The 3rd International Multi-Conference on Engineering and Technological Innovation

June 29th - July 2nd, 2010 – Orlando, Florida, USA

PROCEEDINGS
Volume I

Edited by:

Nagib Callaos
Hsing-Wei Chu
Andrés Tremante
C. Dale Zinn

Organized by
International Institute of Informatics and Systemics
Member of the International Federation for Systems Research (IFSR)

COPYRIGHT

ISBN-13: 978-1-936338-01-6 (Collection)   ISBN-13: 978-1-936338-02-3 (Volume I)

PROGRAM COMMITTEE

Chairs: Hsing-Wei Chu (USA) and C. Dale Zinn (USA)

Adascalitei, Adrian	Technical University	Romania
Aguilar T., Fernando J.	Almeria University	Spain
Ahmad, Ashraf	Princess Sumya University for Technology	Jordan
Al Obaidy, Mohaned	Gulf College	Oman
Al-Aomar, Raid	Jordan University of Science and Technology	Jordan
Alshara, Osama	Higher Colleges of Technology	Uzbekistan
Ariwa, Ezendu	London Metropolitan University	UK
Ayuga T., Esperanza	Universidad Politécnica de Madrid	Spain
Barberà, Elena	Universitat Oberta de Catalunya	Spain
Batos, Vedran	University of Dubrovnik	Croatia
Baykal, Yahya K.	Cankaya University	Turkey
Benbouziane, Mohamed	University of Tlemcen	Algeria
Bernardino, Jorge	Instituto Superior de Engenharia	Portugal
Blankenbach, Karlheinz	Pforzheim University	Germany
Bönke, Dietmar	Reutlingen University	Germany
Brinksmeier, Ekkard	University of Bremen	Germany
Bubnov, Alexey	Institute of Physics	Czech Republic
Calenzo, Patrick	IM2NP	France
Carbon, Margarita	The Boeing Company	USA
Castaño-Moraga, Carlos	Universidad de Las Palmas de Gran Canaria	Spain
Chang, Maiga	TELDAP	Taiwan
Chang, Ruay-Shiung	National Dong Hwa University	Tokelau
Chang, Wen-Kui	Tunghai University	Taiwan
Chen, Yuhua	University of Houston	USA
Chen, Zhi	University of Kentucky	USA
Cheng, Xiaochun	University of Reading	UK
Chinesta, Francisco	ENSAM	France
Chiou, Yin-Wah	Nanhua University	Taiwan
Clarke, Tim	University of Wales	UK
Curran, Kevin	University of Ulster	UK
Davies, B. J.	IJAMT	UK
Davim, J. Paulo	University of Aveiro	Portugal
Dierneder, Stefan	Linz Center of Mechatronics Gmbh	Austria
Dilnutt, Rod	William Bethwey and Associates	Australia
Dodig-Crnkovic, Gordana	Mälardalen University	Sweden
Dudek, Agnieszka	Wroclaw University of Environmental	Poland
Ehmann, Kornel F.	Northwestern University	USA
Erkollar, Alptekin	University of Applied Sciences Wiener Neustadt	Austria
Estrems, Manuel	Universidad Politécnica de Cartagena	Spain
Faura, Felix	Universidad Politécnica de Cartagena	Spain

Franco, Patricio	Universidad Politécnica de Cartagena	Spain
Fu, Yonggang	Shanghai Jiao Tong University	China
Fúster-Sabater, Amparo	CSIC	Spain
Gelbukh, Alexander	National Polytechnic Institute	Mexico
Gibson, Thomas L.	General Motors Research and Development Center	USA
Glovnea, Romeo P.	Brunel University	UK
Gomes, Samuel	UTBM	France
Goriachkin, Oleg	Volga State Academy	Russian Federation
Greenberg, William	Virginia Tech	USA
Grenmyr, Gustav	Chalmers University of Technology	Sweden
Gschweidl, Manfred	University of Applied Sciences Vorarlberg	Austria
Hepdogan, Seyhun	University of Central Florida	USA
Hetzer, Dirk	T-Systems Nova Deutsche Telekom	Germany
Higashiyama, Yoichi	Ehime University	Japan
Hirschberg, Wolfgang	Graz University of Technology	Austria
Holifield, David	University of Wales	UK
Huang, Yo-Ping	Tatung University	Taiwan
Ioannides, Stathis	SFK	Netherlands
Iovan, Stefan	Romanian Railway	Romania
Jiménez R., Lourdes	University of Alcala	Spain
Jingchao, Chen	University Donghua	China
Kalogiannis, Konstantinos	Brunel University	UK
Kaminski, Jacek	Chalmers University of Technology	Sweden
Keiski, Riitta L.	University of Oulu	Finland
Khamba, J. S.	Punjabi University	India
Kobielarz, Magdalena	Wroclaw University of Technology	Poland
Koeglmayr, Hans-Georg	Pforzheim University of Applied Sciences	Germany
Koleva, Maria	Bulgarian Academy of Sciences	Bulgaria
Kruczynski, Klaus	Leipzig University of Applied Sciences	Germany
Kuropka, Piotr	Wroclaw University of Technology	Poland
Lahlouhi, Ammar	University of Biskra	Algeria
Lamo, Yngve	Bergen University College	Norway
Lappas, Georgios	TEI of Western Macedonia	Greece
Li, Longzhuang	Texas A & M University	USA
Li, Man-Sze	IC Focus Ltd.	UK
Litvin, Vladimir	California Institute of Technology	USA
Liu, Jun	University of Ulster	UK
Lloret Mauri, Jaime	Polytechnic University of Valencia	Spain
López Román, Leobardo	University of Sonora	Mexico
Luh, Guan-Chun	Tatung University	Taiwan
Macke, Janaina	Universidade de Caxias do Sul	Brazil
Mansikkamäki, Pauliina	Tampere University of Technology	Finland
Mares, Cristinel	Brunel University	UK
Masoumi, Nasser	University of Tehran	Iran
Masunov, Artëm E.	University of Central Florida	USA
Mbobi, Aime M.	Red Knee Inc.	Canada
Mehrabian, Ali	University of Central Florida	USA
Mikhaylov, Ivan A.	University of Central Florida	USA
Mosley, Pauline	Pace University	USA
Mueller, Christoph Jo.	INCAP GmbH	Germany
Mura, Gianluca	Politecnico di Milano University	Italy

Murugan, Natarajan	Coimbatore Institute of Technology	India
Nagai, Yasuo	Tokyo University of Information Sciences	Japan
Nagar, Atulya K.	Liverpool Hope University	UK
Narasimhan, V. Lakshmi	Western Kentucky University	USA
Newman, Stephen	University of Bath	UK
Nicolescu, Cornel M.	The Royal Institute of Technology	Sweden
Nippa, Markus	Pforzheim University of Applied Sciences	Germany
Oberer, Birgit	University of Klagenfurt	Austria
Ong, Pangleen	University of Kentucky	USA
Orosco, Henry	University of Houston	USA
Ouwerkerk, David B.	General Motors Advanced Technology Center	USA
Ozkul, Tarik	American University of Sharjah	UAE
Paredes, Julia	GGC	USA
Pasquinelli, Melissa	North Carolina State University	USA
Pennington, Richard	Georgia Gwinnett College	USA
Pérez, Carlos A.	Colombian Institute of Petroleum	Colombia
Petit, Frédéric	École Polytechnique de Montréal	Canada
Podaru, Vasile	Academia Tehnica Militara	Romania
Potorac, Alin Dan	University of Suceava	Romania
Praus, Petr	Charles University	Czech Republic
Pursell, David	Georgia Gwinnett College	USA
Rai, Bharatendra K.	University of Massachusetts	USA
Ramessur, Roshan	University of Mauritius	Mauritius
Remondino, Marco	University of Turin	Italy
Revetria, Roberto	Università Degli Studi di Genova	Italy
Rieder, Mathias	University of Applied Sciences Vorarlberg	Austria
Rodger, P. M.	University of Warwick	UK
Rodríguez L., Gloria I.	National University of Colombia	Colombia
Rossbacher, Patrick	Graz University of Technology	Austria
Sahli, Nabil	Ministry of Technologies of Communication	Tunisia
Sala, Nicoletta	Universitá della Svizzera Italiana	Italy
Sarate, João A. R.	CESF	Brazil
Sauder, Deborah	Georgia Gwinnett College	USA
Schaeffer, Donna M.	Marymount University	USA
Schaetter, Alfred	Pforzheim University	Germany
Schumacher, Jens	University for Applied Sciences Vorarlberg	Austria
Shiraishi, Yoshiaki	Nagoya Institute of Technology	Japan
Siddique, Mohammad	Fayetteville State University	USA
Singh, Harwinder	Guru Nanak Dev Engineering College	India
Spearot, James A.	General Motors Research and Development Center	USA
Staretu, Ionel	Transilvania University of Brasov	Romania
Sulema, Yevgeniya	National Technical University of Ukraine	Ukraine
Suomi, Reima	University of Turku	Finland
Suzuki, Junichi	University of Massachusetts	USA
Szotek, Sylwia	Wroclaw University of Technology	Poland
Tchier, Fairouz	King Saud University	USA
Timpte, Candace	Georgia Gwinnett College	USA
Toussaint, Luis	UTBM	France
Trimble, Robert	Indiana University of Pennsylvania	USA
Tsoi, Mai Yin	GGC	USA
Valakevicius, Eimutis	Kaunas University of Technology	Lithuania

Vasinek, Vladimir	Technical Univerzity of Ostrava	Czech Republic
Venkataraman, Satyamurti	AIAMED	India
Vinod, D. S.	Sri Jayachamarajendra College of Engineering	India
Wallner, Daniel	Graz University of Technology	Austria
Wang, Lei	University of Houston	USA
Warwick, Jon	London South Bank University	UK
Whiteley, Rick	Calabash Educational Software	Canada
Yaghmaee, Mohammad H.	Ferdowsi University of Mashhad	Iran
Yanagisawa, Hideaki	Tokuyama College of Technology	Japan
Yingling, Yaroslava	North Carolina State University	USA
Yoon, Changwoo	ETRI	South Korea
Zalewski, Romuald I.	Poznañ University of Economics	Poland
Zaretsky, Esther	Hebrew University	Israel
Zelinka, Tomas	Czech Technical University in Prague	Czech Republic
Zhu, Hui	Soochow University	China
Zobaa, Ahmed	University of Exeter	UK

ADDITIONAL REVIEWERS

Abdelatif, Hassini	Laboratory of Application and Analysis of Radiations	Algeria
Abramchuk, George Heorhi	IEEE	Canada
Acma, Bulent	Anadolu University	Turkey
Agodzo, Sampson Kwaku	Kwame Nkrumah University of Science	Ghana
Aigner, Werner	University of Linz	Austria
Albayrak, Y. Esra	Galatasaray University	Turkey
Alhassan, Mohammad	Purdue University Fort Wayne	USA
Al-Jufout, Saleh	Tafila Technical University	Jordan
Altan, Metin	Anadolu University	Turkey
Álvarez, Francisco	Universidad Autónoma de Aguascalientes	Mexico
Alvarez, Isabel	Autonomous University of Barcelona	Spain
Andreev, Rumen	Bulgarian Academy of Sciences	Bulgaria
Arifin, Achmad	Institute of Technology Sepuluh Nopember	Indonesia
Ariton, Doinita	Danubius University of Galati	Romania
Arrabales Moreno, Raúl	Universidad Carlos III de Madrid	Spain
Aruga, Masahiro	Tokai University	Japan
Ashur, Suleiman	Purdue University Fort Wayne	USA
Assel, Matthias	High Performance Computing Center Stuttgart	Germany
Atem De Carvalho, Rogério	Instituto Federal Fluminense	Brazil
Badescu, V.	Polytechnic University of Bucharest	Romania
Balas, Valentina	Aurel Vlaicu University of Arad	Romania
Barreiro, J.	Universidad de León	Spain
Batovski, Dobri A.	Assumption University of Thailand	Austria
Bayraktar, Emin	SUPMECA	France
Bayraktar, Seyfettin	Yildiz Technical University	Turkey
Bellec, Jacques-H	University College Dublin	Ireland
Benfdila, Arezki	Université Mouloud Mammeri	Algeria
Bernabé, Gregorio	University of Murcia	Spain
Bhuiya, A.	Alberta Electric System Operator	Canada
Bigan, Cristin	Ecological University of Bucharest	Romania
Bilbao, Josu	IKERLAN	Spain
Bilich, Feruccio	University of Brasilia	Brazil
Botella, Federico	Miguel Hernandez University	Spain
Boulanger, Jean-Louis	UTC Heudiasyc	France
Branzila, Marius	Technical University Gh. Asachi	Romania
Burke, Jeffrey	National Pollution Prevention Roundtable	USA
Buzzi, Maria Claudia	CNR	Italy
Camins, Angel Sevilla	Inbionova Biotech S.L.	Spain
Cannavò, Flavio	University of Catania	Italy
Cardoso, Pedro	Universidade do Algarve	Portugal
Carrió Pastor, M Luisa	Universidad Politécnica de Valencia	Spain
Castillo-Guerra, Eduardo	University of New Brunswick	Canada
Cavalini, L. T.	Fluminense Federal University	Brazil

Chahir, Youssef	Université de Caen Basse	France
Challoo, Linda	Texas A & M University	USA
Chehri, Abdellah	Université Laval	Canada
Chen, Haiguang	ShangHai Normal University	China
Chen, Jing-Heng	Feng Chia University	Taiwan
Chen, Ping-Hei	National Taiwan University	Taiwan
Chen, Shi-Jay	National United University	Taiwan
Chen, Yu-Ru	National Cheng Kung University	Taiwan
Chiou , Chuang Chun	Dayeh University	Taiwan
Chis, Monica	SIEMENS	Romania
Chou, Tien-Yin	Feng Chia University	Taiwan
Choudhary, Rahim	Serco North America	USA
Chow, James	University of Toronto	Canada
Cotet, Costel Emil	University Politehnica of Bucharest	Romania
Csáki, Tibor	University of Miskolc	Hungary
Dai, Fengzhi	Tianjin University of Science and Technology	China
Dereli, Türkay	University of Gaziantep	Turkey
Diab, Hassan	Sherbrooke University	Canada
Dijkstra, Jan	Eindhoven University of Technology	Netherlands
Djeffal, Lakhdar	University of Batna	Algeria
Dong, Baoli	Zhejiang University of Science and Technology	China
Doughan, Mahmoud	Lebanese University	Lebanon
Drid, Saïd	University of Batna	Algeria
Dusane, Devendra H.	Institute of Bioinformatics and Biotechnology	India
Eichberger, Arno	Graz University of Technology	Austria
Eisenhauer, William	Portland State University	USA
El Abed, Haikal	Technical University Braunschweig	Germany
Elwany, Hamdy	Alexandria University	Egypt
Enayati, Babak	National Grid	USA
Erives, Hector	NMT	USA
Ersoy, A.	Istanbul Üniversitesi	Turkey
Fathy, Sherif Kassem	King Fasial University	Saudi Arabia
Feldmann, Birgit	University of Hagen	Germany
Fidan, Ismail	Tennessee Tech University	USA
Figueroa, José	IBM	USA
Fiorentino, Michele	Politecnico di Bari	Italy
Fiorini, Rodolfo A.	Politecnico di Milano	Italy
Flammini, Francesco	University Federico II of Naples	Italy
Florea, Adrian	Lucian Blaga University	Romania
Floros, Andreas	Ionian University	Greece
Fryza, Tomas	Brno University of Technology	Czech Republic
Gallucci, Fausto	University of Twente	Netherlands
Gandhi, Meenakshi	Guru Gobind Singh Indraprastha University	India
Ganis, Matthew	Pace University	USA
Garro, Alfredo	University of Calabria	Italy
Georgiev, Georgi Nikolov	University of Veliko Tirnovo	Bulgaria
Gheno, Simoni	Federal University of Sao Carlos	Brazil
Gheorghies, C.	Dunarea de Jos University of Galati	Romania
Ghosh, Sadhana	National Institute of Industrial Engineering	India
Gichev, Dinko	Burgas Free University	Bulgaria
Glossman-Mitnik, Daniel	Grupo Nanocosmos	Mexico
Glowacki, Miroslaw	University of Science and Technology	Poland
Goel, Arun	NIT Kurukshetra	India
Goi, Chai Lee	Curtin University of Technology	Malaysia
Grochocki, Luiz Rodrigo	Pontificia Universidade Catolica do Parana	Brazil
Gu, Fei	Zhejiang University	China

Gujarathi, Ashish M.	Birla Institute of Technology and Science	India
Gunawan, Indra	Auckland University of Technology	New Zealand
Gutiérrez-Torres, C.	Instituto Politécnico Nacional	Mexico
Hadzilias, Elias A.	IESEG	Greece
Halilcevic, Suad	University of Ljubljan	BIH
Hardman, John	Florida Atlantic University	USA
Hasan, Mohamad K.	University of Missouri	USA
Hashmi, Mohammad	Cardiff University	UK
Hassini, Abdelatif	University of Oran Es-Senia	Algeria
He, Hongyu	Louisiana State University	USA
Hennequin, Sophie	LGIPM	France
Hirz, Mario	Graz University of Technology	Austria
Hoeltzener, Brigitte	ENSIETA	France
Hu, Huabin	Chinese Academy of Sciences	China
Huang, Shian-Chang	National Changhua University of Education	Taiwan
Ibhadode, Akii	University of Benin	Nigeria
Ibrahim, Hamidah	Universiti Putra Malaysia	Malaysia
Ibrahim, Rahinah	Universiti Putra Malaysia	Malaysia
Ijioui, Raschid	Rwth Aachen University	Germany
Ingber, Lester	Lester Ingber Research	USA
Isley, Sara L.	Beckman Coulter	USA
James, A. A.	University of KwaZulu-Natal	South Africa
Jaroslav, Heinrich	HBH Projekt Spol. s r.o.	Czech Republic
Jarz, Ewald	University of Applied Sciences Kufstein	Austria
Jastroch, Norbert	MET Communications	Germany
Jiang, Jinlei	Tsinghua University	China
Jimeno, Antonio	University of Alicante	Spain
Jones, Albert	National Institute of Standards and Technology	USA
Kacem, Imed	University Paul Verlaine Metz	France
Kain, Sebastian	Technical University of Munich	Germany
Kakanakov, Nikolay	Technical University of Sofia	Bulgaria
Kamrani, Ehsan	Ecole Polytechnique de Montreal	Canada
Kässi, Tuomo	Lappeenranta University of Technology	Finland
Kavikumar, J.	Universiti Tun Hussein On Malaysia	Malaysia
Kawamura, Hidenori	Hokkaido University	Japan
Ke, Bwo-Ren	National Penghu University	Taiwan
Khudayarov, Bakhtiyar	TIIM	Uzbekistan
Kim, C. G.	Chungnam National University	South Korea
Kim, Hyun-Jun	Samsung Electronics	South Korea
Kochikar, Vivekanand	Infosys Technologies	India
Koh, Min-Sung	Eastern Washington University	USA
Kondylakis, Haridimos	University of Crete	Greece
Kontaxakis, Georgios	Universidad Politécnica de Madrid	Spain
Kouachi, Saïd	Qassim University	Algeria
Ksenofontov, Alexandre	Moscow Engineering Physics Institute	Russian Federation
Kundu, Anirban	Netaji Subhash Engineering College	India
Kureshi, Nadeem	CASE	Pakistan
Kutlu, Önder	Selcuk University	Turkey
Lasmanis, Aivars	University of Latvia	Latvia
Lavric, Vasile	University Politehnica of Bucharest	Romania
Lee, Joo Hwan	Seoul National University	South Korea
Lei, Chi-Un	University of Hong Kong	Hong Kong
Leucci, Giovanni	University of Salento	Italy
Li, Dayong	Shanghai Jiao Tong University	China
Li, Hongyan	Peking University	China
Li, Yu-Chiang	Southern Taiwan University	Taiwan

Lin, Chih-Ting	National Taiwan University	Taiwan
Lin, Chun Yuan	Chang Gung University	Taiwan
Lincke, Susan	University of Wisconsin	USA
Liu, Tingyang Lewis	National Kaohsiung Normal University	Taiwan
Liu, Yajun	South China University of Technology	China
Lopez de Lacalle, Luis	University of the Basque Country	Spain
Lopez, Jorge	University of Texas	USA
Louta, Malamati	Technological Educational Institute of Western Macedonia	Greece
Luqman, Chuah A.	Universiti Putra Malaysia	Malaysia
Mahdoum, Ali	Centre de Developpement des Technologies Avancees	Algeria
Mahlia, T. M. Indra	University of Malaya	Malaysia
Mahmoudi, Saïd	Faculté Polytechnique de Mons	Belgium
Maldonado, J. L.	Centro de Investigaciones en Óptica A.C.	Mexico
Marietto, Marcio	Faculdade de Tecnologia de Tatuí	Brazil
Markakis, Euaggelos	Technological Educational Institute of Crete	Greece
Marlowe, Thomas	Seton Hall University	USA
Marzouk, Osama	National Energy Technology	USA
Masotti, Andrea	Sapienza University of Rome	Italy
Mattila, Anssi	Laurea University of Applied Sciences	Finland
McCall, John	Robert Gordon University	UK
McCormick, John	Institution of Engineering and Technology	UK
Medina, Dulce	Universidad Autónoma Metropolitana	Mexico
Mierlus-Mazilu, Ion	Technical University of Civil Engineering	Romania
Minea, A. A.	Technical University Gh. Asachi	Romania
Mitra, Debojyoti	Sir Padampat Singhania University	India
Mokhtari, A.	UAV	France
Morana, Giovanni	Catania University	Italy
Morel, Eneas N.	UNT	Argentina
Morini, Mirko	University of Ferrara	Italy
Moschim, Edson	State University of Campinas	Brazil
Motieifar, Alireza	University of British Columbia	Canada
Murakami, Akira	Kyoto University	Japan
Naddeo, A.	Università di Salermo	Italy
Nagamalai, Dhinaharan	Woosong University	South Korea
Nasrullayev, Nazim	Baku State University	Azerbaijan
Navas, Luis Manuel	University of Valladolid	Spain
Nayak, P. K.	Tata Institute of Fundamental Research	India
Neaga, Elena Iirna	Loughborough University	UK
Neves, Luís F. F.	University of Oklahoma	USA
Nikolaos, Bilalis	Technical University of Crete	Greece
Nisar, Humaira	Gwangju Institute of Science and Technology	South Korea
O`Shaughnessy, Douglas	INRS-EMT	Canada
Omidvar, Hedayat	National Iranian Gas Company	Iran
Ong, Pang-Leen	Emitech Inc	USA
Onur Hocaoglu, Fatih	Anadolu University	Turkey
Otto, Tauno	Tallinn University of Technology	Estonia
Oueslati, Walid	Laboratoire de Physique des Matériaux Lamellaires	Tunisia
Pedamallu, Chandra Sekhar	New England Biolabs Inc	USA
Pfliegl, Reinhard	AustriaTech Ltd	Austria
Pieters, Cornelis (Kees)	University for Humanistics	Netherlands
Pinto Ferreira, Eduarda	Instituto Superior de Engenharia do Porto	Portugal
Plouffe, B.	Northeastern University	USA
Pogarèiæ, Ivan	Polytechnic of Rijeka	Croatia
Pormann, John B.	Duke University	USA
Poursaberi, A.	University of Tehran	Iran
Prabuwono, Anton Satria	Universiti Kebangsaan Malaysia	Malaysia

Price, Howard	Equipment Design	USA
Puslecki, Zdzislaw	Adam Mickiewicz University	Poland
Pyatt, Kevin	Eastern Washington University	USA
Radneantu, Nicoleta	American University	Romania
Raibulet, Claudia	University of Milano	Italy
Rajagopalan, Mathrubutham	Anna University Chennai India	India
Rana, Mukhtar Masood	Anglia Ruskin University	UK
Ribakov, Y.	Ariel University Center of Samaria	Israel
Riesbeck, Christopher	Northwestern University	USA
Rocha, Rui	University of Coimbra	Portugal
Rodrigues, Jose Alberto	Instituto Superior de Engenharia de Lisboa	Portugal
Rodriguez-Florido, M. A.	Canary Islands Institute of Technology	Spain
Romanov, Sergey	Pavlov Institute of Physiology	Russian Federation
Romdhane, L. B.	University of Sherbrooke	Tunisia
Ros, L.	Universidad Politécnica de Cartagena	Spain
Rot, Artur	Wroclaw University of Economics	Poland
Rydhagen, Birgitta	Blekinge Institute of Technology	Sweden
Saglam, Necdet	Anadolu University	Turkey
Sahin, Omer Sinan	Selcuk University	Turkey
Salay Naderi, Mohammad	Tavanir Holding Company	Iran
Sánchez , Caio	University of Campinas	Brazil
Sane, Vijay	American Chemical Society	India
Sanin, Cesar	The University of Newcastle	Australia
Santos Filho, Esmeraldo	Universite Pierre et Marie Curie	France
Santulli, C.	Universita di Roma	Italy
Schlue, Benjamin	ARCADIS Consult GmbH	Germany
Schmidt, Jon A.	Burns and McDonnell	USA
Sen, Taner Z.	US Department of Agriculture	USA
Seppänen, Marko	Tampere University of Technology	Finland
Serodio, Carlos M. J. A.	UTAD University	Portugal
Shannaq, Boumedyen	Nizwa University	Oman
Shikoski, J.	A. D. EMO	Macedonia
Siricharoen, Waralak V.	University of the Thai Chamber of Commerce	Thailand
Sirjani, Mojtaba B.	Norfolk State University	USA
Sittidech S., Punnee	Naresuan University	Thailand
Skawinska, Eulalia	Poznan University of Technology	Poland
Skoko, Hazbo	Charles Sturt University	Australia
Sllame, Azeddien M.	Al Fateh University	LBY
Sosa, Horacio	College of Professional	USA
Spence, Kelley L.	North Carolina State University	USA
Spina, Edison	Escola Politecnica da Universidade de Sao Paulo	Brazil
Šraml, Matjaž	University of Maribor	Slovenia
Srovnal, Vilem	Technical University of Ostrava	Czech Republic
Stematiu, Dan	Technical University of Civil Engineering Bucharest	Romania
Su, J. L.	Tongji University	China
Süral, Haldun	METU	Turkey
Sutherland, Trudy	Vaal University of Technology	South Africa
Sutikno, Tole	Universitas Ahmad Dahlan	Indonesia
Tamimi, Abdelrahman	Palestinian Hydrology Group	Israel
Tarng, J. H.	National Chiao Tung University	Taiwan
Taylor, Craig	University of Southern California	USA
Tian, Y. B.	Singapore Institute of Manufacturing Technology	Singapore
Ting, Kwun-lon	Tenn Tech University	USA
Tobar, Carlos Miguel	Pontifical Catholic University of Campinas	Brazil
Torras, Eulalia	UOC	Spain
Travieso-González, Carlos M.	University of Las Palmas	Spain

Trifonova, Silvia	University of National and World Economy	Bulgaria
Tseng, Juin-Ling	Minghsin University of Science and Technology	Taiwan
Tsytsarev, Vassiliy	Washington University	USA
Turcu, Cristina	SRAIT	Romania
Ulusoy, A. H.	Eastern Mediterranean University	Turkey
Unold, Jacek	Wroclaw University of Economics	Poland
Vaganova, Natalia	Institute of Computational Mathematics	USA
Van Swet, Jacqueline	Fontys OSO	Netherlands
Vance, James	The University of Virginia`s College	USA
Veeraklaew, Tawiwat	Chulachomklao Royal Military Academy	Thailand
Verstichel, Stijn	Ghent University	Belgium
Vidal-Naquet, Guy	SUPELEC	France
Vimarlund, Vivian	Linköpings Universitet	Sweden
Vintere, Anna	Latvia University of Agriculture	Latvia
Virvilaite, Regina	Kaunas University of Technology	Lithuania
Vizureanu, Petrica	Technical University Iasi	Romania
Vukadinovic, Dinko	University of Split	Croatia
Wang, Chao	Tianjin University	China
Wegner, Tadeusz	Poznan University of Technology	Poland
Wei, Xunkai	Beijing Aeronautical Technology Research Center	China
Wen, Fuhliang	St. John's University	Taiwan
Wen, Guihua	South China University of Technology	China
Wolfengagen, Viacheslav	Educational and Consulting Center JurInfoR	Russian Federation
Woo, Dae-Gon	Yonsei University	South Korea
Xiaoguang, Rui	University of Science and Technology of China	China
Yang, Lei	Tongji University	China
Yindi, Zhao	China University of Mining and Technology	China
Younis, Adel	University of Victoria	Canada
Yusof, Yusri	University Tun Hussein Onn Malaysia	Malaysia
Zacharewicz, Greg	Université Bordeaux 1	France
Zainal Abidin, Azizan	Universiti Teknologi Petronas	Malaysia
Zakrzewska, Danuta	Technical University of Lodz	Poland
Zalewski, Romuald	Poznan University of Economics	Poland
Zamarreño, C. R.	Public University of Navarra	Spain
Zampieri, Douglas E.	UNICAMP	Brazil
Zeiner, Herwig	Joanneum Research	Austria
Zeman, Klaus	Johannes Kepler University Linz	Austria
Zeng, Zhigang	Wuhan University of Technology	China

ADDITIONAL REVIEWERS FOR THE NON-BLIND REVIEWING

Abo El Magd, Mohamed	Saud University	Saudi Arabia
Abramchuk, George Heorhi	IEEE	Canada
Ahuja, I. P. Singh	Punjabi University	India
Albayrak, Y. Esra	Galatasaray University	Turkey
Al-Fuqaha, Ala	Western Michigan University	USA
Ali, Adel	University of Minnesota	USA
Al-Marzouqi, Mohamed	UAEU	UAE
Alonso López, Javier	Universitat Politecnica de Catalunya	Spain
Altan, Metin	Anadolu University	Turkey
Altemose, Brent	Sabre Safety	USA
Alves de Morais, Soraya	Universidade Estadual da Paraiba	Brazil
Al-Zuhair, Sulaiman	UAE University	UAE
Annunziata, Francesco	University of Cagliari	Italy
Arunkumar, Thangavelu	Vellore Institute of Technology	India
B., Ravishankar	BMSCE	India
Badran, Omar	Applied University	Jordan
Báez, Elsa	Universidad Autónoma Metropolitana	Mexico
Bagherzadeh, Nader	University of California	USA
Barreiro, J.	Universidad de León	Spain
Basson, Henri	Laboratoire Informatique du Littoral	France
Baugh, Joseph	University of Phoenix	USA
Bellamy, Al	Eastern Michigan University	USA
Bochniarz, Zbigniew	University of Minneapolis	USA
Boniolo, Giovanni	European Institute of Oncology	Italy
Borja-Medina, Julio C.	Tecnológico de Monterrey	Mexico
Bouslimani, Yassine	University of Moncton	Canada
Brake, Mary	Eastern Michigan University	USA
Breen-Smyth, Marie	Aberystwyth University	UK
Bubnov, Alexey	Institute of Physics	Czech Republic
Bugarin, Eusebio	Instituto Tecnológico de Ensenada	Mexico
Zhang, C.	The University of Western Ontario	Canada
Casolo, Federico	Politecnico di Milano	Italy
Castier, Marcelo	UAE University	UAE
Cavallucci, Denis	INSA	France
Chang, Shun-Chang	Dayeh University	Taiwan
Chate, Andris	Riga Technical University	Latvia
Ciurana, Quim	University Girona	Spain
Cohen, Jerry	Burns and Levinson	USA
Coni, Mauro	University of Cagliari	Italy

Costea, Carmen	Academy of Economic Studies Bucharest	Romania
Dai, Fengzhi	Tianjin University of Science and Technology	China
Dalto, Edson José	Ibmec Business School	Brazil
Del Pozo, Dionisio	Robotiker	Spain
Dereli, Türkay	University of Gaziantep	Turkey
Deutsch, Johnathan	The CUNY Graduate Center	USA
Dhar, Vijay	National Institute of Urban Affairs	India
Dijkstra, Jan	Eindhoven University of Technology	Netherlands
Djeffal, Lakhdar	University of Hadj Lakhdar Batna Algeria	Algeria
Doma, Salah	Alexandria University	Egypt
Dreschhoff, Gisela	University of Kansas	USA
El Abed, Haikal	Technical University Braunschweig	Germany
El-Naas, Muftah	UAE University	UAE
Ersoy, Sezgin	Marmara Universitesi	Turkey
Fales, Roger	University of Missouri	USA
Feng, Zhihua	Soochow University	China
Filgueiras de A. F., Petrônio	Universidade Federal da Paraiba	Brazil
Fitzgerald, Sue	Metropolitan State University	USA
Foley, Michael	Aberystwyth University	UK
Fray, Derek	University of Cambridge	UK
Fuerte-Esquivel, Claudio	Universidad Michoacana de San Nicolás de Hidalgo	Mexico
Fusco, José Paulo Alves	Universidade Paulista	Brazil
Galetto, Maurizio	Polytechnic University of Torino	Italy
Gallucci, Fausto	University of Twente	Netherlands
Gandhi, Meenakshi	Guru Gobind Singh Indraprastha University	India
Gao, Wensheng	Tsinghua University	China
Gaytán-Iniestra, Juan	Tecnológico de Monterrey	Mexico
Ghare, A. D.	National Institute of Technology	India
Ghasem, Nayef	UAEU	UAE
Gibson, David	The Association of Building Engineers	UK
Gichev, Dinko	Burgas Free University	Bulgaria
Girot, Franck	ENSAM Bodeaux	Spain
Godart, Claude	ESSTIN	France
Goel, Arun	National Institute of Technology	India
Gómez, Jorge Marx	Universität Oldenburg	Germany
Grewal, Mohinder S.	California State University	USA
Guo, Shu-Mei	National Cheng Kung University	Taiwan
Hamidian, Karim	California State University	USA
Han, Qingkai	Northeastern University	China
Harris, Thomas	Vanderbilt University	USA
Hasan, Mohamad K.	Kuwait University	Kuwait
Hashemi, Hassan	California State University	USA
Hasselbring, Wilhelm	Christian-Albrechts-Universität	Germany
Hemachandran, K.	Assam University	India
Hernández, Roberto	Universidad Autónoma Metropolitana	Mexico
Herrera, Enrique Jaime	CIATEJ	Mexico
Hessel, Fabiano	PUC-RS	Brazil
Holzer, Thomas	National Geospatial-Intelligence Agency	USA
Hung, Yi-Hsuan	Normal University Taipei	Taiwan
Jackson, Richard	Reader in International Politics	UK
Jansons, Juris	Institute of Polymer Mechanics	Latvia

Jastroch, Norbert	MET Communications	Germany
Jiang, Peixue	Tsinghua University	China
João Mansur, Webe	UFRJ	Brazil
Kadowaki, Makoto	UNICAMP	Brazil
Kahraman, Cengiz	Istanbul Technical University	Turkey
Kaiser, Stephan	Ingolstadt School of Management	Germany
Kalita, U. C.	RGI	India
Kalra, Prem Kumar	Indian Institute of Technology Kanpur	India
Kamrani, Ehsan	Ecole Polytechnique de Montreal	Canada
Kekäle, Tauno	University of Vaasa	Finland
Kelly, Daniel	The University of Adelaide	Australia
Kerns, Edward	Lafayette College	USA
Khusainov, Denis	Kiev National University	Australia
Kim, Soon-Chul	Electronics and Telecommunications Research Institute	South Korea
Kimbrell, Scott	Georgia Institute of Technology	USA
Kokowski, Michal	Polish Academy of Sciences	Poland
Kozak, Drazan	Mechanical Engineering Faculty	Croatia
Kumar, Rakesh	S. V. National Institute of Technology	India
Kundu, Anirban	Netaji Subhash Engineering College	India
La Prad, Jim	Western Illinois University	USA
Lai, Weng Kin	MIMOS Berhad	Malaysia
Lawal, Ganiyu Ishola	University of Lagos	Nigeria
Lee, Jonathan	National Central University	Taiwan
Lengerke Perez, Omar	UFRJ	Brazil
Leon, Juan	Instituto de Física Fundamental	Spain
Li, Shutao	Hunan University	China
Li, Tsu-Shan	National Cheng Kung University	Taiwan
Li, Xiaohua	Florida International University	USA
Lin, Chih-Ting	National Taiwan University	Taiwan
Lin , Chun-Cheng	National Kaohsiung University of Applied Sciences	Taiwan
Liu, Bo	Oak Ridge National Laboratory	USA
Liu, Junhua	Xi'an Jiaotong University	China
Liu, Weidong	Tsinghua University	China
Livingston, Debra	University of the Sunshine Coast	Australia
Lukasz, Miroslaw	Wroslaw University of Technology	Poland
Lukianowicz, Czeslaw	Politechnika Koszalinska	Poland
Ma, Hui	Northeastern University	China
Maginnis, Jean	Maine Center for Creativity	USA
Mahmoudi, Saïd	Faculté Polytechnique de Mons	Belgium
Makwana, AjayJ	S.V. National Institute of Technology	India
Marcondes, Carlos	Federal Fluminense University	Brazil
Marcos, Mariano	University of Cadiz	Spain
Marietto, Marcio	Faculdade de Tecnologia de Tatuí	Brazil
Marlowe, Thomas	Seton Hall University	USA
Marsavina, Liviu	Politehnica University of Timisoara	Romania
Mathur, Mukesh	National Institute of Urban Affairs	India
Matousek, Vaclav	West Bohemian University Pilsen	Czech Republic
Maughmer, Mark	Pennsylvania State University	USA
Medina, Dulce	UAM Azcapotzalco	Mexico
Melo, Carlos	Tribunal de Contas do Estado de Minas Gerais	Brazil
Middleton, Howard	Griffith University	Australia

Morales, Luis Javier	Universidad Veracruzana	Mexico
Moreno, Javier	Centro de Investigación y Desarrollo de Tecnología Digital	Mexico
Motieifar, Alireza	University of British Columbia	Canada
Muñoz, Daniel	UAM	Mexico
Nagy, Gehan	Higher Technological Institute	Egypt
Nakoua, Mohamed	UAEU	UAE
Narayanan, S.	Wright State Research Institute	USA
Nastenko, Yevgen	Amosov Institute of Cardiovascular Surgery	Ukraine
Neaga, Elena Iirna	Loughborough University	UK
Nenovsky, Nikolay	University of National	Bulgaria
Nikl, Jiri	University of Hradec Kralove	Czech Republic
Nisar, Humaira	Gwangju Institute of Science and Technology	South Korea
Nolazco Flores, Juan A.	ITESM	Mexico
Nonato S. S., Raimundo	IBMEC Business School	Brazil
Ogata, Craig	Argonne National Laboratory	USA
Ojolo, Sunday	University of Lagos	Nigeria
Ostwald, Marian	Poznan University of Technology	Poland
Otto, Tauno	Tallinn University of Technology	Estonia
P., Balamuralidhar	Tata Consultancy Services	India
Pal, Arpan	Tata Consultancy Services Limited	India
Parekh, Askok	S. V. National Institute of Technology	India
Passoni, Lucía	Universidad Nacional de Mar del Plata	Argentina
Patterson, Rob	Simon Fraser University	Canada
Pelikan, Emil	Academy of Sciences of the Czech Republic	Czech Republic
Pellom, Bryan	University of Colorado	USA
Peon Escalante, Ignacio	Instituto Politecnico Nacional	Mexico
Pillai, Harish	Indian Institute of Technology Bombay	India
Pinto Ferreira, Eduarda	Instituto Superior de Engenharia do Porto	Portugal
Prasad, Rajender	Indian Institute of Technology Roorkee	India
Puslecki, Zdzislaw	Adam Mickiewicz University	Poland
Qu, Fanqi	Wuhan University	China
Qureshi, Suhail A.	UET	Oman
Ragoub, Lakhdar	Alyamamh University	Saudi Arabia
Rahimi, Masoud	Razi University	Iran
Ramos-Paz, Antonio	Universidad Michoacana de San Nicolás de Hidalgo	Mexico
RAWIDEAN, Mohamed	MIMOS	Malaysia
Ray, Daniel	The University of Virginia`s College	USA
Ray, Pradip Kumar	IIT Khargpur	India
Ren, Lixia	Department of Chemistry and Biochemistry	USA
Rivera-Colon, Nilo	LiveTimeNet Inc.	USA
Robledo, Carlos Walter	National University of Cordoba	Argentina
Rodrigues, José	Universidade Federal Fluminense	Brazil
Roman, Mihai	Academy of Economic Studies Bucharest	Romania
Romanovsky, Boris	MSU	Russian Federation
Sadoway, Donald	Massachusetts Institute of Technology	USA
Saini, Jaswinder Singh	Thapar University	India
Samyudia, Yudi	Curtin University	Malaysia
Sanin, Cesar	The University of Newcastle	Australia
Saqib, Asghar	University of Engineering and Technology	Pakistan
Schumack, Mark	University of Detroit Mercy	USA
Schürmann, Volker	University of Applied Sciences Bochum	Germany

See, Kye Yak	Nanyang Technological University	Singapore
Seppänen, Marko	Tampere University of Technology	Finland
Shannaq, Boumedyen	Nizwa University	Oman
Shen, Yu-Lin	University of New Mexico	USA
Shokry, Hanaa	Higher Technological Institute	Egypt
Siddique, Mohammad	Fayetteville State University	USA
Singh, Diwan	NIT Kurukshetra	India
Singh, Rupinder	GNDEC	India
Sitou, Wassiou	Technische Universität München	Germany
Smith, Marilyn	Georgia Institute of Technology	USA
Soares Filho, Djalma	Petrobras	Brazil
Song, Fugen	Donghua University	China
Su, J. L.	Tongji University	China
Suell Dutra, Max	UFRJ	Brazil
Suratgar, Amir	Arak State University	Iran
Sutherland, Trudy	Vaal University of Technology	South Africa
Sutikno , Tole	Universitas Ahmad Dahlan	Indonesia
Swauger, Lane	KCI Technologies	USA
Takala, Josu	University of Vaasa	Finland
Tang, Yu	University of Electronic Science and Technology	China
Tavakolian, Kouhyar	Simon Fraser University	Canada
Taveira R., Railda S.	Universidade Estadual da Paraiba	Brazil
Thomas, Valerie	Georgia Institute of Technology	USA
Thompson, Andrew	Solail Synchrotron	France
Ting, Kwun-Lon	Tennessee Tech University	USA
Tomczyk, Andrzej	Rzeszow University of Technology	Poland
Tomoko, Hirayama	Doshisha University	Japan
Tosic, Bratislav	UNS	Russian Federation
Tu, Kuo-Yang	NKFUST	Taiwan
Turut, Abdulmecit	Ataturk University	Turkey
Valerio, Angel	ITESM	Mexico
Verma, Nischal K.	Indian Institute of Technology	India
Vite, Manuel	Instituto Politécnico Nacional	Mexico
Von Eye, Alexander	Michigan State University	USA
Vybiral, Bohumil	University of Hradec Kralove	Czech Republic
Wang, Chao	Tianjin University	China
Wang, Chuanyang	Soochow University	China
Wei, Li	Chongqing Jiaotong University	China
Wen, Fuhliang	St. John's University	Taiwan
Weng, George J.	Rutgers University	USA
Wolfengagen, Viacheslav	Educational and Consulting Center JurInfoR	Russian Federation
Xiaoqiang, Wan	Chongqing Jiaotong University	China
Xin, Ying	Innopower Superconductor Cable	China
Xu, Baoming	Beijing Jiao Tong University	China
Xue, Dongfeng	Dalian University	China
Yin, Runyuan	Shanghai Ocean University	China
Yu, Donggang	The University of Newcastle	Australia
Zainal Abidin, Azizan	Universiti Teknologi Petronas	Malaysia
Zakaria, Maamar	Zayed University	UAE
Zalewski, Romuald	Poznan University of Economics	Poland
Zambon, Renato	USP	Brazil

Zhang, Bailing	Xi'an Jiaotong-Liverpool University	China
Zhang, Guibing	China University of Geosciences	China
Zhou, Bing	University of Adelaide	Australia
Zhu, Qingxin	UESTC	China
Zogahib, André Luiz N.	State University of Amazonas	Brazil
Zuhair, Sulaiman	UAEU	UAE

HONORARY CHAIR
William Lesso

PROGRAM COMMITTEE CHAIRS
Hsing-Wei Chu
C. Dale Zinn

GENERAL CHAIR
Nagib Callaos

ORGANIZING COMMITTEE CHAIRS
Belkis Sánchez
Andrés Tremante

CONFERENCE PROGRAM MANAGER /
HARDCOPY PROCEEDINGS PRODUCTION CHAIR
Maria Sánchez

TECHNICAL CONSULTANT ON COMPUTING SYSTEM /
CD PROCEEDINGS PRODUCTION CHAIR
Juan Manuel Pineda

SUBMISSION QUALITY CONTROL SUPPORT
Leonardo Contreras

META REVIEWERS SUPPORT
Maria Sánchez
Dalia Sánchez

SYSTEM DEVELOPMENT, MAINTENANCE AND DEPLOYMENT
Dalia Sánchez
Keyla Guédez
Nidimar Diaz
Yosmelin Marquez

OPERATIONAL ASSISTANTS
Marcela Briceño
Cindi Padilla

HELP DESK
Riad Callaos
Louis Barnes
Katerim Cardona
Arlein Viloria
Pedro Martínez

The 2nd International Symposium on
Energy Engineering, Economics and Policy: EEEP 2010
in the context of
The 3rd International Multi-Conference on Engineering and Technological Innovation: IMETI 2010

GENERAL CHAIRS
Andrés Tremante
Nagib Callaos (IMETI)

ORGANIZING COMMITTEE CHAIRS
Jorge Baralt
Friedich Welsch
Belkis Sánchez

PROGRAM COMMITTEE **Chairs:** Andrés Tremante (Venezuela)
Nagib Callaos (Venezuela)

Akinnikawe, Oyewande	Texas A and M University	USA
Anchliya, Abhishek	Texas A and M University	USA
Arena, Umberto	Second University of Naples	Italy
Baruah, Debendra Chandra	Tezpur University	India
Basso, Giuliano	IEEE	Belgium
Brezet, J. C.	Delft University of Technology	Netherlands
Carrasquero, José Vicente	Universidad Simón Bolívar	Venezuela
Cázares-Rangel, Víctor M.	UANL	Mexico
Cerny, Vaclav	University of West Bohemia	Czech Republic
Cha, Seung Tae	Korea Electric Power Research Institute	South Korea
Ehlig-Economides, Christine	Texas A and M University	USA
Elkamel, Ali	University of Waterloo	Canada
Fukase, Masa-Aki	Hirosaki University	Japan
Gustavsson, Rune	Blekinge Institute of Technology	Sweden
Hansen, Martin Otto L.	Technical University of Denmark	Denmark
Kan, S. Y.	Delft University of Technology	Netherlands
Kim, Yong Hak	Korea Electric Power Research Institute	South Korea
Klapp, Jaime	Instituto Nacional de Investigaciones Nucleares	Mexico
Lefevre, Thierry	Center for Energy Environment Resources Development	Thailand
Mastellone, Maria Laura	Second University of Naples	Italy
Melioli, Riccardo	Università Degli Studi di Genova	Italy
Naddeo, Alessandro	Università di Salermo	Italy
Platt, Glenn	CSIRO Energy Technology	Australia
Rahman, Anuar Abdul	Pusat Tenaga Malaysia	Malaysia
Revetria, Roberto	Università Degli Studi di Genova	Italy
Riaz Moghal, Mohammad	University College of Engineering and Technology	Pakistan
Shin, Jeong Hoon	Korea Electric Power Research Institute	South Korea
Tam, Wing K.	Swinburne University of Technology	Australia
Velte, Clara Marika	Technical University of Denmark	Denmark
Zaccariello, Lucio	Second University of Naples	Italy
Zobaa, Ahmed	Cairo University	UK

ADDITIONAL REVIEWERS

Al-Ammar, Essam	King Saud University	Saudi Arabia
Barzev, Kiril	University of Rousse	Bulgaria
Bode, Sven	Arrhenius Institute for Energy and Climate Policy	Germany
Bruna, Elena	Yale University	USA
De Benedictis, Michele	Politecnico di Bari	Italy
Duta, Anca	Transilvania University of Brasov	Romania
Ferrando, Emanuele	Selex Galileo	Italy
Hammad, Mahmoud	GNREADER	Jordan
Herreros, José Martín	Universidad de Castilla - La Mancha.	Spain
Hiser, Eric	Jorden Bischoff	USA
Jayaraj, Simon	NIT Calicut	India
Khatib, Hisham	World Energy Council	Jordan
Kulp, W. David	Georgia Institute of Technology	USA
Lu, Shuai	Pacific Northwest National Laboratory	USA
Ma, Jian	Pacific Northwest National Laboratory	USA
Mansour, Mohy	Cairo University	Egypt
Muensit, Nantakan	Prince of Songkla University	Thailand
Muensit, Supasarote	Prince of Songkla University	Thailand
Owen, Alan	Robert Gordon University	UK
Pimentel, David	Cornell University	USA
Popescu, Mihai O.	Politehnica University Bucharest	Romania
Reich, Nils	Utrecht University	Netherlands
Sampara, Chaitanya	Nanostellar Inc.	USA
Schenk, Peer	The University of Queensland	Germany
Shi, Yu	University of Wisconsin-Madison	USA
Singh, Kaushlendra	West Virginia University	USA
Sparber, Wolfram	Eurac Research	Italy
Van Dyk, Ernest	Nelson Mandela Metropolitan University	South Africa
Vijay, Virendra Kumar	Indian Institute of Technology Delhi	India
Ward Jr., Marvin	Center for Public Finance Research	USA
Yang, Qiujing	The Pennsylvania State University	USA
Zhao, Jinquan	Hohai University	China

ADDITIONAL REVIEWERS FOR THE NON-BLIND REVIEWING

Alizadeh, Behnam	Iran University of Science and Technology	Iran
Augutis, Juozas	Lithuanian Energy Institute	Italy
Borges, Amadeu	University of Trás-os-Montes and Alto Douro	Portugal
Burtraw, Dallas	Resources for the Future	USA
Carvajal Mariscal, Ignacio	Instituto Politécnico Nacional	Mexico
Casini, Dante	University of Pisa	Israel
Clarke, Joe	University of Strathclyde	UK
Denholm, Paul	National Renewable Energy Laboratory	USA
Elia, Nicola	Iowa State University	USA
Ferrando, Emanuele	Selex Galileo	Italy
Graditi, Giorgio	ENEA	Italy

Grimoni, José	USP	Brazil
Han, Aijie	University of Texas	USA
Hanks, Dallas	Utah State University	USA
Kalantar, Mohsen	Iran University of Science and Technology	Iran
Khatib, Hisham	World Energy Council	Jordan
Kristolaitis, Ricardas	Lithuanian Energy Institute	Italy
Kukushkin, Nikolai	Russian Academy of Sciences	Russian Federation
Lin, Feng	Wayne State University	USA
Liu, Yongwen Liu	Shanghai Jiao Tong University	China
McCalley, James	Iowa State University	USA
McCann, Richard	Aspen Environmental Group	USA
Mesarovic, Miodrag	Energoprojekt Consulting Engineers Co.	Serbia
Mishra, Sukumar	Indian Institute of Technology Delhi	India
Neyestani, Nilufar	Iran University of Science and Technology	Iran
Polupan, Giorgiy	Instituto Politécnico Nacional	Mexico
Potter, Cameron	IEEE	Australia
Rizzo, Gianfranco	Università degli Studi di Salerno	Italy
Sharma, M P	Indian Institute of Technology	India
Shayanfar, Heidarali	Iran University of Science and Technology	Iran
Shayeghi, Hossein	Iran University of Science and Technology	Iran
Starr, Andrew	University of Hertfordshire	UK
Van Dyk, Ernest	Nelson Mandela Metropolitan University	South Africa
Wang, Fei	Zhejiang University	China
Zheng, Jie	University of Akron	USA
Zobaa, Ahmed	Cairo University	UK

Number of Papers Included in these Proceedings per Country

(The country of the first author was the one taken into account for these statistics)

Country	# Papers	%
TOTAL	**126**	**100.00**
United States	38	30.16
Japan	8	6.35
Brazil	7	5.56
Italy	7	5.56
Mexico	7	5.56
China	6	4.76
Australia	5	3.97
Canada	5	3.97
Germany	4	3.17
India	4	3.17
Hong Kong	3	2.38
Spain	3	2.38
Argentina	2	1.59
Israel	2	1.59
Taiwan	2	1.59
United Kingdom	2	1.59
Austria	1	0.79
Colombia	1	0.79
Czech Republic	1	0.79
Egypt	1	0.79
Estonia	1	0.79
France	1	0.79
Ireland	1	0.79
Latvia	1	0.79
Lithuania	1	0.79
Malaysia	1	0.79
Netherlands	1	0.79
New Zealand	1	0.79
Poland	1	0.79
Romania	1	0.79
Serbia	1	0.79
Singapore	1	0.79
Slovakia	1	0.79
South Africa	1	0.79
South Korea	1	0.79
Ukraine	1	0.79
United Arab Emirates	1	0.79

Foreword

Engineering activities are based on the development of new Knowledge (*Scientia*), new 'made things' (*Techné*), and/or new ways of working and doing (*Praxis*). *Scientia*, *Techné,* and *Praxis* are three important dimensions of a comprehensive conception of Engineering as a whole. Engineering, as *Scientia*, is mostly developed in academia; as *Techné,* is practiced in industry generating technological innovations; and as *Praxis,* is carried out in technical and non-technical organizations, supporting managerial activities and technical procedures, via methodical and methodological design and implementation. This is why Engineering provides one of the most solid academic and professional substrata for bridging among universities, industries and governments.

Publications and conferences related to Engineering are usually oriented to one of its three dimensions. While this is an adequate thing to do when disciplinary focus is sought, it does not represent Engineering as a whole and it misses the very important synergic relationships among the three kinds of engineering activities mentioned above. This is why a group of scholars, professionals, and consultants, in the field of engineering, considered the possibility of organizing a conference where presentations would not be reduced to one specific Engineering dimension, but would foster the participation of academics, practitioners, and managers in the three dimensions of Engineering, in the same conference, so they can synergistically interact with each other. A consequence of this purpose is the organization of IMETI 2010, where submissions were accepted for the presentation of:

- **New knowledge** (Engineering as *scientia*);
- **New products and services**, i.e. technological innovations (Engineering as *techné*);
- **New technical and managerial methods and methodologies** (Engineering as *praxis*);
- **New meta-engineering** (Engineering of Engineering activities) knowledge, innovations, and methodologies.

The 7th International Conference on Cybernetics and Information Technologies, Systems and Applications (CITSA 2010) and The 8th International Conference on Computing, Communications and Control Technologies (CCCT 2010) have been organized in the context of IMETI 2010, because both are mainly oriented to Engineering and Technology. Both of them are International Multi-Conferences organized with the purpose of providing a communicational forum to researchers, engineers, practitioners, developers, consultants, and end-users of computerized, communications, and/or control systems and technologies in the private and the public sectors. This multi-disciplinary forum provides the opportunity to share experience and knowledge by facilitating discussions on current and future research and innovation. Participants can explore the implications of relationships between new developments and their applications to organizations and society at-large.

One of the primary objectives of CITSA 2010, CCCT 2010 and, in general, IMETI 2010 is to promote and encourage interdisciplinary cross-fertilization and knowledge communication. They encourage systemic thinking and practice, including the analogical thinking that characterizes the Systems Approach, which is, in most cases, the required path to logical thinking, scientific hypothesis formulation, and new design and innovation in engineering.

CITSA 2010 and CCCT 2010 are spin-offs from the International Conference on Information Systems, Analysis and Synthesis (ISAS), and the World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI) which are yearly events that have been held in the last 15 years as a forum for Information Systems researchers, practitioners, consultants, and users who have been interchanging ideas, research results, and innovations in the area of Information Systems. Analytical as well as synthetical thinking represent the conceptual and methodological infrastructures that support the papers presented in ISAS conferences. Synthetical thinking supported papers in the Information Systems area, as well as in its relationships (analogies, "epistemic things", "technical synthetical objects", hybrid systems, cross-fertilization, etc.) with other areas.
The Organizing Committees of IMETI/CITSA/CCCT 2010 invited authors to submit original works, analogy-based hypothesis, innovations, experience-based reflections and concepts, specific problems requiring solutions, case studies, and position papers that explore the relationships among the disciplines of computers, communications and control, and the social and industrial applications within these fields.

On behalf of the Organizing Committee, I extend our heartfelt thanks to:

1. the 625 members of the Program Committee from 63 countries;
2. the 673 additional reviewers, from 80 countries, for their **double-blind peer reviews**;
3. the 451 reviewers, from 58 countries, for their efforts in making the **non-blind peer reviews**. (Some reviewers supported both: non-blind and double-blind reviewing for different submissions)

A total of 2480 reviews made by 1124 reviewers (who made at least one review) contributed to the quality achieved in IMETI 2010. This means an average of 5.84 reviews per submission (425 submissions were received). Each registered author could get information about: 1) the average of the reviewers' evaluations according to 8 criteria, and the average of a global evaluation of his/her submission; and 2) the comments and constructive feedback made by the reviewers, who recommended the acceptance of his/her submission, so the author would be able to improve the final version of the paper.
In the organizational process of IMETI 2010, including CITSA 2010 and CCCT 2010, about 425 papers/abstracts were submitted. These pre-conference proceedings include about 126 papers, from 36 countries, that were accepted for presentation,. We extend our thanks to the invited sessions organizers for collecting, reviewing, and selecting the papers that will be presented in their respective sessions. The submissions were reviewed as carefully as time permitted; it is expected that most of them will probably appear in a more polished and complete form in scientific journals.

This information about IMETI 2010 is summarized in the following table, along with the other collocated conferences:

Conference	# of submissions received	# of reviewers that made at least one review	# of reviews made	Average of reviews per reviewer	Average of reviews per submission	# of papers included in the proceedings	% of submissions included in the proceedings
WMSCI 2010	**711**	**1841**	**3586**	**1.95**	**5.04**	**211**	**29.68%**
IMETI 2010	**425**	**1124**	**2480**	**2.21**	**5.84**	**126**	**29.65%**
IMSCI 2010	**321**	**720**	**1751**	**2.43**	**5.45**	**121**	**37.69%**
CISCI 2010	**622**	**1174**	**3321**	**2.83**	**5.34**	**194**	**31.19%**
TOTAL	**2079**	**4859**	**11138**	**2.29**	**5.36**	**652**	**31.36%**

We are also grateful to the co-editors of these proceedings for the hard work, energy, and eagerness they displayed preparing their respective sessions. We express our intense gratitude to Professor William Lesso for his wise and opportune tutoring, for his eternal energy, integrity, and continuous support and advice as Honorary President of WMSCI 2010 and its collocated conferences, as well as for being a very caring old friend and intellectual father to many of us. We also extend our gratitude to Professor Belkis Sanchez, who brilliantly managed the organizing process. Special thanks to Dr. C. Dale Zinn for chairing CCCT 2010 Program Committee (PC) and for co-chairing IMETI 2010 PC, to Professor Hsing-Wei Chu for co-chairing the IMETI 2010 PC and being General Co-Chair of CCCT 2010; to Professor Michael Savoie for being Co-General Chair of CCCT 2010 and CITSA 2010; to Professor José Ferrer for chairing the CITSA 2010 Organizing Committee; to professors Andrés Tremante and Belkis Sánchez for co-chairing the IMETI 2010 Organizing committee.

We also extend our gratitude to Drs. W. Curtiss Priest, Louis H. Kauffman, Leonid Perlovsky**,** Stuart A. Umpleby, Eric Dent, Thomas Marlowe**,** Ranulph Glanville, Karl H. Müller, and Shigehiro Hashimoto, for accepting to address the audience of the General Joint Plenary Sessions with keynote conferences, as well as to Drs. Sam Chung, Dr. Susu Nousala, Robert Lingard for accepting our invitation as Keynote Speakers at the Plenary Session of IMETI 2010.

Many thanks to Professors Friedrich Welsch, Thierry Lefevre, José Vicente Carrasquero, Angel Oropeza, and Freddy Malpica for chairing and supporting the organization of the focus symposia and conferences in the context of, or collocated with, IMETI 2010. We also wish to thank all the authors for the quality of their papers.

We extend our gratitude as well to Maria Sanchez, Juan Manuel Pineda, Leonisol Callaos, Dalia Sánchez, Keyla Guedez, Nidimar Díaz, Yosmelin Márquez, Riad Callaos, Marcela Briceño, Pedro Martínez, Louis Barnes, and Katerim Cardona for their knowledgeable effort in supporting the organizational process and for producing the hard copy and CD versions of the proceedings.

Professor Nagib C. Callaos,
IMETI 2010 General Chair

VOLUME I

CONTENTS

Engineering Concepts, Relations and Methodologies

Engineering Education

Engineering Technology Transfer

Pentahedral Honeycomb with Skew Hexagonal Faces

Edgars BERVALDS
Latvian Academy of Sciences, Akademijas laukums 1-205,
Riga, LV-1050, Latvia

Modris DOBELIS
Department of Computer Aided Engineering Graphics, Riga Technical University,
Azenes iela 16/20-439, Riga, LV-1048, Latvia

ABSTRACT

This paper proposes a generally new type of macro and megaspace filling honeycombs having quasiregular pentahedral cells with skew hexagonal faces. An existence of spatial cells named pentahedra is demonstrated by topological transformations of hexagonal prismatic honeycomb and is based on a recently discovered *Phi* relationship within a regular hexagonal tessellation. A geometric symmetry of this honeycomb is studied, filling the space with no gaps or overlaps. Finally it is pointed out that abstract or skeletal analogues of pentahedral honeycomb have effective practical uses by synthesis of artificial man-made macromedium, especially the like of orbital large scale structural systems.

Keywords: Topological Transformations, Skew Faces, Pentahedral Honeycomb

1. INTRODUCTION

Two earlier made discoveries lie in the ground of necessity to carry out these investigations within spatial macro structural geometry.

The first lies within a structural mechanics. In 1993 it was demonstrated that stiffness components of the topological invariants of a spatial bar system are not identical by identically used volume of the material [2]. The numerical value of this characteristic is the most highest of cell like a lattice of diamond, exceeded for example 5.9 times a topological stiffness component of a traditionally used triangular lattice.

It had been discovered in 2007 that vertices of regular hexagonal tessellation are Phi centres with very low variance that corresponds to the 11-th series of the Fibonacci convergence sequence on Phi [3, 4]. It means that an existence of such a rational geometric ratio could secure the highest specific mechanical stiffness exactly of the hexagonal structure. Thereby here is a geometrical problem to construct a honeycomb like a spatial analogue of planar regular hexagonal tiling.

2. TOPOLOGICAL TRANSFORMATIONS OF A HEXAGONAL PRISMATIC HONEYCOMB

We start with building up an infinite horizontal layer of regular hexagonal prisms (Fig. 1). Let a be the length of edges of their base hexagons, and let y be the height of each prism, i.e., the distance between the horizontal base planes (it will be determined below).

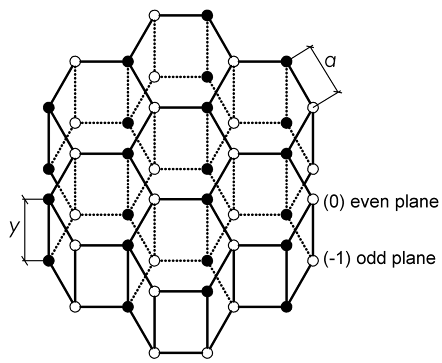


Figure 1: A layer of hexagonal prisms:
(0) and (-1) – even and odd horizontal planes of vertices, respectively; a – length of edges of hexagons;
y – thickness of a layer.

We further make up a hexagonal prismatic honeycomb by arranging layers one above the other so that bases of prisms in neighbouring layers coincide. The horizontal planes between layers are labelled by consecutive integers: …, -4, -3, -2, -1, 0, 1, 2, 3, 4… We also colour vertices of the grid in black and white so as ends of every horizontal edge are coloured differently, while ends of every vertical edge have the same colour. This colouring is used to describe a certain deformation within the layer: (a) we move all black points lying in the planes labelled by even integers (including 0) vertically up by an amount x, and all the white points down by the same amount x (x will be determined later on); (b) we act the other way (black vertices down, white ones up) in the odd planes (Fig. 2).

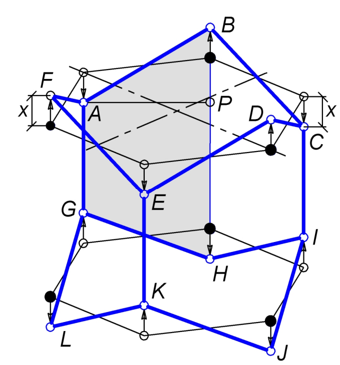


Figure 2: Topologic transformations of hexagonal prism:
x – amount of the vertical displacement of vertices.

During the topological transformation vertical edges of former hexagonal prismatic honeycomb are divided in two groups such that one group consists of those of increased length ($y+2x$) and the other group – of reduced length ($y-2x$). It means that instead of 6 vertical rectangles of each former hexagonal prism we have obtained 3 regular once convex hexagonal faces, but instead of 2 horizontal hexagons, 2 regular twice convex faces. In this way, we obtain a peculiar polyhedron of new kind with five skew hexagonal faces which will be called pentahedron.

We are now going to determine the corresponding values of the parameters x and y. Let us observe one vertex O within obtained pentahedral honeycomb (Fig. 3) connected with vertices J, N, C and H. In this figure, the point I was moved up, and the points J, N, and H were moved down from their common plane, while the point C was moved down from the nearest upper plane. The point I has now only four adjacent vertices (in contrast to five in the former hexagonal prismatic honeycomb). As we want the four edges incident to I to be of equal length, and the six angles between those edges of the same size, the pyramid $JNHC$ has to be regular and the point I has to be its focus.

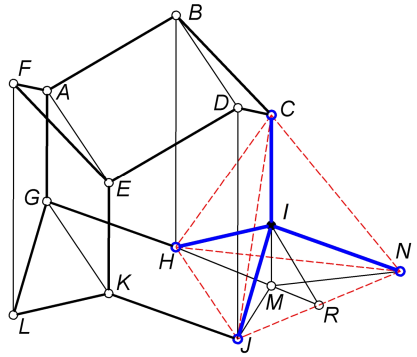


Figure 3: The pentahedral honeycomb:
regular tetrapod circumscribed by
triangular pyramid.

Let us draw a regular triangular pyramid $JNHC$ with focus I, denote the centre of its base triangle by M, and the midpoint of the edge JN by R (Fig. 3). Recall that a was the initial length of edges of hexagons, x was the amount of the vertical shift of vertices, and y was the thickness of a layer (i.e., the length of vertical edges). Therefore $IM = NM = HM = a$, $IM = 2x$, and $MC = y$.

Since the pyramid is regular, the edges of pentahedron incident to I are of equal length, which we denote by z:

$$JI = NI = HI = CI = z\ .$$

The angles between these lines also are equal; let α stand for their size. Then

$$\angle JIN = \angle NIH = \angle HIJ = \angle JIC = \angle NIC = \angle HIC = a\ .$$

Next we will determine the values of the above-defined parameters x and y so that the pyramid $JNHC$ could be regular, indeed. Using Eg. (1) - (8) and performing the calculations, we finally will also obtain the value of z with respect to y and x as well as the size of the angle α:

$$JN = NH = JH = JC = NC = HC = \sqrt{3}a \tag{1}$$

$$y = MC = \sqrt{NC^2 - MN^2} = \sqrt{3a^2 - a^2} = \sqrt{2}\,a \tag{2}$$

$$IM = \frac{1}{4}MC = \frac{1}{4}\sqrt{2}a \tag{3}$$

The focus of a regular pyramid divides its height in the ratio 1:3. Then

$$x = \frac{1}{2}IM = \frac{1}{8}\sqrt{2}a \tag{4}$$

$$IR = \sqrt{IM^2 + MR^2} = \sqrt{\frac{a^2}{8} + \left(\frac{MN}{2}\right)^2} = \sqrt{\frac{a^2}{8} + \frac{a^2}{4}} = \sqrt{\frac{3}{8}}a \tag{5}$$

$$NR = \frac{1}{2}JN = \frac{\sqrt{3}}{2}a \tag{6}$$

$$a = 2\angle RIN = 2\arctan\frac{NR}{IR} = 2\arctan\left(\frac{\sqrt{3}}{2}\cdot\frac{\sqrt{8}}{\sqrt{3}}\right) = \tag{7}$$
$$= 2\arctan\sqrt{2} \approx 109{,}47°$$

$$z = JI = \sqrt{JM^2 + IM^2} = \sqrt{a^2 + \frac{a^2}{8}} = \frac{3}{\sqrt{8}}a = \frac{3}{4}\sqrt{2}\,a \tag{8}$$

Summing up: pentahedral honeycomb is a structure determined by a single parameter z. In order to construct this structure, a hexagonal prismatic honeycomb has to be constructed first. In this latter one the length of an edge of each hexagon is $a = \frac{4}{3\sqrt{2}}z = \frac{2}{3}\sqrt{2}z$, the height of each prism (i.e., the length of its lateral edges) is $y = \sqrt{2a} = \sqrt{2}\frac{2}{3}\sqrt{2}z$.

Then vertices of the prisms have to be shifted by $x = \frac{\sqrt{2}}{8}a = \frac{\sqrt{2}}{8}\cdot\frac{2}{3}\sqrt{2}z = \frac{1}{6}z$. In this way, the faces of prismatic hexagonal honeycomb are transformed into skew hexagonal faces of the pentahedral honeycomb. All the angles between the adjacent edges in the pentahedral honeycomb will be $a = 2\arctan\sqrt{2} \approx 109.47°$.

Let us check one more geometrical relation within a pentahedral cell having skew hexagonal faces. In Fig. 2 depicted is one deformed lateral face $ABHG$ of an initial hexagonal prism. It is an equilateral trapezoid; the sides AG, GH, and AB are edges of a cell in the newly constructed pentahedral honeycomb, $\angle AGH = \angle GAB = a = 2\arctan\sqrt{2} \approx 109.47°$ and AP is the height of the trapezoid. We should to check that $BP = 2x = \frac{1}{3}z$.

Indeed,

$$BP = AB\cdot\sin\angle PAB = AB\cdot\sin(\angle GAB - \frac{p}{2}) =$$
$$= z\cdot\sin(2\arctan\sqrt{2} - \frac{p}{2}) \approx 0{,}333333\cdot z \approx \frac{1}{3}z.$$

3. SYMMETRY OF THE PENTAHEDRAL HONEYCOMB

Symmetry, being the most inherent property of pentahedrons and spatial pentahedral tessellations, has been studied both on the level of a separate cell and of a honeycomb as a whole by means of symmetry groups (Fig. 4).

The pentahedron has 12 vertices, 15 edges, and 5 faces – the skew hexagons $ABCDEF$, $GHIJKL$, $ABCIHG$, $EDCIJK$, and $AFEKLG$. In a pentahedral honeycomb, every vertex is incident to 4 equal edges, and the angle between any two of them is $a = 2 \cdot \arctan\left(\sqrt{2}\right) \approx 109{,}47°$.

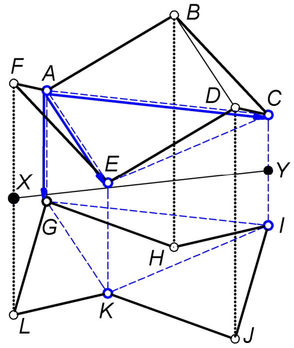


Figure 4: On a symmetry of pentahedral cell: .... bend lines of skew hexagonal faces with midpoints *X* and *Y*; _ _ _ _ _ _ triangular prism, ⟶ vectors of the symmetry.

3.1 Polyhedral Symmetry Group

The symmetry group of a pentahedron, i.e. the group of orthogonal transformations of a space, that map the figure onto itself, is isomorphic to the symmetry group D_3 of a regular triangular prism. This follows from the fact that every orthogonal transformation which maps a pentahedron onto itself is completely determined by an orthogonal transformation which maps onto itself the regular prism *ACEGIK*. The following transformations are the generators of the symmetry group of a pentahedron:

a) Rotation through 120° on the central vertical axis of the pentahedron,
b) Symmetry about the central horizontal plane,
c) Symmetry about the line *XY*.

3.2 Symmetry of the Honeycomb

There are three independent vectors the translation along which maps a honeycomb onto itself. These translations are the base elements of the translation group of the honeycomb which is isomorphic to the free Abelian group $Z^3 = Z \oplus Z \oplus Z$. More specifically, the base consists of

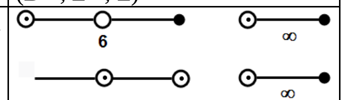
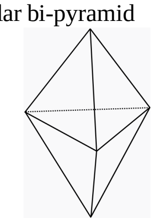
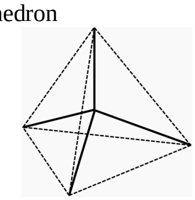
a) The vertical translation along the vector $\overrightarrow{AG} + \overrightarrow{FL} = \overrightarrow{AG} + \frac{5}{3}\overrightarrow{AG} = \frac{8}{3}\overrightarrow{AG}$,
b) The translation in the horizontal plane along the vector $\overrightarrow{AC}$,
c) The translation in the horizontal plane along the vector $\overrightarrow{AE}$.

The (full) symmetry group of the honeycomb is generated by

a) Generators of D_3,
b) Generators of Z^3, and
c) The rotation of the honeycomb through 60° on the vertical axis of a pentahedral cell.

An infinite pentahedral tessellation includes space tunnels structures with skew hexagonal faces. They have six-fold spiral symmetry, for polyhedrons packed along vertical axis repeat both after a rotation through 60° and a translation by a distance equal to the length of each edge. The main characteristics of symmetry properties of a pentahedral honeycomb are given in Table 1.

Table 1. The main characteristics of symmetry properties of a pentahedral honeycomb.

No	Characterization	Explanation
1.	Type	Convex uniform honeycomb
2.	Family	skew hexagon faces polyhedra
3.	Cell type	{6.4}
4.	Face type	{6}
5.	Schläfly symbol	{6.6}
6.	Coxeter group	(D^3, Z^3, Z)
7.	Coxeter – Dunkin diagram	
8.	Cells / edge	{4.3}4
9.	Faces / edge	63
10.	Cells / vertex	{4.3}
11.	Faces / vertex	66
12.	Edges vertex	6
13.	Dual	triangular bi-pyramid 
14.	Vertex figure	tetrahedron 
15.	Internal angle	109.47°
16.	Symmetry group	D3
17.	Other properties	isogonal and isotoxal polyhedra, n=6 fold helical symmetry of Z tunnel

4. ON A SKELETAL APPROACH TO THE PENTAHEDRAL HONEYCOMB

The long standing challenge of designing and constructing new crystalline solid state materials from molecular building blocks had been successfully started [6]. This success concerning a reticular synthesis (or chemistry) of robust materials with highly porous frameworks and with predetermined chemical properties had been achieved by investigation of abstract (skeletal) micro and nanostructural polyhedra.

This same conceptual approach could be employed for the creation of non-default macro and megastructural skeletal systems by predetermining such mechanical properties like minimum mass or maximum stiffness. Branko Grünbaum made a special study of abstract polyhedra, in which he developed an early idea. He defined a face as a cyclically ordered set of vertices, and allowed faces to be skew as well as planar [5]. Moreover in modern computer graphics any polyhedron gives rise to a graph or skeleton with corresponding vertices and edges.

It has been proved in a Chapter 2 that macrospatial pentahedral honeycomb has the same skeletal graph or skeleton as a nanospatial lonsdaleite or hexagonal diamond, which vertices are like junctions of tetrapod shape. It allows creating structural building constructions like hybrid bar systems. Thereby tetrapod shape junctions of bars or finite superelements of this system must be shells without moments of deflection like the most effective macro and megaconstruction from the point of view of used volume of material [1].

5. CONCLUSIONS

1. The pentahedral honeycomb has been obtained by topological transformations (stretching) of a regular prismatic hexagonal tessellation and is quasi-regular (vertex and edge transitive) with cells having three regular once convex and two regular twice convex faces.
2. Pentahedral honeycomb is a third type of discrete symmetry groups equipped with a topology or an infinite space group which combines elements of both point groups and lattice groups and also include such an extra transformation like screw axis.
3. Pentahedra are homeomorphic to hexagonal prisms uniform polyhedra consisting of regular skew hexagonal faces and congruent vertices. So they have exactly the same size and shape and are second space tiling polyhedra after cubic one, tilling space without holes and overlaps.
4. It could foresee a widely use of pentahedral lattices for a synthesis of minimum mass and maximum stiffest large scale structural especially orbital systems or builds on Earth natural satellites.

ACKNOWLEDGMENTS

Authors acknowledge financial support from the Latvian Council of Science grant No. 09.1526/230 "Investigations of geometrical, topological and mechanical properties within spatial analogues of hexagons".

REFERENCES

[1] E. Bervalds, *Hybrid Bar Constructions of Steerable Mirror Antennae.* **Papers of URSI Riga Meeting**, Vol. 1, September 4-6, 1990, 125-133.
[2] E. Bervalds, *Topological Transformations and Design of Structural Systems*. **Proc. of the World Congress on Optimal Design of Structural Systems**. Rio de Janeiro, Brazil, Vol. 1, August 2-6, 1993, pp. 153-160.
[3] E. Bervalds, *Existence of the Phi relationship within a regular hexagonal tessellation.* **Latvian J. of Physics and Techn. Sciences**, 2007, No. 2, pp.57-58.
[4] E. Bervalds, M. Dobelis, *Geometric Properties of a Regular Lattice caused by Phi Relationship*. **Proc. of the 13th International Conference on Geometry and Graphics**. Dresden, Germany, August 4-8, 2008, 8 p.
[5] B. Grünbaum, *Are Your Polyhedra the same as My Polyhedra.* **Discrete and Computational Geometry**. The Goodman – Pollack Festschrift, eds. Springer, New York, 2003, pp. 461-488.
[6] O.M. Yaghi, M. O'Keeffe, N.W. Ockwing, , H.K. Chae, M. Eddaoudi, J. Kim, *Reticular synthesis and the design of new materials*. **Nature**, (423), 12, 2003, pp. 705-713.

DNA Separation by Directed Optical Transport

Avital Braiman,[1] Fedor Rudakov,[2] and Thomas Thundat[3]
[1] Division of Engineering, Brown University, Providence, RI, 02912, USA
[2] Computer Science and Mathematics Division
[3]Biosciences Division
Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA

A wide variety of biological and medical applications require separation of biomolecules of a particular size from a mixture. Usually separation is achieved by electrophoretic transport of charged biomolecules through a sieving media with a uniform electric field. Separation of biomolecules using a uniform electric field has serious drawbacks, the most significant one being that in order to separate biomoelcules of a particular size, complete separation of the sample by size is required. Furthermore, since biomolecules are subject to diffusion, separation is accompanied by band broadening.

We present an alternative design for biomolecule separation that permits selective separation of biomolecules from a mixture (complete separation of the sample by size is not required)[1]. In addition, our design permits independent translocation of biomolecules of different sizes along two dimensional pathways while keeping the biomolecules concentrated. Separation is achieved by optically directed transport of biomolecules through a sieving media.

In our design, a laser beam is projected onto a photoanode that is in contact with an electrolyte, thus creating a highly localized electrical field trap. Charged biomolecules within the electrolyte medium migrate toward the center of the photoelectrophoretic trap. Moving the focus of the laser beam along the photoelectrode, translocates the trap and consequently results in migration of the trapped biomolecules within the medium. When the photoelectrode is in contact with a sieving media such as gel, the ability of biomolecules to follow the trap is size dependent. For a sample consisting of biomolecules of different sizes the speed of the photoelectrophoretic trap can be selected such that the sample splits in two molecular packets. One molecular packet contains biomolecules with mobilities higher than the minimal mobility required for following the trap, while biomolecules in the other molecular packet have lower mobilities and therefore cannot keep up with the trap. After the trap has moved a significant distance from its initial position, the biomolecules outside of the trap do not experience any substantial electric field and consequentially remain close to their initial position. Thus, in order to separate biomolecules of a particular size from the mixture, the velocity of the trap should be selected in such a way that only the biomolecules of interest follow the trap while the next largest in size biomolecules would be incapable of keeping up with the trap. The speed of the photoelectrophoretic trap should be subsequently increased so that only the biomolecules of interest drop out of the trap while the other biomolecules would keep translocating with the trap.

We demonstrated experimentally concentration and separation of a sample into two concentrated molecular packets consisting of DNA fragments of different sizes. Photoconcentration and separation was performed in 1.5% agarose gel. The diameter of the photoconcentrated sample was 470 μm, corresponding to only ~5% of the original sample area. In addition, we performed Monte-Carlo simulations of the DNA separation. We solved the Langevin equation for a DNA particle subject to diffusion and the electric field and obtained qualitative agreement between our computational and experimental results. We believe that well-established electrophoretic techniques can be integrated with photoelectrophoretic transport to gain more control on the separation process and achieve higher resolution.

---

[1] A. Braiman, F. Rudakov, and T. Thundat, "Highly selective separation of DNA fragments using optically directed transport" **Applied Physics Letters**, Vol. 96, No. 5, 2010, pp. 53701/1-4.

A Hybrid Approach for Malay Text Summarizer

Norshuhani ZAMIN, Arina GHANI
Department of Computer and Information Sciences
Universiti Teknologi PETRONAS
Bandar Seri Iskandar, 31750 Tronoh, Perak, MALAYSIA.

ABSTRACT

Summarization is the art of generating the main points of a lengthy text document by removing redundant and less important information without losing the meaning of the original text. Summaries are significantly shorter than the original text and take a broad overview of the source material. With the increasing volume of digital information today, people find the manual process of summarization as hectic and time consuming. Having an automated text summarization system for electronic documents would very much help to encourage people to read, giving quick access to information thus helping them to a faster decision making process. Although many research and commercial text summarization tools are available but no research is officially reported for Malay language. Malay text summarizers are coming into demand when a lot of information in Malay language can now be accessed freely via the Internet. This paper presents a hybrid approach to an automated text summarization system for Malay language. The base system is built on SUMMARIST system and is expanded by combining with EstSum system. Experimental results show that expanding training data size significantly contributes to the performance. In general, our system produced acceptable results at the best case of 76% and the worst case of 31%.

Keywords: Malay Text Summarization, Statistical Approach, Text Mining, Natural Language Processing.

1. INTRODUCTION

Based on the survey of Malaysian reading profile in 1996 done by the Malaysia National Library, the average of reading activity for Malaysian is approximately two books in a year. Further survey made in 2005 on 60,441 Malaysian reported that the 'problem' has not shown significant improvement. Malaysia is far behind among most of well-develop countries in reading activities. One of the major causes to this problem is the reading habits develop very slowly in low income family as compared to higher income family. The mushrooming of the digital Malay texts in the Internet motivates us to develop an automated summarization system to encourage the reading habits of Malaysian by consuming less time to read lengthy documents.

The earliest research in text summarization was done in 1960s and the growth of the interest in this research continues in recent years. Most of the work found is for English text summarization but with the increasing demand of this tool for other languages, a number of research and development found for Estonian [1] Scandinavian [2], Thailand [3], Persian [4], Swedish [5] and the five in one system known as SUMMARIST [6] for summarizing English, Japanese, Arabic, Spanish, Indonesian and Korean text.

A summary is categorized into two types: 1) *indicative* and 2) *informative* summary [7]. Indicative summary is a summary that highlights only the topic of the text while informative summary describes the central information in the text. As this research is concerned with the synopsis of the text, the results shall produce informative summary. However, a development of text summarization tool for Malay language which has a totally different grammatical structure to English is not only the challenge but also the accuracy of the results will be the important issue to be discussed. How good a summary is depending on the percentage of essence preserved in the summary and cohesion between one sentences to another. As it is difficult to measure the quality of the Malay summaries with no baseline research in existence, therefore, the outcome of this research will be compared with the analysis of human-made summaries. Thus, the objective of this research is to develop a Malay text summarizer giving the accuracy of at least 60% resemblances with the manual summaries. The development and evaluation considers summarization of various types of documents such as news articles, magazines, reports and story books.

2. RELATED WORKS

In this section we briefly present some of the research literature related to our work.

Text Mining

Text mining refers to computational methods to discover previously unknown meaningful information from unstructured text. Text mining is closely related to data mining – finding interesting patterns and trends in large dataset. The only difference is text mining deals with natural language text while data mining requires structured databases or facts. The purpose of text mining is to link together the extracted information to form new facts or new hypotheses to be explored further [8]. In recent years, research in text mining covers diverse areas which include term association discovery, document clustering, text summarization and text categorization.

Text mining consists of three basic steps: 1) *Text Preparation* – the preprocessing of text to extract meaningful terms or features 2) *Text Processing* – the use of computational methods to identify interesting patterns in preprocessed text 3) *Text Analysis* – the evaluation of extracted output [9].

Text Summarization

Text summarization is "the process of distilling the most important information from a text and to produce an abridged version for a particular task and user" [23]. Automatic text summarization refers to the use of computational methods to automatically derive the summary of a given text. Over the past half a century, text summarization research has been explored by the Natural Language Processing (NLP) community. The increasing availability of online information has necessitated intensive research in this area.

There are two main methods to automatic text summarization, *abstraction* and *extraction* [10]. Abstraction is a difficult technique yet promising where it generates new sentences from the original sentences through a process called paraphrasing. This technique involves syntactic and semantic study for the particular language and is useful for meaningful applications. On the other hand, the extraction method has been the current state of the art and commonly used by most of the existing tools. This method weighted each sentence in the original text with some specified characteristics and selects the original sentence and juxtapose in the summary. The basic of extraction-based summarization where each sentence is measured through special predefined properties, selecting the most relevant sentences based from the value of the properties and put them together in a summary. This research learnt the basic extraction based technique from [11].

A research in [12] categorized the approaches of the automatic text summarization into three: 1) *Shallow Approach* – the simplest approach of all approaches where a summary is produced by extracting sentences from information source. However, the challenge in this approach is to preserve the original context when breaking the sentences 2) *Deeper Approach* – this approach produces summary called 'abstract' where some of the text in summary may not be found in the original text. It finds the most specific generalization of concepts from texts and uses this for the summary and 3) *Hybrid Approach* – a combination of extraction techniques with Natural Language Processing techniques. This research proposed a hybrid approach using the morphological and part of speech tagging methods from SUMMARIST [6] and the statistical scoring methods from EstSum [1].

Among the aims of text summarization are the single document and multiple documents summarization [23]. In the single document summarization, a summary that characterized the content of a single document is produced. Whilst, the multiple document summarization takes a group of documents as input and a condensation of the content of the entire group is produced as the summary. Multiple-document summarization has turned out to be much more complex than summarizing a single document. This research is focusing on the single document summarization.

Malay Text Summarizer

Malay is not only a native language for Malaysia but also one of the languages used in Indonesia, Brunei, Singapore, and southern Thailand. The Malay language is rich in colloquial, idiomatic expressions and literary allusions and like other languages, it possess its own unique structure and grammar. As the Malay language is used within the South East Asia region, it has become one of the less resourced languages in the world. Due to this, limited number of computational linguistic research was found related to Malay language. Although there are many studies related to Malay language, however, to the best of our knowledge, none has been officially reported on Malay text summarization. However, other computational linguistic studies for Malay language exist such as the Information Retrieval [13], Essay Marking [14], Novelty Detection [15] and Machine Translation [16]. Recently, an open source tool for Malay language corpus analysis was found in an ongoing research [17]. The tool provides access to Malay tokenisers, lemmatisers and part-of-speech taggers that are vital for the Malay linguistic research.

Other Text Summarizers

SUMMONS [19] is the first example of multi-documents summarization system. It summarizes news articles about terrorism from different news agencies and produces a briefing merging relevant information about each identified event. SUMMONS architecture consists of two major components: 1) *Content Planner* – selects the important information from the combination of input templates with instantiated slots of predefined semantics and 2) *Linguistic Generator* – selects the right words to express the information in grammatical and coherent text.

An automated text summarization system called EstSum [1] is able to summarize newspaper articles in Estonian language. It constructs short summaries of text by selecting the key sentences that describes the document. The sentences are classified using a weighted combination of statistical, linguistic and typographic aspects such as the position, format and type of sentence, and the frequency of each word appeared in the system. It achieves up to 60% accuracy on the evaluation done against the human-made summaries or newspaper articles.

SUMMARIST [6] aims at generating both *abstracts* and *extracts* for arbitrary English and other languages texts. In this research *extract* is defined as "portions extracted verbatim of the original (they may be single words or whole passages)" and *abstract* as "novel phrasings describing content of the original (which might be paraphrases or fully synthesized text)". SUMMARIST combines statistical techniques and symbolic word knowledge derived from WordNet [18] – a large lexical database of English. It's technique lies on the following 'equation':

$$\textit{summary} = \textit{topic identification} + \textit{interpretation} + \textit{generation} \quad (1)$$

The purpose of *topic identification* is to filter the input which retains only the most important central topics using various techniques such as stereotypical text structure, cue words, high frequency indicator phrases and discourse structure. *Interpretation* processes the topics, rephrases and compresses them. This process is vital to achieve further compaction and to remove redundancies, rephrase sentences and to merge related topics into more general one. *Generation* process aims to reformulate the interpreted data into new text. The SUMMARIST architecture is illustrated in Figure 1.

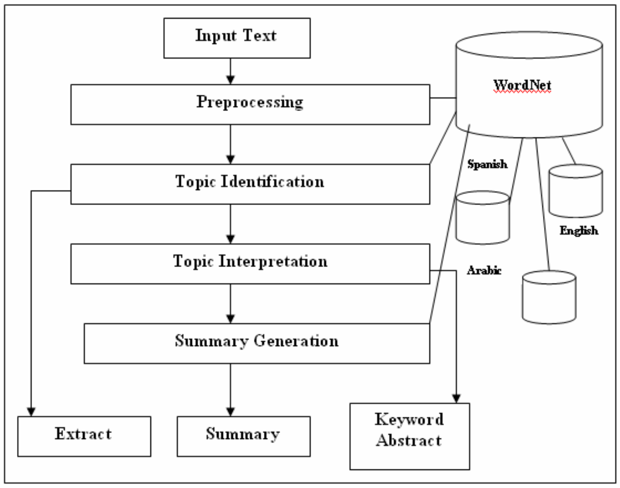


Figure 1. SUMMARIST architecture

SweSum [2] is a Swedish text summarizer. The sentences are extracted based on a combination of linguistic, statistical and heuristic methods. SweSum works in three different passes: 1) *Tokenization, Scoring and Keyword Extraction* – the input text is split into sentences. Word boundaries are identified searching for periods, exclamation and question marks. The sentences are then scored by using statistical, linguistic and heuristic methods. 2) *Sentence Ranking* - The score of each word in the sentence is calculated by a set of parameters, which can be adjusted by the user, and total score is accumulated. Sentences containing common content words get higher scores and 3) *Summary Extraction* – the final summary file is created in HTML format. These processes are schematically represented in Figure 2. The lexicon is a database consists of key / value pairs where the key is the inflected word and the value is the stem / root of the word in Swedish.

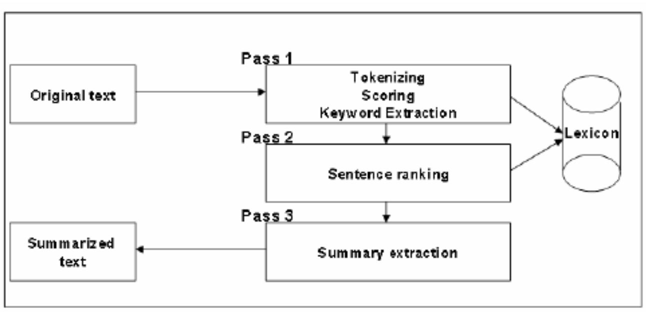


Figure 2. SweSum architecture

FarsiSum [20] is a text summarizer for Persian built based on SweSum modules. The system is implemented as a HTTP client/server application. The FarsiSum's tokenization module uses Persian's stop-list in Unicode format and a small set of heuristic rules. The stop-list is a file including the most common verbs, pronouns, adverbs, conjunctions, prepositions and articles in Persian. Figure 3 depicts the FarsiSum architecture with the each of the summarization steps numbered accordingly. The system is located on the server side and the client is a browser. The summarization steps are described as follows:

Step 1: The browser sends a summarization request to the Web server where FarsiSum is located. The URL of the document to be summarized is attached to the request where the original text is in Unicode format.

Step 2-5: The document is summarized in three phases similar to SweSum. The words in the document are compared with the words in the Persian stop-list.

Step 6: The summary is returned to the HTTP server. The browser then renders the summarized text to the screen.

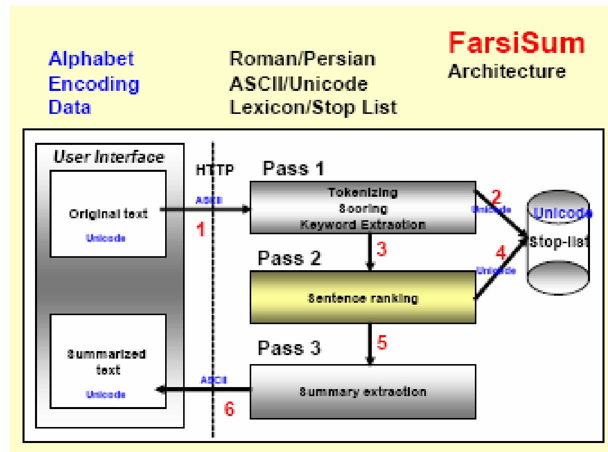


Figure 3. FarsiSum architecture

3. PROPOSED APPROACH

In this research, we choose to work with extraction method. We believe that with further refinement to the formula that will be discussed in this section, we can maximize the retention of the important information in the Malay text. We have considered adopting some techniques from existing successful research. Basically we divide the whole summarization process into three phases:
1) *Preprocessing* 2) *Text Extraction* and 3) *Sentence Selection.*

In the *Preprocessing* phase, we used the technique introduced in SUMMARIST [6]. The preprocessing algorithm considers only two modules in SUMMARIST: 1) *Tokenizer* and 2) *Token Frequency Counter*. There are two forms of *Tokenizer* which are word tokenizer and sentence tokenizer. Word tokenizer chunks each and every word in the input text to produce a set of tokenized text. The boundary of each word is determined based on the white space found between words. While sentence tokenizer chunks the text into sentences by taking the full stop (.) as the boundary between sentences. The tokenizer algorithm for Malay text is developed to recognize Malay words.

In the *Token Frequency Counter* module, the number of occurrence of each word appears in the original text is counted. The highest frequency will be considered as the keyword of the text. However, for some of the commonly used words like articles – "the (English) = itu (Malay)" and "is (English) = ialah (Malay)" will be ignored. From this rank of word frequency, we select the first 10 words with the highest number of occurrence. Referring to these selected words, all sentences containing any of these words will be merged together as a preprocessed text. The original text is now simplified based on the frequency score of words. In the *Text Extraction* phase, we applied the Edmundson's statistical formula [21, 22] shown below and our main reference is the recent EstSum research [1]:

$$\mathbf{W(s)} = \frac{\boldsymbol{\alpha}\mathbf{P(s)} + \boldsymbol{\beta}\mathbf{F(s)} + \boldsymbol{\gamma}\mathbf{K(s)}}{\mathbf{P(s)} + \mathbf{F(s)} + \mathbf{K(s)}} \qquad (2)$$

Where,
W(s) – weight function of sentence **s**;
P(s) – position-based score function;
F(s) – format-based score function;
K(s) – keyword-based score function;
α, β, and γ are constants.

The α, β, and γ that have been previously adjusted by hand using a manually created training corpus, act as the tuning parameters. A combination module evaluates (2) to give each sentence a unique weight. The scores are computed based on three properties identified by experts: 1) *Position* – the score given to the location of the sentence. The regularities in text structures of many genres are helpful to rank the sentences based on their location in the text. For example, the first sentence appears in the text tend to contain important information and thus, will be given a higher score 2) *Format* – the score given to the font style and format. For example, the word written in bold or italic will be given higher score as it addressed the importance of the word and 3) *Keyword* – the score given to the frequency of word appeared in the text. The expression P(s) + F(s) + K(s) in (2) is a normalization factor. The constant α, β and γ act as the tuning parameters which have been previously adjusted by hand using a manually created training corpus are used to evaluate P(s), F(s) and K(s) respectively. The total score gives a sentence a unique weight. Finally, in the *Sentence Selection* phase, by considering some threshold values, sentences with the higher score will be merged and taken as summary. The overall architecture of our proposed work is shown in Figure 4.

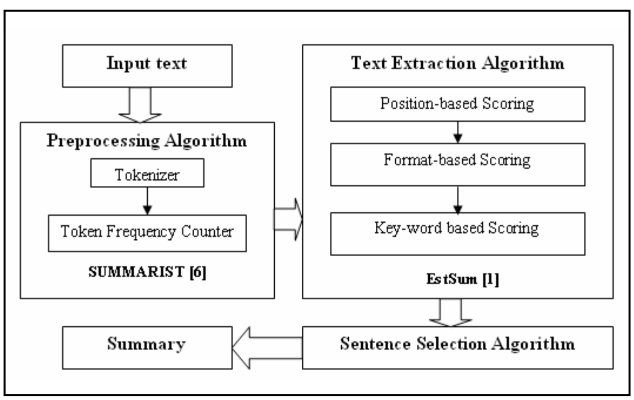


Figure 4. Malay Text Summarizer architecture

4. EXPERIMENTS

This section briefly describes the methodology used in our system evaluation.

Methods

Our experiment requires a training corpus. Due to the non existence of Malay corpus for text summarization, we create our own corpus consists of summaries compiled by four Malay language experts. A total of 10 original Malay news articles covering general, business and sports news were given to each of the four Malay language experts for manual summarization. Each expert submitted 10 summaries limited to 30% of the length of the source text giving 40 hand-created summaries. The length of the source text is varying upon the type of the text. On average, it consists between 800 – 1200 words. The process is illustrated in Figure 5.

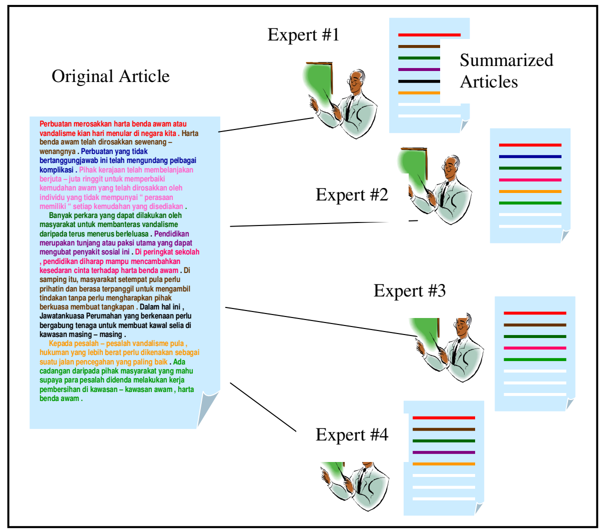


Figure 5. Process of creating the training corpus

Scoring Mechanisms

The baseline of the scoring criteria is obtained from [1] with minor modifications made to suit the different structure of Malay text. As our training corpus of extracts is relatively small (only 40 summaries), we manually examined and compared each of the original text and its summaries. For the *Position-based Scoring*, we assigned the appropriate weight to a sentence by investigating the location of the sentence appeared in the summary using the following rules: 1) The first 3 sentences of the original text 2) The first 3 sentences of each paragraph in the original text and 3) The first 3 sentences after each subtitle in the original text .An example of a Position-based Score for one of the summaries is shown in Table 1.

Table 1.An Example of a Position-based Score

Feature	% in extract	Given score (W(s))
1st sentence in article	100	10
2nd sentence in article	77	8
3rd sentence in article	45	5
1st sentence in paragraph	70	7
2nd sentence in paragraph	38	4
3rd sentence in paragraph	58	6
1st sentence after subtitle	40	3
2nd sentence after subtitle	35	3
3rd sentence after subtitle	26	2

In addition to the *Position-based Scoring*, we defined a hypothesis that a paraphrased sentence found in an expert's summary will be treated as a multi-sentence during the manual investigation. Paraphrasing allows putting together multiple sentences in author's own word which is critical in natural language. This has been a challenge for text summarization research over a decade ago [24]. However, a research in [25] is found to be a promising start for automatically generating sentence paraphrases. At this development stage, we do not consider any paraphrasing in our system's generated summary.

For the *Format-based Scoring*, we considered the sentence based on the style of font (default, bold, italic) and punctuation marks (exclamation marks, question marks, double quotes). Unlike in [1], we excluded the score of the figure captions and text author as they were not present in our training data. Whilst for the *Keyword-based Scoring*, we used a general Malay word frequency table that is generated from all the 10 original texts by our *Token Frequency Counter* module. This helps to estimate whether the word appears more frequently in original text than it normally does in the summarized text. This scoring takes the following rules: 1) The words belonging to the title (article headline) and subtitles are given higher scores and 2) All the other words are given similar lower weight.

Evaluation Metrics

The comparison of our proposed system's generated summaries is done against the human experts' summaries due to the non existence of commercially available Malay text summarizer. A survey in [26] describes and compares various human and automatic metrics to evaluate summaries. We employ the performance measures commonly used in the traditional Natural Language Processing task – *Recall*, *Precision* and *F1 Score*. These scores quantify how close the system's extract to human's. *Precision* shows the accuracy of the extracted sentence, *Recall* reflects how many good sentences the system has missed and *F1 Score* is a weighted average of the *Precision* and *Recall* [27]. Given an input text (original text), human-made summary and system generated summary, the following metrics are applied:

$$\textbf{Precision} = \textbf{correct} / (\textbf{correct} + \textbf{wrong}) \quad (3)$$

$$\textbf{Recall} = \textbf{correct} / (\textbf{correct} + \textbf{missed}) \quad (4)$$

$$\textbf{F1 Score} = 2 \times (\textbf{Precision} \times \textbf{Recall}) / (\textbf{Precision} + \textbf{Recall}) \quad (5)$$

Where,

correct – the number of sentences extracted by the system and the human;

wrong - the number of sentences extracted by the system but NOT by the human;

missed - the number of sentences extracted by the human but NOT by the system.

The generated summary is judged *correct* if it contains sentences that were tagged in the human's summary or *partially correct* if the summary provides sufficient context for the passage. The generated summary is judged *wrong* if needed context was totally misleading or if the summary did not contain the expected passage at all. Finally, the generated summary is judged *wrong* if there is insufficient context for the passage. One standard marks a sentence as in the summary only when all four human experts agree.

5. RESULTS

Table 2 shows the performance of our system. The second, third and fourth row of the table show the statistics of the three test collections. All the summaries agreed at a fixed-length compression rate of 30%.

Table 2. System performance evaluation

Type of Test Collection	*General News*	*Business News*	*Sports News*
No. of Document	4	3	3
Average no. of sentences per document	110	85	94
Average no. of sentences per summary	33	23	26
Average Precision (%)	85.2	77.6	66.3
Average Recall (%)	38.0	22.3	31.0
Average F1 Score (%)	76.0	31.0	42.2

The average number of sentence per summary in *General News* is relatively high in comparison with the other test collections because the number of sentences in the body of documents is higher. Consequently the performance of *Domestic News* is better on average than the other test collections. The reasons, based from the feedbacks provided by the human experts, why Domestic News outperforms others are as following: 1) They are different from each other in genre 2) The *Business News* provides an extremely thorough analysis such as the stock market, foreign exchange market and mutual funds. The technical expressions that are regularly iterated reduced the average score and 3) most summaries generated by the system are hard to understand. 4) The number of the training data is relatively small for drawing any final statistical conclusions. However, further investigation need to be done in future in order to reveal clearer reasons.

Accuracy is any natural language application such as text summarization, machine translation, speech processing is always a big issue. As for text summarization, evaluation is an important aspect to ensure whether the system has reached the goal to resemble the human made summaries. Indeed, naturally that it is hard to find two similar human made summaries for the same language in the world. A study in [28] found that at best there was about 70% average agreement between two human made summaries. On average, our system produces summaries that are about 50% similar to the manually created summaries. Although the agreement between the compute generated summary and the human summaries is quite low but it can be a promising start for a Malay text summarization research.

6. CONCLUSIONS AND FUTURE WORKS

This paper presents a Malay text summarization system using a hybrid approach – the preprocessing module introduced by the SUMMARIST [6] and the statistical scoring methods described in the EstSum's text extraction module [1]. Experiment shows that using the combination of both techniques, the system is able to extract the most important sentence from Malay news articles.

This is a cost-effective solution to reduce users' consuming time in document reading without losing the general issues for users' comprehension. Summary helps users to easily decide its relevancy to their interests and acquire desired documents with less mental loads.

Since the research in this area is still at its immature stage there are many things to be investigated in the future. One of the problems that should be given a highlight is the widespread use of disparate metrics. It is found that there is no standard human or automatic evaluation metric in text summarization to compare different systems and establish a baseline. Hence, in future, to increase the decision accuracy, we plan to conduct the following evaluation as proposed in [29]: 1) *Quantitative Measures* – involve the categorization of decision relevancy, summary time and summary length and 2) *Qualitative Measures*- involve user preferences and detailed feedback as to why the summary was or was not acceptable for a given task.

6. ACKNOWLEDGMENTS

We would like to thank the conference's anonymous reviewers for their valuable comments. We are grateful to Dr. Lai Weng Kin of MIMOS, Malaysia; Professor Dr. Alan Oxley and Dr. Mohd Nordin Zakaria of Universiti Teknologi PETRONAS, Malaysia for their suggestions which helped to improve this work. We are indebted to the four Malay language experts for their efforts to manually summarize given documents in time.

7. REFERENCES

[1] M. Kaili and M. Pilleriin, "EstSum - Estonian Newspaper Text Summarizer", **In the Proceedings of the 2nd Baltic Conference on Human Language Technologies**, 2005.

[2] D. Hercules, "A Text Summarizer for Swedish", **In Technical Report TRITA-NA-P0015**, 2000.

[3] C. Jaruskulchai and C. Kruengkrai, "A Practical Text Summarizer by Paragraph Extraction for Thai", **In the Proceedings of the 6th International Workshop on Information Retrieval with Asian Languages**, Vol. 11, 2003, pp 9 – 16.

[4] M. Mazdak,, "FarsiSum – A Persian Text Summarize", **Master Thesis**, Department of Linguistics, Stockholm University, 2004.

[5] H. Dalianis, "SweSum – A Test Summarizer for Swedish", **Technical Report TRITANAP0015**, 2000.

[6] H. Eduard and L. Chin Yew, "Automated Text Summarization in SUMMARIST", **In Proceedings of the Workshop on Intelligent Scalable Text Summarization**, 2000, pp. 18-24,

[7] H. Sagiion and G. Lapalme, "Generating indicative-informative summaries with sumUM", **Computational Linguistics**, Vol. 28 , No. 4, 2002, pp. 497 – 526.

[8] M. A. Hearst, "Untangling Text Data Mining," **In Proceedings of the Association Computational Linguistic,** Vol. 37, 1999, pp. 3–10.

[9] T.M Chang and W.F Hsiao, "A Hybrid Appraoch to Automatic Text Summarization", **In Proceedings of the 8th IEEE International Conference on Computer and Information Technology,** 2008.

[10] I, Mani, **Automatic Summarization,** John Benjamins, 2001.

[11] H.P. Luhn, "The Automatic Creation of Literature Abstract", In **IRE National Convention**, 1958, pp. 60-68.

[12] A.M. Diola, J.T.T.O Lopez, P.F. Torralba, S.So and A.Borra, "Automatic Text Summarization", **In Proceedings of the 2nd National Natural Language Processing Research Symposium,** 2004.

[13] S.Y.Tai, C.S.Ong and N.A.Abdullah, "On Designing an Automated Malaysian Stemmer for the Malay Language", **In the Proceedings of the 5th International Workshop Information Retrieval with Asian Language**, 2005.

[14] A. Selamat and K.B. Yee, "Web-Based Automated Essay Marking System for Historical Malay Text Using Nearest-Neighbor Technique**", In Proceedings of International Conference on Knowledge Management**, 2005.

[15] A.T.Kwee, F.S.Tsai and W.Tang, "Sentence Level Novelty Detection in engliah and Malay", **Lecture Notes in Artificial Intelligence**, Vol. 5476, 2009, pp. 0 -51.

[16] L.C. Tong, "English-Malay Translation System : A Laboratory Prototype", In Proceedings of COLING, 1986.

[17] T. Baldwin and S.Awab, "Open Source Corpus Analysis Tools for Malay", **In the Proceedings of the 5th International Conference on Language Resource and Evaluation,** 2006.

[18] WordNet is available at http://wordnet.princeton.edu/

[19] K. R. McKeown and D. R. Radev, "Generating Summaries of Multiple News Articles", **In the Proceedings of ACM Special Interest Group on Information Retrieval (SIGIR)**, 1995.

[20] M. Hassel and N. Mazdak, "FarsiSum – A Persian Text Summarizer", **In the Proceedings of 20th International Conference on Computational Linguistics,** 2004.

[21] H.P Edmundson, "New Method in Automatic Extracting", **In Journal of the Association for Computing Machinery**, Vol 16, No. 2, 1969, pp. 264-285.

[22] I. Mani, **Automatic Summarization**, Amsterdam: John Benjamins Publishing and Co., 2001.

[23] I. Mani and M.T. Maybury, **Advances in Automatic Text Summarization**, MIT Press, 1999.

[24] M. Shigeru and Y. Kazuhide, "Some Research Topics and Future Prospects in Text Summarization. Statistical Method, Paraphrasing, and More", **In Joho Shori Journal**, Vol.43, No.12, 2002, pp.1310-1316.

[25] R. Barzilay and L. Lee, "Learning to Paraphrase: An Unsupervised Approach Using Multiple-Sequence Alignment", **In the Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL)**, 2003.

[26] C.Y. Lin and E. Hovy, "Manual and Automatic Evaluation of Summaries", **In Proceedings of the ACL-02 Workshop on Automatic Summarization**, 2002, pp 45-51.

[27] E. Hovy, Text Summarization, **The Oxford Handbook of Computational Linguistic**, Oxford University Press, 2005, pp. 584 -598.

[28] M. Hassel, "Exploitation of Named Entities in Automatic Text Summarization for Swedish", **In Proceeding of Nordic Conference on Computational Linguistic**, 2003.

[29] T.F. Hand, "A Proposal for Task-based Evaluation of Text Summarization Systems", **In the Proceedings of a Workshop Association for Computational Linguistics**, 1997, pp. 31-38.

Electroencephalographic Signal Based Clustering of Stimulated Emotion Using Duffing Oscillator

Pavel BHOWMIK

Department of Electronics and Tele-Communication Engineering, Jadavpur University, Calcutta 700032, West Bengal, India.

P. Bhowmik, S. Das, D. Nandi, A. Chakraborty, A. Konar, A. K. Nagar

P. Bhowmik, S. Das and A. Konar are with the dept. of Electronics and Tele-Communication Engineering., Jadavpur University, Calcutta- 32. (E-mails: bpavel88@gmail.com, sauvik.095@gmail.com, konaramit@yahoo.co.in).
D. Nandi is with Calcutta Inistitute of Engineering and Mangement, Cacutta-24 (E-mail: dolphia85@gmaail.com)
A.Chakraborty is currently with the dept. of Computer Science and Engineering, St. Thomas' College of Engineering and Technology, Calcutta-23. She is also a visiting faculty of ETCE Dept., Jadavpur University. (E-mail: aruna_stcet@rediffmail.com).
A. K. Nagar is with the dept. Of Computer Science, Liverpool Hope University, Liverpool, U.K. (E-mail: nagara@hope.ac.uk).

ABSTRACT

Duffing Oscillator is well known for its chaotic behaviour. This paper aims at clustering of emotions from the EEG response to external audio-visual stimulus used for excitation of a subject. The EEG signal corresponding to a specific emotive stimulus is used as an excitation input to a non-linear Duffing Oscillator dynamics, and the phase trajectory plots of the two state variables of the oscillator dynamics exhibits significant differences for various emotion-excitatory stimuli. Experimental investigations reveal that injection of Gaussian noise with a Signal-to-Noise Ratio as low as 25 dB retains the results of emotion clustering, indicating robustness of clustering. Further, with different pre-classified audio-visual stimulus responsible for excitation of a specific emotion, the phase-portraits obtained from EEG data of the subject have substantial similarity, indicating accuracy in clustering.

Keywords: Duffing oscillator, EEG, Emotion clustering and Gaussian Noise.

1. INTRODUCTION

Perception involves interpreting sights, sounds, smells and touch. Perception is relatively younger discipline in Artificial Intelligence, and we are afraid that there are fewer works on perception about emotions. Researchers, however, are keen to develop new models and techniques to understand and recognize emotions from external manifestations, such as crying, laughing etc. this paper deals with classification of emotion aroused with audio-visual stimulus from electroencephalographic (EEGs) signals. Biologists believe that most of our high level understanding process involving emotions is due to the interaction of neural and hormonal activities. EEGs that represent neural activities of the brain might help us in better understanding human emotions than other widely used modes including facial expression [1], [8], [17] and voice [2], [6], [17].

In recent times, researchers have started paying attention to Electroencephalography (EEG) [19], functional Magnetic Resonance Imaging (fMRI) [11], [18], Positron Emission Tomography (PET) [18], and Magnetoencephalography (MEG) [7] –based information to correctly determine the emotional response to external stimulus. However, unfortunately, very little of brain functioning could be identified until this time, and consequently, almost no interesting results have been reported so far on emotion clustering from the above modes of information extraction. The primary objective of this paper is to classify emotion of a subject from his/her EEG signal, obtained through audio-visual excitation of the subject. In our early research [4], [5], we classified the input stimulus based on their power of excitation on a specific emotion. We used these stimuli in the present experiment, and would like to examine whether stimulus used for excitation of same emotion would ultimately map EEG's to a unique pattern.

In order to examine similarity among EEG patterns corresponding to a specific emotion excitatory stimulus, we employ a Duffing Oscillator and the response of the oscillator to the EEG signal as excitation input is recorded. Phase trajectories are built up with two state variables of the oscillator dynamics, and similarity in the chaotic behavior in the phase trajectory is noted for similar stimulus. This fundamental observation reveals that EEG obtained for arousal of a specific emotion has a unique characteristic. Thus emotion classification by EEG signal should result in good accuracy in comparison to other traditional means of emotion clustering from voice and facial expressions.

The paper is classified into 5 sections. In section 2, we briefly outline the state space representation of a Duffing Oscillator dynamics and how phase trajectories were obtained from the time-response of the oscillator dynamics. In section 3, we represent the experimental results for emotion clustering by noting similarity in phase trajectory. The effect of noise on the EEG signal is studied in section 4. Conclusions are listed in section 5.

2. THE DUFFING OSCILLATOR DYNAMICS AND PHASE RESPONSE

In this section, we propose a specialized non-linear oscillator dynamics, which has a proven chaotic behavior [3], [13], [14], [16] in its temporal response. The dynamics of Duffing Oscillator has a similarity with typical spring-mass load system of a conventional mechanical process [12], [15]. However, the spring in the present context, being a non-linear device, has a restoration force proportional to its cubic linear displacement. Naturally, the restoration force of ideal spring that obeys Hooke's Law is also maintained in the Duffing Oscillator dynamics. Consequently, the restoration force has two components, one following Hooke's Law, while the other is due to a high stiffness condition of the spring, represented by a cubic displacement term. The dynamics of Duffing Oscillator is given in equation (1).

$$\frac{d^2x}{dt^2} + \delta\frac{dx}{dt} + \beta x + \alpha x^3 = \gamma\cos(\omega t) + e(t) \qquad (1)$$

where,

x represents the linear displacement,

$\frac{dx}{dt}$ represents the velocity of a unit mass connected in spring-mass load system,

βx and αx^3 are due to spring restoration force,

$\gamma\cos(\omega t)$ is a fixed excitation input to maintain certain level of oscillation in the response of the dynamics, and

e(t) is the disturbance input to the oscillator.

In this present context, we use the EEG signal as the disturbance input e(t). We took α=1, β=-1, γ=0.826, δ=0.5 and the gain of the EEG signal to be 5. The basic Duffing Oscillator dynamics (1) can equivalently be represented by (2) and (3).

$$\dot{x} = y \qquad (2)$$

$$\dot{y} = \gamma\cos(\omega t) + e(t) - \delta\frac{dx}{dt} - \beta x - \alpha x^3 \qquad (3)$$

At first, the EEG signal, which was obtained in sampled version, was passed through a First-Order-Hold circuit, whose transfer function is given by:

$$G_h(s) = \left[\frac{1+Ts}{T}\right] \times \left[\frac{1-e^{-Ts}}{s}\right]^2 \qquad (4)$$

where

T= sampling time,

s= Laplace-domain operator.

The hold-circuit is used to get a continuous version of discrete EEG signal. Then, a Runge-Kutta algorithm was used to solve the coupled-differential equations (2) and (3), and phase portraits for x against y at different time slots are plotted. One typical phase portrait for an initial value of x(0)= 2 and y(0)= 20 is given in Fig. 2.a for convenience. Since Duffing Oscillator has a non-linear dynamics, as shown in the block diagram in Fig. 1, it is apparent that for varying initial conditions, the phase portraits could have been different shapes. However, experimental instances reveal that a chaotic response of the dynamics prevails even for redefining new initial conditions. Fig. 2.b and 2.c illustrate this behavior with different initial conditions.

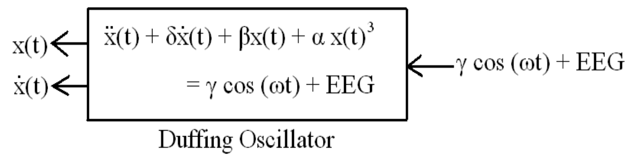


Fig 1: Block diagram of a Duffing Oscillator

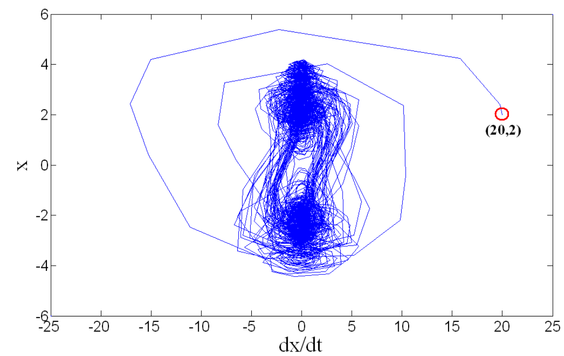


Fig. 2.a: Phase trajectory of anger with initial condition at x(0)= 2, y(0)= 20.

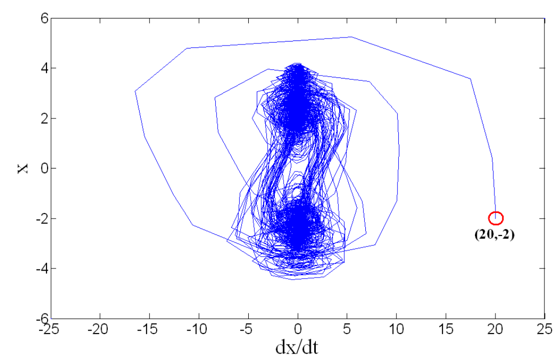


Fig. 2.b: Phase trajectory of anger with initial condition at x(0)= -2, y(0)= 20.

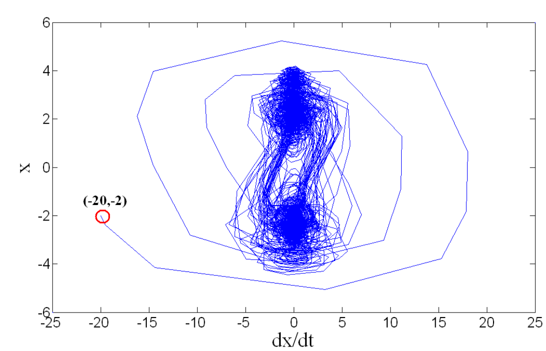


Fig. 2.c: Phase trajectory of anger with initial condition at x(0)= -2, y(0)= -20.

3. EMOTION CLUSTERING FROM EEG SIGNALS USING DUFFING OSCILLATOR

The time-continuous EEG signal obtained through first order hold circuit is used to excite the Duffing Oscillator, [9],

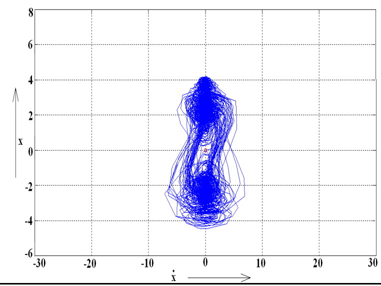
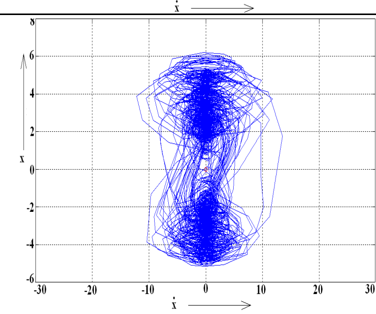
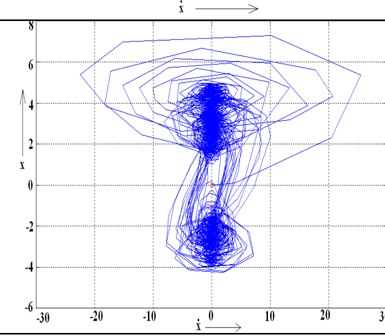
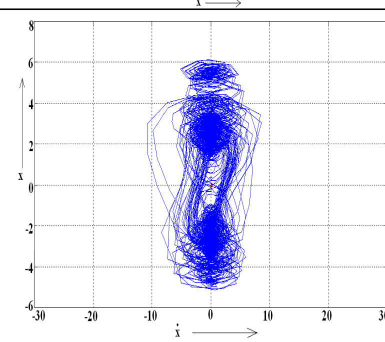
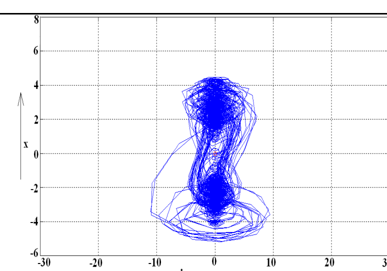
[10], [14], [19] and the response of the oscillator is obtained by solving the differential equation using Runge-Kutta method. The experiment was conducted with 15 audio-visual stimuli, each 3 of which correspond to exciting a specific emotion. The principles of automatically identifying the best audio-visual stimulus, appropriate for excitation of a given emotion is briefly outlined below.

To identify the right audio-visual stimulus responsible for arousal of a given emotion, we classified the stimulus manually with the help of 50 observers, most of whom are University students and faculties. Each observer was asked to classify a given audio-visual stimulus into 5 emotion-arousing classes - anger, fear, joy, relaxation and sadness. He/she used a 100-point scale, and assigned individual score to the entire possibility space of 5 emotions, such that sum of the scores assigned to a given audio-visual stimulus is equal to 100. For 50 observers, we determine the mean and variance of their assignments to a particular emotion-prone category, and evaluate the ratio of mean/variance for each of the 5 emotions. The emotion having the largest mean/variance ratio is considered the best category for a given stimulus. The experiment was repeated for 50 such stimuli, and the mean/variance ratio of the winning emotion for each stimulus is identified. A sorting algorithm is then applied to rank them in descending order of their mean/variance measure in the specific emotion category. The first 3 stimuli for each category of emotion are then identified from the list. The entire experiment was performed with these 3 stimuli responsible for excitation of a specific emotion. Consequently, for 5 emotions, we have 5×3=15 best-selected audio-visual samples. Table I gives the tabular representation of the results obtained by responses of 50 subjects, each of whom was shown 60 audio-visual stimuli. It is apparent from the Table that the row-sum in Table I is always 100.

TABLE I: ASSESSMENT OF THE AROUSAL POTENTIAL OF SELECTED AUDIO-VISUAL MOVIE CLIPS IN EXCITING DIFFERENT EMOTIONS

Subjects used to access the emotion aroused by the audio-visual clips	Title of audio-visual clips	Percentage arousals of different emotions by a clip				
		Anger	Relax	Joy	Sad	Fear
Subject 1	Clip 1	0	20	80	0	0
Subject 2	Clip 1	0	25	75	0	0
…						
Subject 50	Clip 1	0	12	88	0	0
Subject 1	Clip 2	0	82	0	9	9
Subject 2	Clip 2	0	80	0	12	8
…						
Subject 50	Clip 2	0	84	0	10	6
…						
Subject 1	Clip 60	78	10	0	0	12
Subject 2	Clip 60	80	16	0	0	4
Subject 50	Clip 60	84	8	0	0	8

TABLE II: EXPERIMENTAL RESULTS ON THE CLUSTER-SHAPES

Emotions and remarks on related phase plots	Experimentally obtained cluster shapes (without noise)
Anger The phase trajectory covers the least area and width, confined within $\dot{x} = -8$ to $\dot{x} = 8$.	
Fear An extension is visible to the right side of the main phase trajectory.	
Happiness The upper lobe is highly dispersed, less dense, thus covering the maximum area.	
Relaxation Above the upper lobe, a thick lobe is formed. An extension is formed nearer to the left side of the original upper lobe.	
Sadness Below the lower lobe, another lobe is formed, which is sparse in nature.	

Clustering from phase trajectory

We performed two different experiments for clustering of emotions in the EEG space. First, different audio-visual stimuli were used to excite a specific emotion of a subject, and response of the Duffing Oscillator having initial condition x= 0 and y=0 from his/her EEG signal was obtained. Table II above gives a comparative study of the phase portraits of x against y, formed due to one of the selected audio-visual stimuli for each of the emotions. We noted that for three stimuli responsible for exciting the same emotion, the phase trajectories looked almost similar, indicating the fundamental truth that similar excitations for arousal of a given emotion is responsible for excitation of similar brain-activities, pertaining to similar EEG response. These EEG, when fed to a Duffing Oscillator, thus maintains similarity in these portraits of the oscillator state variables.

Figures 3.a, 3.b, 3.c, 4.a, 4.b, 4.c, for example, demonstrate similarity in the phase portrait for excitation of emotions fear and relaxation from its stipulated stimulus list.

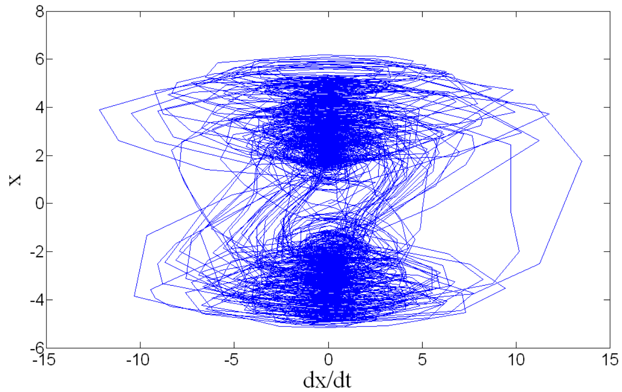


Fig. 3.a: Phase trajectory for Fear due to 1st stimulus of Fear

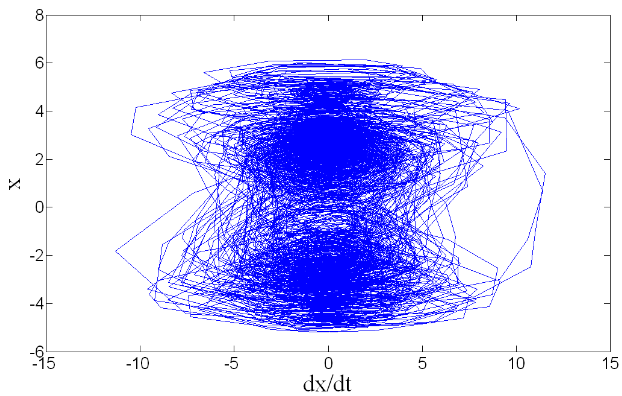


Fig. 3.b: Phase trajectory for Fear due to 2nd stimulus of Fear

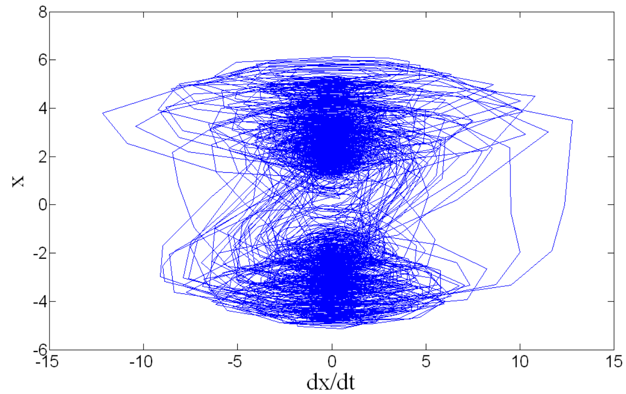


Fig. 3.c: Phase trajectory for Fear due to 3rd stimulus of Fear

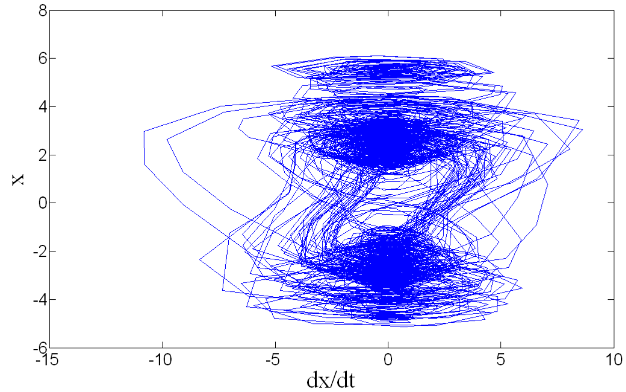


Fig. 4.a: Phase trajectory for Relaxation due to 1st stimulus of Relaxation

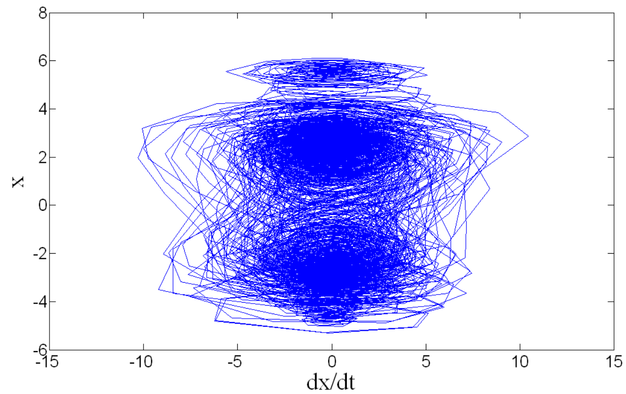


Fig. 4.b: Phase trajectory for Relaxation due to 2nd stimulus of Relaxation

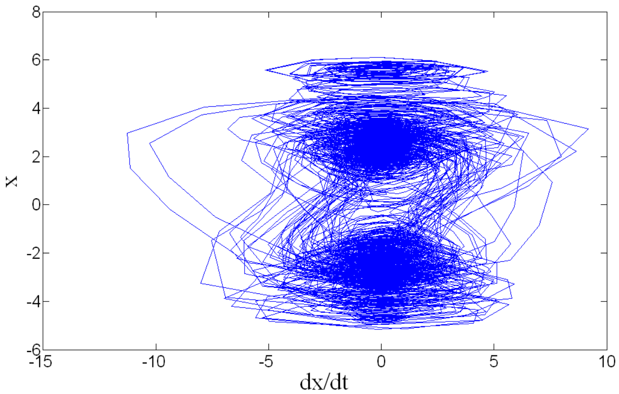


Fig. 4.c: Phase trajectory for Relaxation due to 3rd stimulus of Relaxation

4. EFFECT OF NOISE ON EMOTION CLUSTERING FROM DUFFING OSCILLATOR RESPONSE

In this section, we experiment by adding noise to the original signal corresponding to a specific emotion, and note the changes in the phase portrait obtained from the Duffing Oscillator response. It is interesting to note that when Signal-to-Noise Ratio of the EEG signal is maintained to a level of 25 dB, the phase portraits maintains similarity, indicating robustness in emotion clustering.

Figures 5.a, 5.b, 5.c, 6.a, 6.b, 6.c demonstrate the behavior in the phase portrait for different level of Signal-to-Noise Ratio as indicated in the figure caption. It is also noteworthy that when the Signal-to-Noise Ratio goes below a threshold, misclassification starts, by noting differences in the phase portraits for a given emotion.

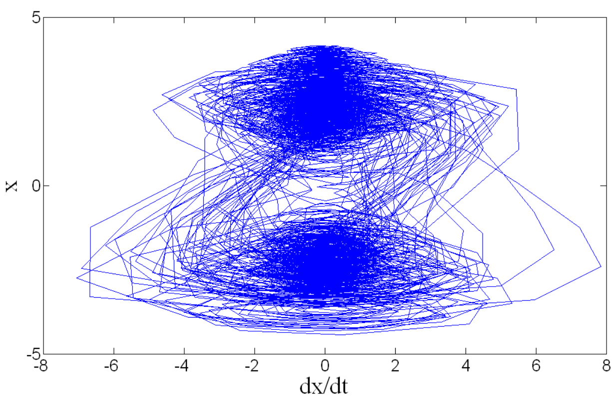


Fig.5.a: Phase trajectory for Anger when the EEG signal is corrupted by a noise of SNR 30dB

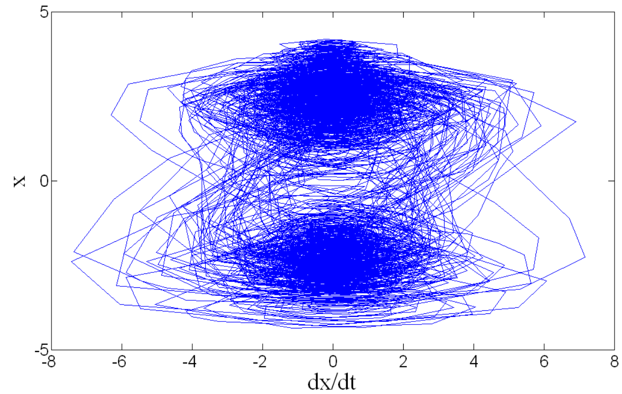


Fig. 5.b: Phase trajectory for Anger when the EEG signal is corrupted by a noise of SNR 25dB

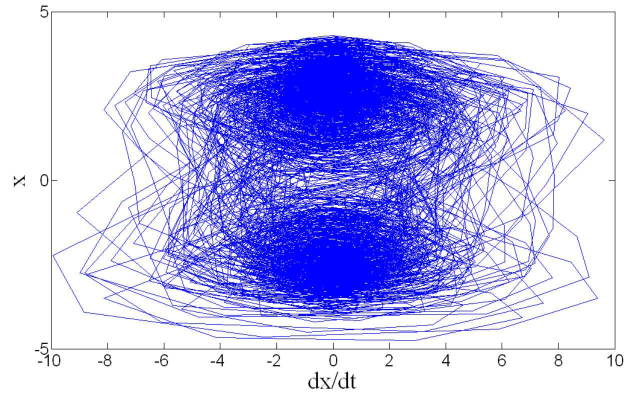


Fig. 5.c: Phase trajectory for Anger when the EEG signal is corrupted by a noise of SNR 20dB

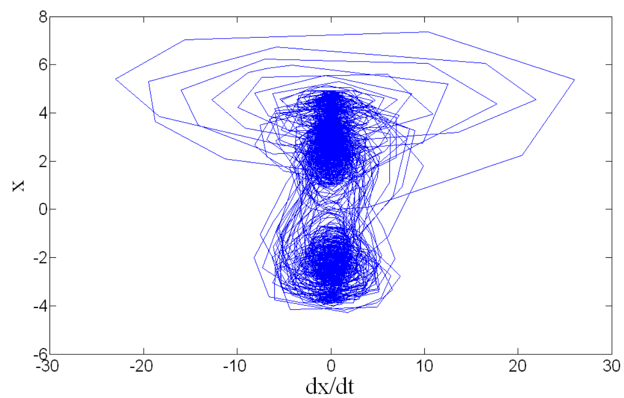


Fig. 6.a: Phase trajectory for Joy when the EEG signal is corrupted by a noise of SNR 30dB

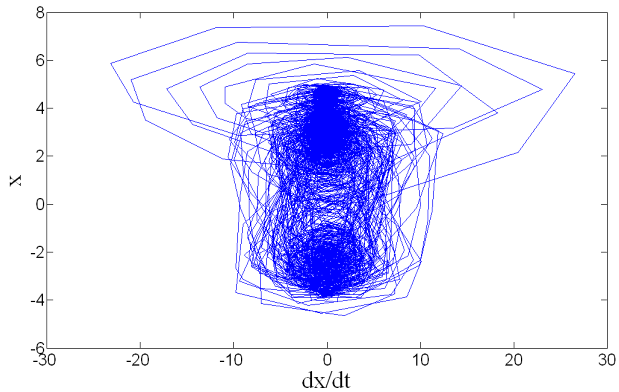


Fig. 6.b: Phase trajectory for Joy when the EEG signal is corrupted by a noise of SNR 25dB

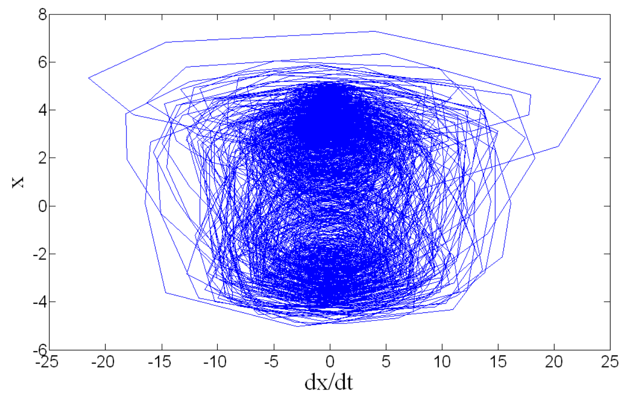


Fig. 6.c: Phase trajectory for Joy when the EEG signal is corrupted by a noise of SNR 20dB

5. CONCLUSIONS

The paper attempted to cluster emotions from the stimulated EEG signals using Duffing Oscillator as a medium. EEG signals aroused with specific emotion excitatory stimulus were supplied as an input to the Duffing Oscillator, the phase portrait corresponding to the response of which is plotted. Similarity in phase portraits is considered clustering of EEG in the phase-space. Consequently, clustering of emotions can be undertaken by determining similarity of the EEG signals. A noise analysis undertaken reveals that the clustering of emotions can be clearly visualized in the phase portrait, as long as the Signal-to-Noise Ratio is maintained above a prescribed threshold (25 dB). It is also obtained from the experiments that excitations responsible for arousing specific emotion have similar EEG signals, which can be clustered easily in the phase space from the response of the Duffing Oscillator. In brief, the similarity in chaotic behavior of the phase portraits resembles the similarity in EEG and consequently, similarity in emotions. The paper thus opens up a new methodology of emotion clustering from the EEG signals in the phase-space.

REFERENCES

[1] Adolphs R, Damasio H, Tranel D, Damasio AR., "Cortical systems for the recognition of emotion in facial expressions," Journal of Neuroscience 1996;16:7678±87.
[2] Bezooijen Rv, "The Characteristics and Recognizability of Vocal Expression of Emotions," Walter de Gruyter, Inc., The Netherlands, 1984.
[3] Cristian Bonatto, Jason A. C. Gallas, and Yoshisuke Ueda, "Chaotic phase similarities and recurrences in a damped-driven Duffing oscillator", Phys. Rev. E 77, 026217 (2008).
[4] Aruna Chakroborty, "Cognitive Cybernetics – A Study of the Behavioral Models of Human, Interactions", Ph D Thesis, Jadavpur University.2005
[5] Aruna Chakroborty and Amit Konar, "Emotional Intelligence: A Cybernetic Approach", Springer 2009.
[6] Dellaert F., Polzin T., Waibel A., "Recognizing emotion in speech," Spoken Language, 1996. ICSLP 96. Proceedings. Fourth International Conference on, Volume: 3, 3-6 Oct. 1996. Pages: 1970 - 1973 vol.3.
[7] Dolan RJ, Heinze HJ, Hurlemann R, Hinrichs H. Magnetoencephalography (MEG) determined temporal modulation of visual and auditory sensory processing in the context of classical conditioning to faces. Neuroimage 2006;32(2):778–89.
[8] Ekman P., Friesen W. V., "Facial Action Coding System: A Technique for Measurement of Facial Movement," Consulting Psychologists Press Palo Alto, California, 1978.
[9] P. Holmes, "A nonlinear oscillator with a strange attractor," Philosophical Transactions of the Royal Society A, 292, 419-448, 1979.
[10]P. Holmes and D. Rand , "Phase portraits and bifurcations of the non-linear oscillator," International Journal of Non-linear Mechanics, 15, 449-458, 1980.
[11]Tom Johnstone, Carien M. van Reekum1, Terrence R. Oakes and Richard J. Davidson, "The voice of emotion: an FMRI study of neural responses to angry and happy vocal expressions," Social Cognitive and Affective Neuroscience, vol. 1, number 3, Pp. 242-249, October 20, 2006.
[12]Benjamin C. Kuo, "Automatic Control Systems," John Wiley & Sons Inc.
[13]H. Nakano, T. Saito, "Grouping synchronization in a pulse-coupled network of chaotic spiking oscillators," IEEE Trans. Neural Netw., vol. 15, no. 5, pp. 1018-1026, Sep. 2004.
[14]S Novak, RG Frehlich, "Transition to chaos in the Duffing oscillator," Physical Review A, 1982 – APS
[15]K. Ogata, "Modern Control Engineering," Fifth Edition, Prentice Hall
[16]E. Ott, "Chaos in Dynamical Systems (2nd edition)," Cambridge University Press, 2002.
[17]James A. Russell,Jo-Anne Bachorowski, and José-Miguel, "Facial and Vocal Expressions of Emotion," Fernández-Dols Vol. 54: 329-349 (Volume publication date February 2003) (doi:10.1146/annurev.psych.54.101601.145102) First published online as a Review in Advance on October 4, 2002
[18]F. Schneider, W. Grodd, R. E. Gur, U. Klose, A. Alavi and R. C. Gur, "PET and fMRI in the study of emotion," Psychiatry Research: Neuroimaging, Volume 68, Issues 2-3, 7 February 1997, Pages 174-175
[19]Ye Yuan, Yue Li, "Study on EEG Time Series Based on Duffing Equation," bmei, vol. 2, pp.516-519, 2008 International Conference on BioMedical Engineering and Informatics, 2008.

A Survey of Dynamic Adaptation Techniques

Jorge Fox, Siobhán Clarke
Lero - The Irish Software Engineering Research Centre
Distributed Systems Group
School of Computer Science and Statistics
Trinity College Dublin, Ireland

{firstname.lastname}@cs.tcd.ie

Abstract

The growing complexity of software systems as well as changing conditions in the operating environment demand systems that are more flexible and dependable. A possible solution we envisage is the use of mechanisms for effecting behavioural enhancements or changes in running systems. This has been called Dynamic Adaptation (DA). This implies exploring a number of challenges. Some questions that have to be addressed relate to finding mechanisms for: service detection, implementation of behavioural changes during runtime, service interaction and service behaviour modification. This paper introduces a survey of approaches to dynamic adaptation in order to assess their capabilities. We describe a framework for comparing approaches to (dynamic) adaptation (DA) and evaluate selected approaches to DA against this framework. Based on the comparison framework we outline current trends in DA technologies.

Keywords: Software Engineering, Dynamic Adaptation, Software and Systems Development, Run-time Systems

1 INTRODUCTION

Dynamic Adaptation (DA) is gradually becoming a key element in software engineering for a growing range of domains such as: automotive systems, web services, networks, among others. Furthermore, within these domains the requirement to adapt to changing conditions in the environment as well as the need to deploy additional services on heterogeneous platforms, motivates the use of technologies facilitating a higher level of adaptation to changes.

A review of the state of the art on DA, reveals open research areas. Consider, for instance, time-bounded runtime dynamic systems. As will be explored later in this work, DA within time bounds and without feature interference is a research field in which no conclusive results have been achieved. Naturally, there is a number of approaches for adaptation, but at the same time most are static ones. More importantly, the flexibility of adaptation or the degree at which adaptations are achieved, is in most cases limited. Also, in many existing DA frameworks adaptation is achieved by parametrisation or reconfiguration, which may render limited solutions with respect to flexibility and limit further adaptations.

The relevance of DA lies on the growing need for flexible and dependable systems in complex environments. These are environments characterized by the need for: ubiquity, distributed systems, interoperability; as well as controlled and foreseeable adaptation mechanisms.

2 DYNAMIC ADAPTATION

In this section we explore current concepts and definitions related to DA. Adaptability is defined as the ability of software systems to withstand changes in their environment. As Yan et al. mention "a software system will be adaptable provided its software architecture is itself adaptable in the first place" [18].

Adaptive systems are those that posses the ability to adapt at run-time to react to user needs, system intrusions or faults, changing operational environment, resource, and performance variability. We consider dynamic adaptable systems to be a subset of adaptable systems wrt. moment of adaptation. Dynamic adaptable systems perform adaptations at run-time as opposed to performing adaptations at design time.

In this sense, [6] and [9, 10] introduce a thorough review of adaptability and adaptiveness. In Section 4, we inspired on their classification to develop our comparison framework. However, the emphasis of our work is on dynamic *adaptive* systems.

3 COMPARISON OF APPROACHES

In order to classify groups in DA, we first scrutinised the possible lines of research in adaptive systems. This means, the extent to which a system adapts to changes in the environment, whether it is through structural means i. e., architectural adaptation, changes in the parametrisation of the system, or a combination of both. Another set of criteria we found, relates to the degree of anticipation to changes. In other words, the extent to which the adaptation reacts to changes in the environment: fully unanticipated or foreseable changes. Clearly, the former is hard to conceive and even more to implement in its pure form. Second, we classify adaptability according to characteristics we identified as relevant for adaptive systems, such as: degree of anticipation, scope of adaptation changes (i. e., architectural vs. localised), whether it is achieved with composition mechanisms or through parametrisation and whether there is tool support or not. Equally important, some authors (see [6]) consider the relationship between what is called "compositional" as opposed to "parametric" adaptation, and mixed-forms. We consider both as two dimensions in the classification, which can be combined. Our classification criteria is further explained in Section 4. Third, the classification criteria and the approaches we analysed is represented in Table 1, in which we assigned values (ranging from low to medium and high) to the surveyed research teams for each criteria. Assignment of values was based on a review of the literature and available information. Furthermore, our classification schema draws inspiration from [6], in particular on the distinction on composition adaptation as opposed to parametrical adaptation, and anticipated against unanticipated adaptation.

Table 1 shows the classification criteria we propose for (dynamic) adaptive systems. Some criteria can be combined, whilst others may not. Take for instance, achieving adaptation through a high level of parametrisation and localised in one or two precise modules, this is relatively straightforward and is present in most approaches. On the other hand, some combinations may not be attainable like having *total anticipation*, meaning full anticipation to environmental changes and achieving it at runtime. Hence, at this stage of our research, results indicate that these criteria are interdependent. Still, it belongs to work in progress to identify the extents and properties of such relationships.

4 CLASSIFICATION CONCEPTS FOR DA

We briefly introduce the classification concepts we propose to describe current research approaches in DA.

4.1 Unanticipated Adaptation

This concept indicates the degree to which the adaptation triggers and possible adaptation needs are known in advance or not. The higher the level of adaptation to unforeseen changes, the higher the level of the framework in this parameter. We consider that a higher level of adaptation to non foreseeable changes, indicates a more flexible or more generic adaptation framework.

4.2 Scope

This concept refers to the extent to which changes in adaptation spread over the software system. We assign values from low to high according to the following. If the adaptation is limited to a localised component, the approach gets the value low in scope, if adaptation is performed on a reduced number of components it is classified as medium level and finally if the adaptation reaches a system-wide level then it is considered high in scope of adaptation.

4.3 Parametric Adaptation

This criterion indicates whether adaptation is achieved by means of adjusting or fine-tuning predefined parameters in given software entities, such as components, services or methods. A higher parametrisation may indicate a rather inflexible framework, due to a higher dependency on predefined values and parameters.

4.4 Compositional

This classifier signifies that the framework under analysis achieves adaptations through the insertion or replacement of functional units. By functional units we mean components or sets of components or services. A compositional approach usually relies on binding and unbinding mechanisms.

4.5 Tools

We also consider whether the approach has tools to support dynamic adaptable systems, such as a development environment or a runtime monitoring environment. This is the last classification criteria provided in Table 1. We believe this criteria to be of a relatively high importance given the need to facilitate adoption of the approach or framework.

In the following section, we introduce the research teams that we considered representative enough to explore our classification criteria. Selection is based on a thorough review of the literature and subsequent selection of teams that had relevant publications in the field. We also privileged those teams working within a consortium of universities and institutions, or an established research group in academia. The objective of our survey is to explain our classification concepts and identify important traits in the field, rather than introducing an exhaustive review of DA approaches.

5 ADAPTATION TECHNIQUES

We identify three main groups of adaptation techniques, these are dynamically linking and unlinking selected components, use of generic interceptors and reconfiguration techniques. These techniques were selected after reviewing the literature and analysing the current techniques for adaptation.

The following DA technologies represent an overview of various methodologies, methods and techniques in the field. Naturally, there are other approaches than the ones we selected, however we consider this selection to be sufficient ground for comparisons.

In this work we considered technologies that achieve DA by compositional adaptation, but also technologies that achieve adaptation through reconfiguration or by means of interceptors. For an exhaustive list of adaptive frameworks see [6]. Furthermore, we did not include in this survey approaches that focus on very particular issues like "Hypervisor Modules" [12], or that centre on particular problems of DA such as interoperability [7].

5.1 Dynamically Linking and Unlinking Selected Components

This technique is used by the Extensible Service-Oriented Component Framework (iPOJO) [4], In iPOJO, Plain Old Java Object (POJO) components are "injected" by handlers onto the base component. These handlers manage service publication and providing as well as dependencies. When a service satisfies given dependency conditions, then it is published, otherwise it is ignored. Components relate to each other connecting through these dependencies. Components turn invalid when a service provider (dependency) is gone. Therefore creation or activation of components is equivalent to publicising its dependencies, whilst deactivating a component is achieved by eliminating dependencies. In general terms, iPOJO consists of a component model that "injects" Plain Old Java Objects (POJO's) at runtime. This is the overall mechanism through which systems are adapted in this approach. This is mainly realised through the management of dependencies and service providing, while the business logic is set at the level of POJO's. DA is then implemented by means of redirecting dependencies; this is managed by handlers which in turn are selected by meta data indicated in XML files.

A component container handles all the service-oriented computing aspects and separates them from the business logic which remains in the base component. iPOJO provides a runtime component environment that simplifies development of applications over the platform provided by the Open Services Gateway initiative (OSGi). OSGi is a technology aimed at facilitating the interoperability of applications and services through a component integration platform [3, 4]. The concept of service used in iPOJO is rather abstract and seems closer to that of features in a broader sense. This approach makes the implementation dependant on the underlying service runtime framework, which renders this work to have a moderate scope for adaptations. iPOJO provides a high level of compositionality as well as dynamism regarding injection, binding and rebinding of components or POJO's. Whilst at the same time, the scope of adaptation is determined by the underlying framework and its availability, which poses limitations to integration with services or components not running on OSGi.

Another technique is represented by PCOM. PCOM is a distributed application model which supports DA via signalling mechanisms and adaptation strategies, see [2]. In PCOM components are entities that interact with each other in order to fulfil their dependencies. This definition of components resembles that of "services," yet services are more explicitly aimed at cooperating, if needed, to fulfil their own functionality. Applications in PCOM are described by a tree of components and their dependencies, being the root component a sort of "`main()`" program or application identifier.

However, it is not clear in [2] whether dependencies only occur following the branches of the tree or some other relationships are allowed and to what extent these dependencies are transitive. Besides that, the authors acknowledge that arbitrary graphs would cause complications. This can be seen as a limitation in the framework. For the above mentioned reasons, we may consider PCOM as more parametric than compositional. Given that some strategy for adaptation has to be set beforehand it achieves a medium level of unanticipated adaptation. The framework is not as dynamic as ACT (see Sect. 5.3), still does claim to support runtime adaptation, so we considered it highly dynamic as well.

Another group of techniques, closely related to the ones in dynamic linking and unlinking of components propose the use of composition frameworks, filters, paths and injectors. In this category we find a technique that introduces the use of "injectors". Injectors in The Object Infrastructure Framework (OIF) offer a way to facilitate evolution and creation of distributed systems. Its main mechanism is injecting behaviour on the communication path between components [5]. Behaviours may be injected on the client or the server. Instances and methods can have a distinct sequence of injectors. Stubs can be changed during execution fostering the dynamic behaviour of the system. There is a high-level specification language and a compiler to support

Concept/ Approach	ACT (CORBA)	DAiSI	Dynamic TAO	iPOJO	MADAM	MBD DA	PCOM
Unanticipated adaptation	medium				low	low	low
Scope	low	low	medium	low	medium	low	low
Parametric	low	medium	medium		medium		medium
Compositional	medium	low	medium	medium	low	low	low
Tools	low	low				low	low

Table 1. Evaluation of selected research approaches to adaptation

OIF. OIF injectors work with the Common Object Request Broker Architecture (CORBA) stubs with some modifications on skeletons to obtain the injector sequence for each method. The injector may modify the target, the operation arguments, the annotations, and the return value. It can also invoke other remote calls. Injections may perform actions before and after the server action. This allows to modify the flow of control. In OIF components are black-box objects. Injectors are created by two classes, the injector itself and a factory that creates instances of the injector. Injector instances are created by calls on the factory when building CORBA proxies. Injectors are then inserted in the methods using an aspect-oriented programming language.

Client side injectors can change the destination of a request. There are different kinds of injectors: rebind, impatient, insecure, mediating, and balancing. These differ on the decision criteria to select target services. In order to determine the target of a redirection, the injector may rely on a "clerk" which possesses information on the alternatives offered by target services. Clerks can be dynamically arranged in case new services are discovered. To optimize this mechanism clerks can be grouped in a "community" of clerks which share information.

5.2 Dynamic Adaptation with Aspect-orientation

Dynamic adaptation with aspect orientation (AO) in Yang et al [19] is performed in two phases. In the first phase adaptation points are defined and in the second phase the adaptation infrastructure is related to the base program. As Yang et al say, the adaptation infrastructure consists of an adaptation manager and a rule base. Dynamic adaptation is directed through a set of rules. The adaptation kernel is a loose grouping of adaptation managers that are explicitly invoked to check execution conditions and perform adaptations accordingly. At run-time the adapt-ready program is instantiated. Behaviour adaptors in the running program use a filter-chain to trap the respective adaptation manager and determine which rules are satisfied and what corresponding adaptation should be performed.

5.3 Generic Interceptors

The use of generic interceptors is used by approaches like Adaptive CORBA ([16]). These techniques do not modify a component's behaviour, but intercept the messages between components in order to provide for additional behaviour to perform the adaptation. For instance, in an adaptive CORBA template (ACT) generic interceptors are registered with the Object Request Broker (ORB) of a CORBA application at start-up. Interceptors adapt requests, replies and exceptions passing through the ORB. Therefore, the generic interceptors do not modify the component's behaviour. These interceptors have to be previously registered, which restricts the flexibility of the adaptation. See [9].

ACT is a language independent template which can be used to develop an object-oriented framework as well as for enhancing CORBA applications
[16]. It introduces generic interceptors, which are specialised request interceptors registered with the ORB at start-up. Interceptors are static or dynamic. Dynamic interceptors can be registered or unregistered at runtime, while static ones cannot be unregistered with the ORB at runtime. This approach also relies on the notion of weaving for relating the dynamic interceptors at runtime. The concept of generic interceptors provides some underpinnings for unanticipated adaptation, since these interceptors are registered without specific behaviour and may later be enhanced at runtime to implement some needed functionality. For these reasons, we consider this work to achieve a high level of unanticipated adaptation, while achieving only a middle level scope of adaptation since only dynamic interceptors are changed. It also has a middle level of parametrisation since the use of proxies and redirection is needed, and it is highly compositional.

5.4 Reconfiguration Techniques

These techniques aim at adjusting internal or global parameters in order to respond to changes in the environment. Reconfiguration may help to rearrange the elements of a system. Aksit and Choukair [1] identify two major research approaches to reconfiguration: adding configuration elements and the use of component and configuration languages.

Dynamic reconfiguration [13, 14] aims at achieving adaptation at the level of component service usage, component service implementation and configuration adaptation. The first kind of adaptation supports switching components at runtime and selecting services based on some quality property, for instance. The second one, supports altering the behaviour of a component and the realisation of the service it renders. Finally, the third kind of adaptation is oriented to reconfiguring components in a non-localised way, it aims at modifying how components relate and how the services offered are activated or stopped. For more on it see [8]. It works on the basis of a component model for DA and relies on a formal foundation [15]. A related framework is the Dynamic System Infrastructure (DAiSI) [8]. This infrastructure introduces a dynamic adaptive component model which defines how a component has to be structured for DA. Our research indicates that in its current state DAiSI achieves adaptation through parametrisation as well as composition mechanisms. Anticipation to changes seems to be an open issue in this framework, since there is no explicit mechanism to cope with changes and it may not react to unanticipated changes in the environment, rather on those indicated by their configuration component manager (browser). There is a good level of tool support. Another framework is Dynamic TAO, an extension to "The ACE ORB" (TAO). TAO is a standard CORBA Object Request Broker (ORB), see [17]. The salient characteristic of DynamicTAO is the capability of reconfiguring the ORB at runtime "by dynamically linking/ unlinking certain components." [11, 17]. It enables remote reconfiguration and replacement of given ORB components with no need to restart the whole ORB, which is a useful trait for DA. It also provides the means for uploading code with new implementations, which is also essential for DA. Given its reconfiguration and replacement capabilities, we consider it to be highly dynamic. We also consider that the scope of adaptation, meaning the extent to which the system adapts as a proportion of entities with DA capabilities, is in DynamicTAO high, given that the underlying ORB framework allows, at least in principle, for any of the constituent components to be adaptable.

Another framework is Mobility and ADaption enAbling Middleware (MADAM). This framework provides a component model with add-ons for adaptation [6]. With this framework the possible variations for a system are accomplished through the recursive application of predefined realisation plans. Realisation plans are actual composition plans or predefined combinations of components given by the designer. This component model includes an adaptation manager. A composition or adaptation manager is a common mechanism in most adaptive frameworks. Furthermore, MADAM provides a middleware framework for runtime adaptation with: context management, adaptation management and configuration management. Its composition is based on parametric adjustments. Likewise, given that adaptations are predefined in an adaptation plan by a designer, unanticipated adaptation is not possible.

5.5 Model-Based Development of Dynamically Adaptive Software (MBD DA)

Zhang and H.C. Cheng ([20, 21]) have worked on reliability aspects of DA. The authors introduce an approach to realise formal models for the behaviour of adaptive programs. This way, they provide a way to ensure that such adaptations are safe with respect to system consistency. It is based on state-machine representations of adaptive programs. The properties that the program should satisfy throughout its execution are called global invariants. Adaptations are defined as adaptation sets and its behaviour is represented as simple adaptive programs. The properties of the adaptive program are local. Their method takes into consideration dependency analyses for target components, specifically determining viable sequences of adaptive actions and those states in which an adaptive action may be applied safely. This technique supports safe adaptation. MBD DA allows for insertion, removal, and replacement of components, in response to changing external conditions. Their work is explored at the example of a wireless multicast video application. In addition a safe DA process has been developed in a related project [20]).

Their state-machine based formal framework does cover static and dynamic analysis. It is also capable of dealing with runtime systems. Their approach can be supported by different tool suites (see [20]). It offers a medium level of tool support. There was no stronger evidence of a robust tool set available. This work is more focused on providing a formal framework for analysing adaptation programs than on mechanisms or frameworks supporting adaptation itself.

6 CONCLUSIONS

After a review of a number of DA frameworks and approaches, we highlighted salient characteristics for DA systems; particularly the extent of changes or what we called the scope of adaptations, whether these are performed at the underlying framework or on a limited number of components pre-enabled for DA. Also, the level of anticipation to changes is an important attribute, because it determines the capacity of the systems to cope with new services or changes in the environment. Moreover, the particular adaptation approaches may vary depending on the underlying foundation: components, services, or a combination of both. Another aspect is that the adaptation mechanisms themselves are sometimes left to the decision of designers and are specified as parameters on which the system reconfigures or implements the adaptations.

In this work, we identified the need for further research on DA mechanisms; which may allow for higher compositionality and flexibility. Some traits we recognise as significant for DA systems are the breadth of the adaptation, the mechanisms used to achieve adaptation, and the underlying framework or tools available. Another question that influences DA is run-time discovery and replacement of services, and the decision making process behind adaptation.

Finally, there is a need for a framework that allows for runtime discovery or replacement of services, with a runtime environment capable of verifying the reliability of changes and preservation of the execution time bounds of the software system. Specific attributes play a critical role in DA, these are ensuring reliability of the adaptations and preservation of the execution time of the software system after adaptation, and keeping adaptation time within predefined time bounds. In this regards, our survey reveals open areas of research.

ACKNOWLEDGEMENTS

This work was supported in part, by Science Foundation Ireland grant 03/CE2/I303_1 to Lero - the Irish Software Engineering Research Centre (www.lero.ie).

References

[1] Mehmet Aksit and Zièd Choukair. Dynamic, adaptive and reconfigurable systems overview and prospective vision. In *ICDCS Workshops*, pages 84–. IEEE Computer Society, 2003.

[2] C. Becker, M. Handte, G. Schiele, and K. Rothermel. Pcom - a component system for pervasive computing. *Pervasive Computing and Communications, 2004. PerCom 2004. Proceedings of the Second IEEE Annual Conference on*, pages 67–76, 14-17 March 2004.

[3] Clément Escoffier and Richard S. Hall. Dynamically adaptable applications with iPOJO service components. In Markus Lumpe and Wim Vanderperren, editors, *Software Composition*, volume 4829 of *Lecture Notes in Computer Science*, pages 113–128. Springer, 2007.

[4] Clément Escoffier, Richard S. Hall, and Philippe Lalanda. iPOJO: an extensible service-oriented component framework. In *IEEE SCC*, pages 474–481. IEEE Computer Society, 2007.

[5] Robert E. Filman and Diana D. Lee. Redirecting by injector. *Distributed Computing Systems Workshops, International Conference on*, 0:0141, 2001.

[6] Kurt Geihs. Selbst-adaptive software. *Informatik-Spektrum*, 2007. 0170-6012 (Print) 1432-122X (Online).

[7] Robert Hirschfeld and Katsuya Kawamura. Dynamic service adaptation. *Software - Practice and Experience*, 36(11-12):1115–1131, September/October 2006.

[8] Holger Klus, Dirk Niebuhr, and Andreas Rausch. A component model for dynamic adaptive systems. In Alexander L. Wolf, editor, *Proceedings of the International Workshop on Engineering of software services for pervasive environments (ESSPE 2007)*, pages 21–28, Dubrovnik, Croatia, sep 2007. ACM.

[9] Philip. K. McKinley, Seyed Masoud Sadjadi, Eric P. Kasten, and Betty H. C. Cheng. A taxonomy of compositional adaptation. Technical Report MSU-CSE-04-17, Dept. Computer Science and Engineering, Michigan State University, 2004.

[10] Philip K. McKinley, Seyed Masoud Sadjadi, Eric P. Kasten, and Betty H.C. Cheng. Composing adaptive software. *Computer*, 37(7):56–64, 2004.

[11] Philip. K. McKinley, R. E. Kurt Stirewalt, Betty H. C. Cheng, Laura K. Dillon, and Sandeep Kulkarni. Rapidware: Component-based development of adaptive and dependable middleware. Technical report, Michigan State University, 2005.

[12] Thomas Naughton, Geoffroy Valle, and Stephen L. Scott. Dynamic adaptation using xen:thoughts and ideas on loadable hypervisor modules. *First Workshop on System-level Virtualization for High Performance Computing (HPCVirt 2007)*, March 2007.

[13] Marie-Claude Pellegrini and Michel Riveill. Component management in a dynamic architecture. *J. Supercomput.*, 24(2):151–159, 2003.

[14] James M. Purtilo. The polylith software bus. *ACM Trans. Program. Lang. Syst.*, 16(1):151–174, 1994.

[15] Andreas Rausch. DisCComp – a formal model for distributed concurrent components. *Electronic Notes in Theoretical Computer Science (ENTCS)*, 176(2):5–23, 2007.

[16] S. M. Sadjadi and P. K. McKinley. Act: An adaptive corba template to support unanticipated adaptation. *icdcs*, 00:74–83, 2004.

[17] D C Schmidt, B Natarajan, A Gokhale, N Wang, and C Gill. Tao: A pattern-oriented object request broker for distributed real-time and embedded systems. *IEEE Distributed Systems Online*, 2002.

[18] Narayanan Subramanian. *Adaptable software architecture generation using the nfr approach*. PhD thesis, The University of Texas at Dallas, 2003. Supervisor-Lawrence Chung.

[19] Z. Yang, B. H. C. Cheng, R. E. K. Stirewalt, J. Sowell, S. M. Sadjadi, and P. K. McKinley. An aspect-oriented approach to dynamic adaptation. In *WOSS '02: Proceedings of the first workshop on Self-healing systems*, pages 85–92, New York, NY, USA, 2002. ACM.

[20] Ji Zhang and Betty H. C. Cheng. Model-based development of dynamically adaptive software. In *ICSE '06: Proceeding of the 28th international conference on Software engineering*, pages 371–380, New York, NY, USA, 2006. ACM.

[21] Ji Zhang, Betty H.C. Cheng, Zhenxiao Yang, and Philip K. McKinley. Enabling safe dynamic component-based software adaptation. *Architecting Dependable Systems III*, 3549/2005:194–211, 2005.

Architecting Information resources of Brazilian Social Security: approaches of Social Science and Computer Science working together

Claudio Jose Silva RIBEIRO
DATAPREV – Information Technology of Brazilian Social Security,
Rua Prof. Alvaro Rodrigues, 460/303, Botafogo, Rio de Janeiro, RJ, CEP 22280-040, Brazil

ABSTRACT

The increasing use of Information in virtual environments has broadened the problem of Information Management, as it ended by boosting the discussion about its solution towards technology. In this regard, this article introduces part of a research that proposes a realignment of focus in this debate making use of epistemology as it studies the process of planning, understanding and representation of Information resources, with particular emphasis on structuring according to Information Architecture. The theoretical framework started with Information Administration and Information Architecture presuppositions, as well as Project Management Methodology and Development of Information System Methodology. The understanding of information needs was developed making use of Domain Analysis for domain delimitation, understanding and judging the relevance of information within the field under review. Besides that it was complemented by investigations directed towards users through the use of Sense-Making approach, originating in Social Science, and Eliciting Requirements, with origins in Software Engineering. In dialogue with the Information Science and the Computer Science, this paper presents the basic definitions that has being cast some light on the path that it is traveled toward developing and structuring the project of Dataprev about the Content Management System for Brazilian Social Security.

Keywords: Information Management, Domain Analysis, Eliciting Requirements, Sense Making, Information Architecture, Social Security, Brazil

1. INTRODUCTION

The way that society uses and treats information has been changing for the last three decades as a consequence of the emergence of new social, economic or technologic models. These models have promoted a paradigm shift as important as the press invention, or still, as the industrial revolution itself. We notice that information has evolved from a simple and mundane fact, with limited repercussion, to something that is increasingly becoming knowledge to obtain power. The expression "power" must be understood in its broadest sense, like, for instance, to use knowledge in order to gain economic benefits.

Once technology has sped up the dissemination and exchange of information, the discussion about information treatment has catalyzed another transformation process. The large use of information spaces on the Internet (Web Sites), where data and information are available 24 hours a day, 7 days a week, has fertilized further more the change process already on its course.

According to a historical view of information treatment, we come to the subject of Management Information (IM) which, in its genesis, sought to treat the collection and usage of human, technological, financial, material and physical resources to the management of information. These elements must be observed both with a view towards the strategic and operational levels, in order to make the information a useful resource and a part of business strategies for individuals, groups and institutions. The information inventory, the costs involved in the management and use of the information, beyond the identification of the existing informational gaps, were aspects considered in the scope of the studies upon IM, which were maximized with the incorporation of new technologies [1]. Levitan, one of the subject's precursors observes:

"Information Management can be viewed as the keystone of information science and technology. Every activity in the eclectic profession - from user requirements, system design and evaluation to document and knowledge representation, database organization, storage and retrieval techniques, hard and soft technologies applications, as well as repackaging, dissemination, and marketing – relates to the complex process of managing information. " [2]

In this direction, we note that the efforts of gathering, treatment and retrieval of data, information and documents, have also motivated the development of projects at the Brazilian Social Security System environment. Apart from the media in which data, information and documents are materialized, the development of these works is fundamental, considering that the Brazilian Social Security System acts as the keeper of the Brazilian citizens' information.

The volumes involved in this Project point to a large volume of data, with a variety of types originated from many different sources. In the present day, there is a monthly payment of approximately 24 million of Benefits, with distinct Benefit Categories, each with distinct characteristics, demanding a differentiated treatment. This information is stored in very large databases and must be available to Brazilian citizens. In the universe of non-structured documents and information, studies point to a process volume of around 32,000 document movements per day, including administration processes, Benefit processes and other tasks. Considering the average of 15 pages per process, we find a total volume equivalent to 450,000 pages per day [1].

This business scenario, rich and with many nuances, brought us to delimit the context of the research, focusing on the use of documents, leading the project to the establishment of some definitions that were useful to the development of the work. We must remember that documents are means whose historical objective is to enable the communication and to provide the storage of knowledge. Therefore, these definitions were capable to lead us to an understanding of the activities that belong to the communication process, once different documents may bear several meanings, allowing many interpretations, depending on the context in which they are inserted. The concept of Information may have its meaning associated to the document itself or, still, to the knowledge related to its content, both participants of a specific domain [3].

In this direction, to overcome the challenge facing the Social Security information environment, we sought support on the process of assembling an Information Architecture. In this thematic, the development method of researches related to the understanding, to the representation of the needs of information and management of such resources, must be conducted according to the phases of [4]:

- strategic planning;
- creation;
- understanding and mapping;
- capture and collection;
- selection and treatment.

Another approach that guided the study was obtained in Choo [5], for he presents a view of the information management which supports the steps listed above. This author also suggests the activity to be seen as a net of processes that acquire, create, organize, distribute and use information according to the Procedural Model of Information Management:

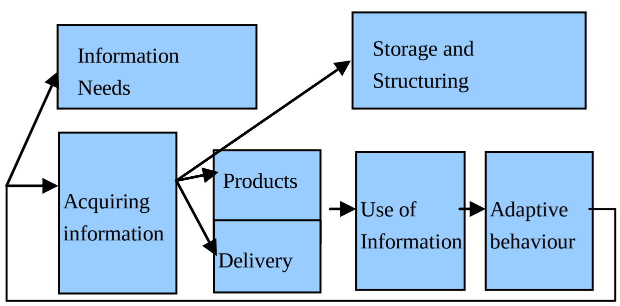


Figure 1: Procedural Model of Information Management [5]

According to a historical view, the definitions of Projects were always more related to areas of Engineering, where there were various projects for constructing buildings, bridges, planes, cars, ships and electronic equipments. Today, however, the project development theme began to create new ties and became part of other knowledge areas, contributing to the organization of new ventures in several fields of study. This view of projects was also used on the structuring of the research [1].

Therefore, we assumed that a project is a set of activities or measures planned to be executed with a determined execution responsibility in order to achieve given goals within a defined scope, with limited deadline and specific resources.

In a complementary way, we sought support on development processes of Information Systems brought by the Computer Science, where the basic cycle of the project is organized by work phases. This cycle may be developed in the phases of [6]:

- Conception;
- Elaboration;
- Construction;
- Transition.

This theoretical framework allowed the structuring of the work according to a set of phases and intermediate products that enabled the beginning of our interdisciplinary journey, seeking to overcome the barriers towards the structuring of the Brazilian Social Security System information resources.

2. UNDERSTANDING THE USER'S OBJECTS OF INTEREST

According to the interdisciplinary vision proposed for the project, we started our way through some support points identified as essential to cope with the challenge of "understanding the users' information needs". They are: the Domain Analysis, Sense-Making and Requirements Elicitation.

Within the Information Science (IS) scope, the domain analysis leads with the delimitation and comprehension of the set of information of a given context, through the understanding of the communication patterns and the relevancy. Some topics are studied in the field of the IS, such as: the use of strategies for the classification of subjects, users' controlled vocabularies and studies, may be understood as a type of Domain Analysis [7].

On the other hand, while working with domains especially within the scope of documents analysis, it is necessary to highlight the activity of the analysis of these documents subjects. According to an analytic-domain approach, a document may serve to different group of users that may use it to different purposes. Therefore, the subjects must not be identified from individual points of view, neither with general or universalized points of view, but must reflect the interests of the groups that use the information systems analyzed in the domain under study [8].

In order to assist in the domain delimitation, it was necessary to get support from other field essential to the development of the Domain Analysis: the concept of information relevance, for it is necessary to clearly identify the context's limits of the subject under analysis [9]. In this direction, relevance now has a major role in the processes of acquisition, organization, storage, preservation, communication, interaction and use of information, mainly when such activities are executed and supported by Automated Systems of Information Retrieval. According to a still preliminary vision, these systems were developed to respond with information potentially relevant to people. Thus, it is possible to discern two interacting worlds: the IT world, with a relevance of a systemic category; and the Social Science world, with a relevance of the individual perception category [10].

Another point that has helped us on the domain delimitation was the possibility of using some models to measure the relevance [11] [12], which permitted to turn into prescriptive the identification process of the domain components under analysis.

In the context of the Computer Science, however, the Domain Analysis is a method used in the development of systems and software engineering, which main object is to assist in the reuse of components of information systems. This work is developed by the Domain Analyst who tries to identify, capture, organize and represent all the relevant information of a domain. This relevant information will be used to develop systems, with the aim of making it reusable when creating new applications [12].

Pietro-Diaz explains that several types of information are generated during the process of systems development, from the requirement's analysis phase until the generation of source code programs. This information is aggregated to the domain knowledge along with the new requirements for the current system, as well as for the requirements identified for the future system. Domain specialists and analysts identify the relevant information and synthesize them. With the support of a domain engineer, the knowledge is organized and grouped in the form of domain models and with reusable components patterns and collections. One of the Domain Analysis goals is to make all this information available to be reused [12].

The vision of creating instruments to enrich the Domain Analysis approach has passed through the analysis, understanding, formulation and exteriorization of a situation. This vision led us to the incorporation in our work proposal of the approach presented by Dervin. In this direction, the interpretation and understanding of the external world must be obtained through observation, for this observation will lead us to the cognitive internal senses related to actions and attitudes. The use of the guidelines proposed by Dervin was an instrument that enabled, according to a more pragmatic view, to build the logic of a series of everyday situation that are frequently undergoing a change process. These guidelines are listed bellow [13]:

- Note the differences of opinion: find the way to think about the diversity, complexity and completeness. However, it's necessary to be careful not to simplify the domain under review neither submerge into a Babel Tower;
- Use metaphors to analyze situations: trip metaphors through periods of time and space, the experiences of the past contributing to the future, supply the differences and promote alternative paths to overcome the differences and, finally, try to estimate the results obtained;
- Try to project systems to serve the users. Think, ask and talk to the users;
- Make interviews in order to understand the identified dislocations, use metaphors to conduct such interviews with the users;
- Analyze the time-space periods, dislocating the individual attention and identifying the person-situation binomial. This relationship is also called sense-making instance. Try to apply the results and generalizations of the project to the instances not to the individuals;

The third support point proposed to this stage led us to some authors who have important works on Requirements Elicitation subject.

In Goguen's vision most of the information desired by Requirements Analysts is present and available at the social context of users and managers, and must be extracted through interviews and questionnaires. Systems, functionalities, entities and associations, must work with synergy, cooperating in order to attend the established goals to this context [14].

Kontoya and Somerville present a series of techniques to carry out the work of requirements investigation. Similarly to what was proposed by Goguen, Kontoya and Somerville also point to interviews and documentary analysis, but observe that it may be necessary to complement these interviews with the use of other investigative approaches. To this end, the authors present, among several other techniques, the construction of scenarios that simulate the interactions between users and systems, the use of prototypes to support the experimentation of such scenarios and the process of observation with a social analyses of the context (called ethnography) [15].

Still in the field of requirements investigation techniques, Robertson and Robertson observe that it is possible to use techniques to construct scenarios with the use of documentary analysis[1]. These authors also present a technique based on requirements records in snow cards or white cards, which are given to the participants of the survey meeting. At the end of these meetings the cards are collected and gathered to be analyzed with further consolidation and recording of the results [16].

Although the views presented to the needs of understanding of information do not exhaust the subject, these views have contributed to the formulation of the project and helped to understand the things of the world, delimiting the domain under analysis and allowing the use of prescriptive orientations that made this work more objective.

With the context delimited and the scope of the project established, we proceeded to the search of an structuring of the information content. In this direction, the use of the Information Architecture brought the basis to our project, for it has illuminated the path to be followed, once it presents 3 basic levels to the representation of processes linked to the cycle of information, as shown below (Figure 2):

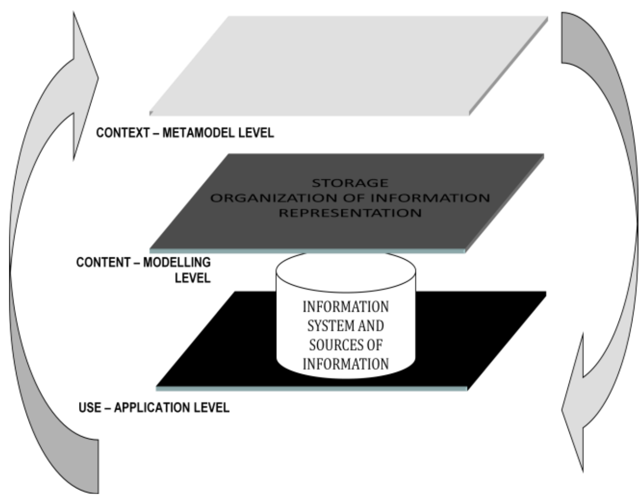


Figure 2: Information Architecture model proposal (adapted from [4])

3. DEVELOPING THE INFORMATION ARCHITECTURE IN A DOMAIN

One of the precursors of the development of the Information Architecture theme was Richard Saul Wurman. He tried to understand how information was collected, organized and presented in a meaningful way for planners, architects and engineers. These professionals are responsible for processing the information to use in projects of urban environments and also for the planning of transport channels. With the evolution and development of the subject, the Information Architecture had its operational area expanded and also began to be applied

[1] The original expression used by Robertson and Robertson is "Document Archelogy".

in order to attend users, enabling them to make use of this information organization.

We may understand that the Information Architecture concept gathers two very broad terms: "Information" and "Architecture".

The first addresses to Architectural issues, or issues of Architecture, where, since ancient times, a field of knowledge studies, plans and organizes the spaces in accordance to the requirements of its users, always seeking to work with appropriate measures and placing objects in their due space with elegance and harmony. This includes an outlook according to a perspective view, both geometric and spatial, besides counting with symmetry in relation with the whole. Lastly, according to economicity aspects of its elements and materials [17].

The second addresses to the definition of Information, which in itself gives rise to a specific study once the information is a complex object that deserves to stand out in any branch of knowledge. To develop the project the information is understood as an input to the generation of knowledge, it is the link between thinking and attitudes, and finally, a set of data endowed with relevance and purpose [9].

According to a historical view the Information Architecture theme was developed from a concept advocated by Zachman [18] and Sowa and Zachman [19], who mainly observed the use of different views in the conception of the architecture of a same product. Evernden [20] extended Zachman's concepts to the information environment, bringing different perspectives – organizational, of business and technological – and created a "big picture" to generate the Information Architecture. This map is composed of 8 factors (categories, understanding, presentation, evolution, knowledge, responsibility, process and meta-information), which are built from a diagnosis of the use of information in organizations [21].

In order to understand the organization of these characteristics and, supported by Evernden and Evernden [21], we proposed the set of recommendations presented below:
To categories:

- Try to classify and group items by similarity;
- Use differentiation to categorize elements of the domain;
- Divide big problems into smaller parcels;
- Allows organizing and structuring information.

To comprehension and understanding:

- To structure information and data from the existing understanding;
- Try to interpret and use information in innovative ways;
- Try to discover new meanings, patterns and tendencies;

To presentation:

- Try to improve the use and understanding of the information;
- Try to communicate ideas and messages;
- Try to persuade the participants by stimulating them to present scenarios of information with impact and passion;

To evolution:

- Keep everything that is relevant up to date;
- Prioritize and control changes;
- Recognize both changes and new ideas;

To knowledge:

- Codify personal knowledge as a corporative information resource;
- Apply personal experience and profiles to the information;
- Learn through feedback and practice;

To responsibility:

- Account for the changes and identify those responsible for them;
- Reconcile differences of opinion;
- Coordinate efforts being developed in the organization;

To process:

- Improve efficiency, effectiveness and productive use of information;
- Improve the value and reuse of information;
- Maximize information feedback;

To meta-information:

- Promote a language and a grammar to information management;
- Develop guidelines and patterns to improve the use of information;
- Promote the use of indexes to the use of information.

Continuing with the structuring in accordance to an Information Architecture, but bringing now the influence of technological aspects, we observe in Rosenfeld and Morville [22] the direction of their work towards the structuring and evaluation of Web environment. Observing aspects of information usability and retrieval, these authors also describe some guiding principles to the development of content projects on Websites. The outline of an Information Architecture passes, necessarily, through the organization of its elements: context, content and users, with all these elements related to the domain being reviewed. They are utilized to help in the definition of large categories that may be used in the preparation and in the framework of the Architecture, validating the propositions congregated in Evernden and Evernden.

In complementation, the guidelines pointed on Hourican's work [23] characterize the Information Architecture as a result of a work process centered in aspects of use, search and retrieval. The Information Architecture starts to develop itself towards the understanding of Structure, People, Processes and Tools.

This author also presents large groups of tools that may help in the implementation of an Information Architecture [23]:

- Content Management Systems – A system that facilitates the capture of the content, the management and publishing of different elements and applications, mainly on Web environment;
- Document Management Systems – A system that supports the Management of documents within companies. May include functionalities of document versioning and use of electronic sheets, and also controls and manages the storage and dissemination of them;
- CRM Systems – A system to manage information of people who interact with the company (costumers, suppliers, partners, etc.);
- Search engines to retrieve and search information, based on mathematical algorithms and categorization schemes, with accounting of the pages accesses;
- Portals that assemble different contents on a single and customized interface;

- E-commerce & E-business facilitating the realization of business and commerce through the internet;
- E-mails management, including search, classification, retrieval, analysis and statistics of the use of these elements.

The contributions gathered on Haverty [24] indicated that the elaboration proposal of an Architecture starts as targets and objectives of high level, the detailing of this work, however, must be conducted in an inductive way and rises from the understanding of the user information needs. Haverty also observes that the Information Architecture needs to incorporate representations by means of diagrams, as a way to facilitate the understanding of the problem.

Hence, supplied by the interdisciplinary look and with a theoretical framework, we begin the task of creating the leading thread that could guide us on this journey. Therefore, our biggest challenge was to define exactly which elements of the domain were necessary to design the Information Architecture of a company, once users will have to interact directly with many of these components.

4. RESULTS

As previously observed, with the justifications brought by these approaches, the project began to be designed and structured. The assembly of these views enabled the junction of the specific stages and objectives established in each stage [1]. We must point out that the stage of Logical Design is still ongoing and, for that matter, its actions were identified as "under development".

Stage: Project Planning
Objective: To help in the planning and investigation of the reason why this project was necessary, trying to find a clear definition about the motivation and definition of the work scope.
Actions Developed:

- Delimitation of the context, contemplating the Mission, Targets, Objectives, Sponsorship, Policies, Cultures and Technologies.
- Establishment of the responsibilities associated to the project.
- Establishment of a plan of communication.
- Identification of both the domain to be analyzed and the users.
- Definition of the approach to investigate the domain.
- Survey and analysis of the works previously developed.
- Identification of the information assets aligned with the business goals of the organization.

Stage: Specification of the Information Resources
Objective: To help in the definition of the work through the understanding of the resources that are generated, delimiting the domain and specifying with detail the elements that shall compose this context.
Actions Developed:

- Survey with detail of the domain content.
- Utilization of the information relevancy to help in the identification of the domain elements
- Definition of the criteria that will help the users in the measurement of the relevancy.
- Elaboration of a preliminary model with a graphic representation of the domain elements, as well as their associations.
- Complementation of the understanding on domain elements using approaches oriented to the users.

Stage: Logical Design of the Information Architecture
Objective: To help in the structuring and organization of the content representation, the navigation and the meta-information, in order to allow the effectively use and retrieval of the content.
Actions under development:

- Organization and classification of resources and contents.
- Identification of the responsible for the contents.
- Identification of the needs of retrieval and search engines
- Utilization of metadata to help the representation.
- Representation of the navigation and the content.
- Utilization of an Information Architecture view.

4.1 THE INTERMEDIARY PRODUCTS GENERATED

The use of Information Architecture led us to an outline of categories and relationships, which permitted the development of a project to the storage of the Information Resources of Processual Research Center and Social Information Register.

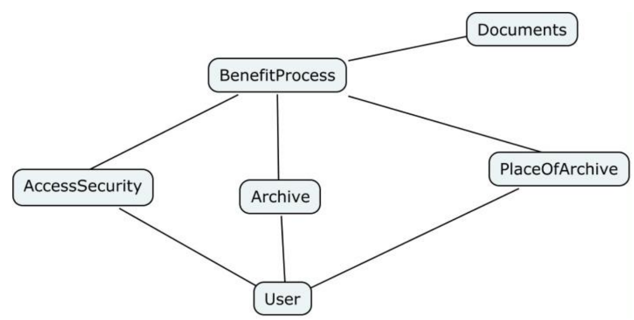


Figure 3: Information Architecture – Processual Research Center

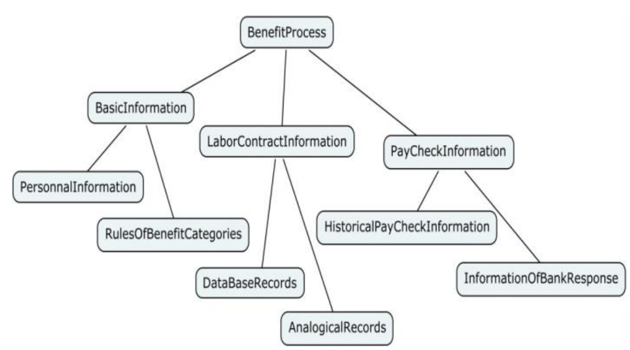


Figure 4:. Information Architecture – Social Information Register

5. FINAL CONSIDERATIONS

As presented at this report introduction, the rising of the Web environment has contributed to increase the amount of problems of Information Management, which caused a shift of focus from the resolution of the problem to the use of technological approaches. Mostly on large information volume environment, the retrieval solutions began to be structured from magical search mechanisms that proposed to bring all the relevant information. As from the implementation of this project and with the support of the theoretical contribution presented here, we hope to reduce the uncertainties of the Information Management, contributing to the process of Content Management in the Web environment, especially to the set of information of the Brazilian Social Security System.

6. REFERENCES

[1] RIBEIRO, C. J. S. **Diretrizes para o projeto de portais de informação: uma proposta interdisciplinar baseada na Análise de Domínio e Arquitetura da Informação.** 2008, UFF-Rio de Janeiro.

[2] LEVITAN, K. B. Information Resource(s) Management - IRM. **Annual Review of Information Science and Technology (ARIST)**, No. 17, p. 227-266, 1982, pp. 227

[3] HJØRLAND, B. Theory and metatheory of information science: a new interpretation. **Journal of Documentation,** Vol. 54, No. 5, 1998, pp. 606-621.

[4] LIMA-MARQUES, M.; MACEDO, F. L. O. Arquitetura da Informação: base para gestão do conhecimento. In: TARAPANOFF, K. (Ed.). **Inteligência, Informação e Conhecimento**. Brasília: IBICT/UNESCO, 2006, pp.241-255.

[5] CHOO, C. W. **A organização do conhecimento: como as organizações usam a informação para criar significado, construir conhecimento e tomar decisões.** 2a. ed. São Paulo: SENAC, 2006, pp. 404.

[6] JACOBSON, I. B., G.; RUMBAUGH, J. **The unified software development process.** Addison-Wesley Longman Publishing Co. Inc., 1999.

[7] HJØRLAND, B.; ALBRECHTSEN, H. Toward a new horizon in information science: Domain-analysis. **Journal of the American Society for Information Science,** Vol. 46, No. 6, 1995, pp. 400-425.

[8] HJØRLAND, B. THE CONCEPT OF 'SUBJECT' IN INFORMATION SCIENCE. **Journal of Documentation,** Vol. 48, No. 2, 1992, pp. 172-200.

[9] RIBEIRO, C. J. S. **Em busca da organização do conhecimento: a gestão da informação nas bases de dados da Previdência Social Brasileira com o uso da abordagem de Análise de Domínio.** 2001, UFRJ-Rio de Janeiro.

[10] SARACEVIC, T. Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science. Part II: Nature and Manifestations of Relevance. **Journal of the American Society for Information Science and Technology,** Vol. 58, No. 13, 2007, pp. 1915–1933.

[11] GREISDORF, H. Relevance thresholds: a multi-stage predictive model of how users evaluate information. **Information Processing & Management,** Vol. 39, No. 3, 2003, pp. 403-423.

[12] PRIETO-DIAZ, R. Domain analysis: an introduction. **SIGSOFT Softw. Eng. Notes,** Vol. 15, No. 2, 1990, pp. 47-54.

[13] DERVIN, B. Sense-making theory and practice: an overview of user interests in knowledge seeking and use. **Journal of Knowledge Management**, Vol. 2, No. 2, 1998, pp. 36-46.

[14] GOGUEN, J. A. Requirements engineering as the reconciliation of social and technical issues In: JIROTKA, M. e (ORGS.), J. G. (Ed.). **In: Requeriments Engineering, Social and Technical Issues**. San Diego: Academic Press, 1994.

[15] KONTOYA, G.; SOMERVILLE, I. **Requirements engineering: process and techniques**. West Sussex, England: John Willey & Sons, 1998.

[16] ROBERTSON, S.; ROBERTSON, J. **Mastering the Requirements Process**. London: ACM-Press/Addison-Wesley, 1999.

[17] VITRUVIUS. **The ten books on architecture**. New York: Dover Pub. Inc., 1960.

[18] ZACHMAN, J. A. A Framework for Information Systems Architecture. **IBM Systems Journal,** Vol. 26, No. 3, 1987, pp. 276-292.

[19] SOWA, J. F.; ZACHMAN, J. A. Extending and Formalizing the Framework for Information Systems Architecture. **IBM Systems Journal,** Vol. 31, No. 3, 1992, pp. 590-616.

[20] EVERNDEN, R. The Information FrameWork. **IBM Systems Journal,** Vol. 35, No. 1, 1996, pp. 37-68.

[21] EVERNDEN, R.; EVERNDEN, E. **Information First: Integrating Knowledge and Information Architecture for Business Advantage**. First. Burlington, 2003.

[22] ROSENFELD, L.; MORVILLE, P. **Information Architecture for the World Wide Web**. 2nd Edition. O´Reilly Media Inc, 2002.

[23] HOURICAN, R. Information Architectures - What are they? **Business Information Review,** Vol. 19, No. 3, 2002, pp. 16-22.

[24] HAVERTY, M. Information architecture without internal theory: an inductive design process. **Journal of the American Society for Information Science and Technology,** Vol**.** 53, No. 10, 2002, pp. 839-845.

In-Home Telerehabilitation as an alternative to face-to-face treatment: Feasability in post-knee arthroplasty, speech therapy and Chronic Obstructive Pulmonary Disease

Michel Tousignant
School of Rehabilitation, University of Sherbrooke
Sherbrooke, Québec, J1H 5N4, Canada

And

Mathieu Hamel & Simon Brière
Research Center on Aging
Sherbrooke, Québec, J1H 4C4, Canada

ABSTRACT

The purpose is to show three technological innovations used in in-home telerehabilitation and results regarding the efficacy pilot studies. Telerehabilitation systems in enhanced by TERAS software, externals sensors and control of the camera. Based on our experience, residential Internet network is of sufficient quality to make in-home teletreatment feasible. Innovative technologies improve teletreatment sessions. Telerehabilitation seems to be a practical alternative for home visits by a physiotherapist for dispensing rehabilitation services.

Keywords

in-home telerehabilitation, teleconsultation, elderly, total knee arthroplasty-TKA

INTRODUCTION

In-home telerehabilitation, defined as the provision of remote rehabilitation services to individuals with persistent and significant disabilities via information technologies and telecommunications in their home [1], is growing as a complementary or alternative intervention to traditional face-to-face therapy in home care and outpatient services. The rationale for in-home telerehabilitation is to expand and facilitate the delivery of rehabilitation services to people who cannot access them due to a shortage of or lack of access to services, long waiting lists for home care services or problems getting to and from the clinic [2]. Clinical care that can be provided via in-home telerehabilitation encompass active treatment and follow-up [3] as opposed to diagnosis and evaluation by teleconsultation.

PURPOSE OF THE PAPER

The purpose of this presentation is to show three technological innovations used in in-home telerehabilitation. Moreover, we present preliminary results regarding the efficacy of in-home telerehabilitation as an alternative to conventional rehabilitation services provided following an acute illness.

TECHNOLOGICAL INFRASTRUCTURE FOR TELEREHABILITATION SERVICES

Based on experience from two previous studies [4, 5], a telerehabilitation platform was developed and refine. The platform includes various components in order to provide a user-friendly experience to both the clinician and the patient at home. While being similar in many ways, two different systems were used to provide telerehabilitation services: an "in-home" system and a clinician system. The telerehabilitation platform and software interface for both systems are illustrated in Figure 1.

The core of these systems is the videoconferencing system (Tandberg 550 MXP), which uses a h.264 video codec and integrates a pan-tilt-zoom (PTZ) wide angle camera and an omnidirectional microphone. The system is mounted over a 20-inch LCD screen, which displays the video received from the other end. Audio can be played using external speakers placed on both sides of the screen (the internal LCD screen speakers are rarely sufficient to provide a satisfactory experience).

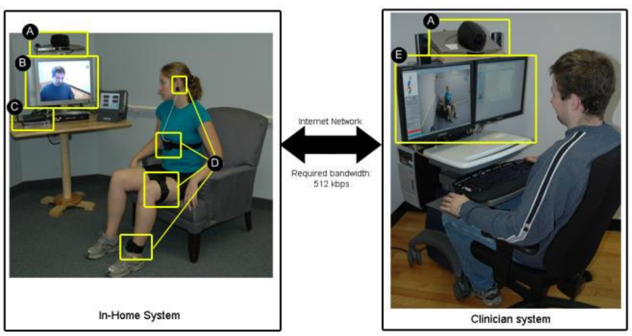


Figure 1 - Telerehabilitation systems. Components of both systems are identified: A) Videoconferencing system, B) LCD screen, C) Router and modem connecting to the internet, D) Sensors and external devices, E) Clinician computer and screen display.

Video and audio data is encrypted and transmitted over a high-speed internet connection, allowing communication using a maximum bandwidth of 512 kbps for both upload and download. The system is also resilient to packet loss and ensures that audio and video are correctly synchronized. The in-home system can include external wireless sensing devices such as oxymeters, respiratory belts, instrumented soles and inertial measurement units (figure 2). These sensors provide additional information in real time to the clinician such as oxygen saturation level, heart rate and anatomical angles. The bandwidth will vary accordingly to the numbers and type of sensors included in the setup.

The clinician system adds a computer to the in-home system. A software interface (TeRA), running on this computer, provides user-friendly control and monitoring of videoconferencing sessions, cameras control, built-in clinical tests, photo and video recording and external sensors and devices support [6]. TeRA workflow is represented in Figure 2. The platform was developed to ensure that interactions between clinicians and clients during the telerehabilitation sessions were not impeded by technology, but facilitated with user-friendly interfaces. A special effort was made to provide a mouse-based interface to intuitively control, from a unique screen through point-and-click or area zoom, PTZ camera functions at both sites. This functionality is represented in the Figure 3.

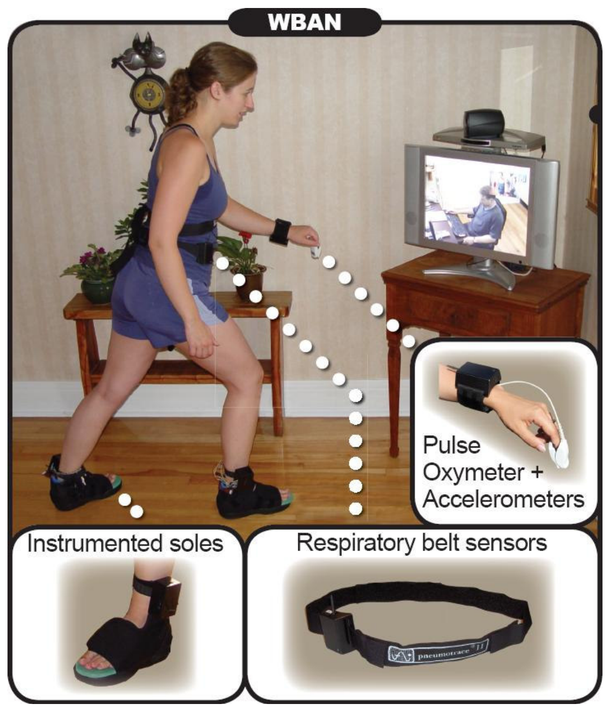


Figure 2 – Externals sensors. A computer can be included in the in-home system to accommodate several external sensors. The computer is wirelessly connected to the router and the sensor network is connected to the computer. Sensors illustrated in the picture include: inertial measurement units, respiratory belts, pulse oxymeters and instrumented soles.

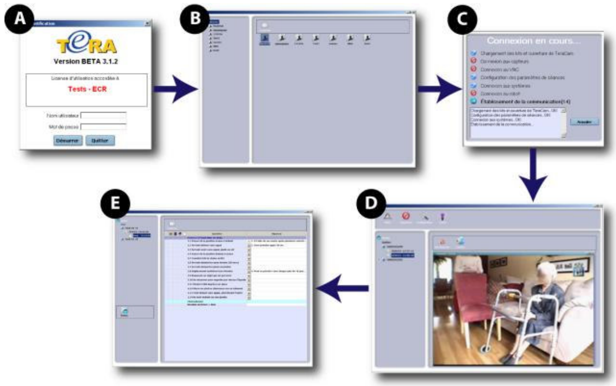


Figure 3 - TERA software. When the software is launched, the A), login screen, appears where access privileges can be controlled as each user has a unique login ID. On a successful login, the B), main screen, is displayed, where the user can select a client and easily connect to it by double-clicking the connect button, bringing in the C) interface displaying the connection process. When the connection is complete, the D) interface display the remote video and provides camera control and E) test completion directly into the software.

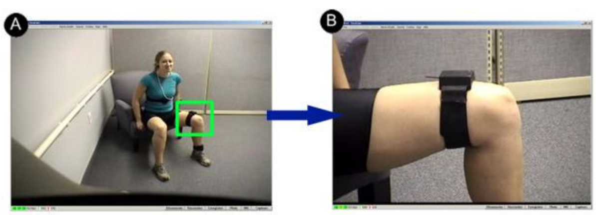


Figure 4 - Camera Control. Camera control is entirely done with the mouse. On a click, the camera centers on that point. In A), the users select an area with the mouse and release the mouse button. The camera moves and centers on the area as shown in B.

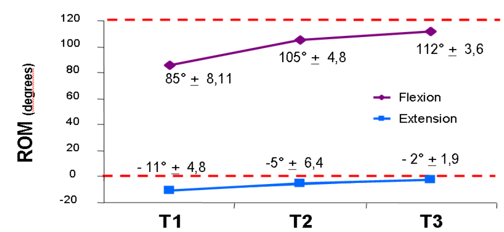
CLINICAL PILOT STUDES USING INNOVATIVE TECHNOLOGY

Post-knee arthroplasty [4]

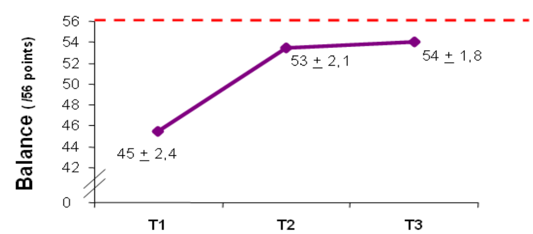
Purpose: The purpose of this study was to investigate the efficacy of in-home telerehabilitation provided following knee replacement surgery (Total Knee Arthroplasty-TKA). Setting was a service center linked to the home of the patient by high speed residential internet service.

Method: A pre/post-test design without a control group was used for this pilot study. Five community-living elders who had knee arthroplasty were recruited prior to discharge from an acute care hospital. Telerehabilitation treatments (16 sessions) were conducted by two trained physiotherapists from a service center to the patient's home. Disability (range of motion and balance) and function (locomotor performance in walking) were measured in face-to-face evaluations prior to and at the end of the treatments by a neutral evaluator. The satisfaction of the health care professional and patient was measured by the use of questionnaires.

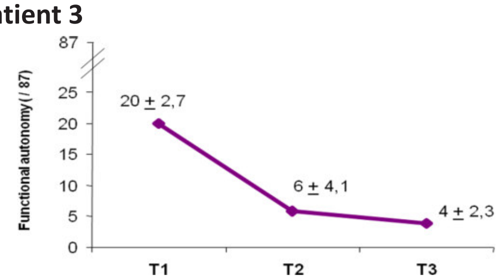
Knee range of motion



Balance



Patient 3



Results: Technology was robust despite some loss of connection in the sessions. Indeed, the satisfaction of the health care professionals regarding the technology and the communication experience during the therapy sessions was similar or slightly lower. One participant was lost during follow-up, which was not due to the technology**.** Clinical outcomes improved for all subjects and these improvements were sustained two months post-discharge for in-home telerehabilitation. The satisfaction of the participants with in-home telerehabilitation services was very high.

SPEECH THERAPY (NOT PUBLISHED)

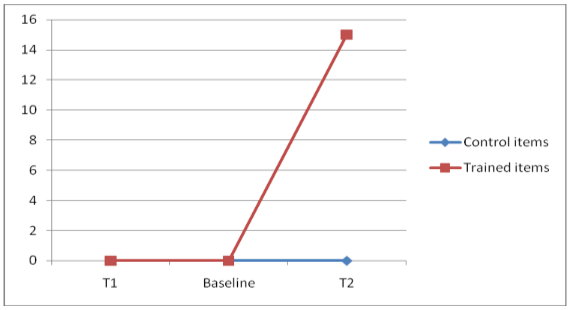
Purpose: The purpose of this study was to investigate the efficacy of speech therapy teletreatment for rehabilitation services provided following a cerebral vascular accident. The setting was a service center and a simulated home (inside the service center) using high-speed residential internet services.

Technology related to speech therapy: The telerehabilitation platform was adapted for the purpose of speech therapy treatment. Indeed, patients must react to visual cues presented by the clinicians. Therefore, an interactive computer was inserted in the platform.

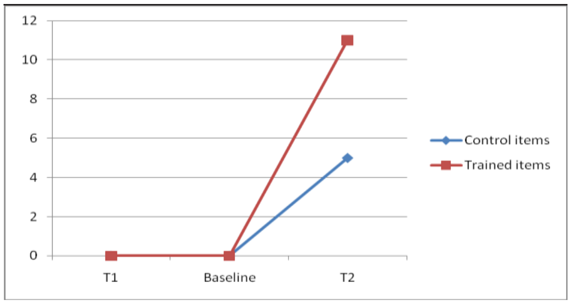
Method: The design used for this study was a pre-post test with a baseline such that the patient is his own control. The in-home speech therapy teletreatment was delivered over a period of two months. Three patients (two women and one man) who have had a cerebral vascular accident (CVA) with language problems were recruited. They were at different stages of their rehabilitation: 2, 6 and 8 months post-CVA. Subjects, depending on failed items in the recognition task at the assessment,, weredivided in two groups for the treatment: half was trained in the in-home teletreatment and the other half was not trained (control). The comparison between the number of successful items before and after treatment served as outcome measures.

Results: Despite the frailty of patients, technology was considered very satisfactory for their treatment. Clinical outcomes improved for the three subjects: they showed huge improvement of the trained items compared to the not trained items

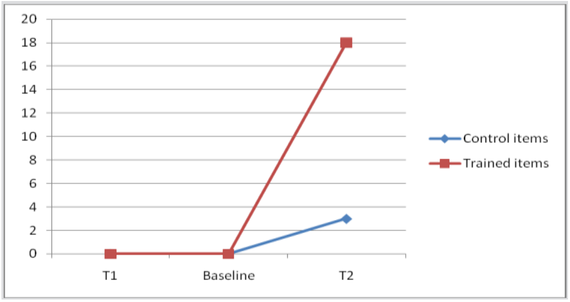
Patient 1



Patient 2



Patient 3



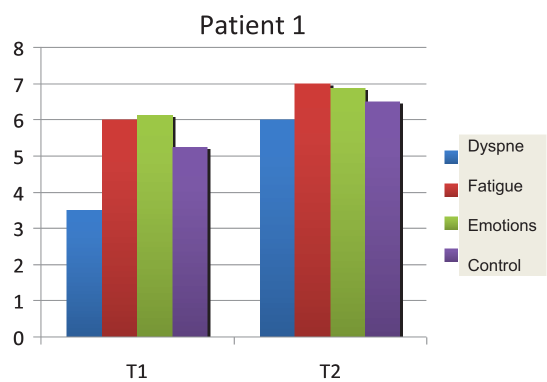
CHRONIC OBSTRUCTIVE PULMONARY DISEASE (COPD) (NOT PUBLISHED)

Purpose: The purpose of this study was to investigate the efficacy of in-home physiotherapy teletreatment for rehabilitation services provided for COPD patients.

Technology related to cardiopulmonary rehabilitation: The telerehabilitation platform was adapted for the purpose of treatment of cardiopulmonary rehabilitation. Indeed, some physiological data must be controlled in direct time, like cardiac frequency and oxygen saturation. Thus, a Nolin probe was inserted in the platform.

Method: A pre/post-test design without a control group was used for this pilot study. One community-living elder with COPD was recruited prior to discharge from his rehabilitation program. Telerehabilitation treatments (16 sessions) were conducted at the patient's home by two trained physiotherapists from the service center. Function (locomotor performance in walking) and quality of life were measured in face-to-face evaluations prior to and at the end of the treatments by an independent evaluator.

Results: Locomotor performance did not change between T1 and T2. However, the four aspects of quality of life improved over time.



CONCLUSION

A residential Internet network is of sufficient quality to make in-home teletreatment feasible. Innovative technologies improve teletreatment sessions. Telerehabilitation seems to be a practical alternative for home visits by a physiotherapist for dispensing rehabilitation services.

ACKNOWLEDGMENT

P. Boissy, H. Corriveau, H. Moffet, N. Marquis, L. Dechêsne, F. Cabana

REFERENCES

1. Cooper, R.A., et al., *Telerehabilitation: Expanding access to rehabilitation expertise.* Proceedings of the IEEE., 2001. 89(8): p. 1174-1191.
2. Wakeford, L., et al., *Telerehabilitation position paper.* American Journal of Occupational Therapy, 2005. 59(6): p. 656-60.
3. Forducey, P.G., et al., *Using telerehabilitation to promote TBI recovery and transfer of knowledge.* NeuroRehabilitation, 2003. 18(2): p. 103-11.
4. Tousignant, M., et al., *In-Home Telerehabilitation for Post-Knee Arthroplasty: A Pilot Study.* International Journal of Telerehabilitation 2009. 1: p. 9-16.
5. Tousignant, M., et al., *In home telerehabilitation for older adults after discharge from an acute hospital or rehabilitation unit: A proof-of-concept study and costs estimation.* Disability and rehabilitation Assistive technology, 2006. 1(4): p. 209-16.
6. Hamel, M., R. Fontaine, and P. Boissy, *In-home telerehabilitation for geriatric patients.* IEEE Engineering in Medicine and Biology Magazine, 2008. 27(4): p. 29-37.

Research on the Fuzzy Comprehensive Evaluation of Performance Analysis in the Data Warehouse Engineering Model Design

Yan Wang
Glorious Sun School of Business and Management, Donghua University
Shanghai City, 200051, China
and
Jiajin Le
School of Computer Science and Technology, Donghua University
Shanghai City, 201620,China
and
Dongmei Huang
College of Information Technology ,Shanghai Ocean University
Shanghai City,201306,China

ABSTRACT

This paper analyzed the model design process of data warehouse project and the various factors which affected the performance in data warehouse model design. Then the author introduced the multistage fuzzy comprehensive evaluation method, and studied its application to the data warehouse's model performance evaluation, and gave the performance measure model of data warehouse based on multistage fuzzy comprehensive evaluation method.

Keywords: Data Warehouse, Fuzzy Comprehensive Evaluation, Performance Analysis and Model Design,

1. Introduction

Data warehouse provides effective system support and information for end users' decision-making process. It is a subject-oriented data system which changes over time changing. The data in it is extracted and converted from many existing businesses data. Its feature determines the design of data warehouse is not only a business-driven engineering, but also data-driven one. Therefore, engineering design must focus on the effective extraction, synthesis, integration, and mining data from the existing database resources. In the development of data warehouse project, data-driven model design is the key point.

In practice the performance issues of data warehouse become increasingly important[1]. The huge scale of data warehouse and the fast-growing data volume lead to the strict requirements on system performance and have special significance on the control and evaluation of the performance in the data warehouse engineering. It is completely necessary that we analysis and evaluate the performance during the data warehouse model design stages.

2. Model design of data warehouse and performance analysis

Data warehouse can not maintain a constant workload. It needs to continue understanding the users' analytical requirements, and provide users with accurate and useful decision-making information, so the establishment of data warehouse is a dynamic process to meet new user demands. Large-scale and continuous expansion of data, as well as rapid growth of user workload make the performance index become an essential target to determine the quality of data warehouse[2].

In the traditional system lifecycle, there is always a lot of analysis and planning, but in the process of data warehouse design, system developers have not enough time to analysis and plan. Engineers are often required to integrate and develop data warehouse in a very short time. So system developers often have not chance to do a comprehensive performance analysis and the capability design .And it is very difficult to manage the performance in the entire data warehouse development process. As the workload is very difficult to be predicted, we often find the performance problems until the last. Therefore, we must continue evaluating and analyzing the performance during the data warehouse development phase.

Model-based data warehouse design comprises conceptual model, logical model, and physical model design. Conceptual model represents the relationship of "business information" in the real world. Logical model uses tables to store data, and it is also the relational model. Physical model involves the physical storage structure of these tables, such as the table index design. The performance analysis of data warehouse will be detailed below.

Partitioning the subject field can give a definite business value and a goal of the data warehouse.

Good Partition can cut down the unnecessary redundancy of system, reduce the relationships between various subjects within the range of possibilities and optimize performance. Subject field's public key can represent the characteristics and data links of the data warehouse, and make the inner relevance of data clear. The force or relevance linked by public key between the subject fields has important impact on data interaction and integration. At last, field attribute group definition must reflect the abstract subject field.

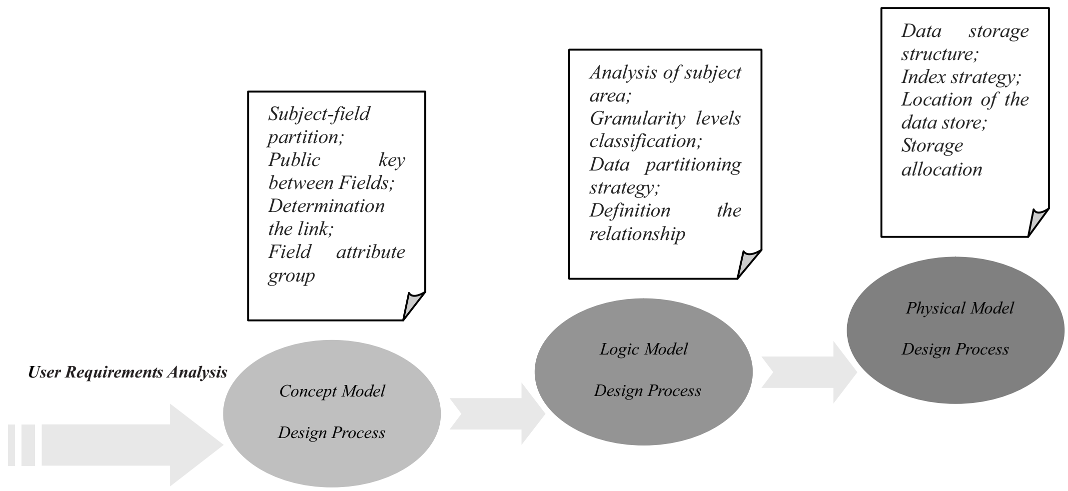


Figure1. Model design and performance factors

2.1 The conceptual model's performance factors

Conceptual model is a business model and requires analysis from business decision-makers, business domain knowledge experts and IT experts in cross-enterprise business system. Conceptual model establishes a robust model based on integration and reorganization of the data in the existing database system. So we must analyze the existing database system at first, make clear the content and organizational structure of the existing database system, and then consider how to build an effective conceptual model.

The original design documents of the database and the data relationship pattern in the data dictionary can clearly reflect the company's existing content, but a conceptual model of data warehouse is a global business-oriented model, and it provides a unified conceptual view for data integration from the various application-oriented databases.

So, determining the subject field of the model can have a good design border, and will not introduce the redundant data. The clear definitions of decision-making type can help designers make sure the best decision optimization point. At the same time, the preference of decision-makers and decision-making information sources are also important factors in the model customization.

In a conceptual model, the performance evaluation has a very vague standard. But subject field partition, public key between fields, determining the link and field attribute group are important factors which affect the performance.

2.2 Logical model's performance factors

Logic model defines the logic implementation of subject which is being loaded, and records the relevant content in the data warehouse's meta-data[3] ,that including the granularity definition, data partitioning strategy, the division of the table and the definition of the sources of data ,etc. We have to consider the performance of the following factors below.

1. Analysis of subject field

In the conceptual model, the basic subject field has been established, but data warehouse design is an iterative process, we must analysis a few basic subject field established in conceptual model, and select the first subject to implement. The primary consideration of select strategy is that subject should be large enough to be built into a system can be applied, but it must be small enough to be facilitated and implemented rapidly. If the subject is too large and complex, we can implement a meaningful subset of it. In each feedback process, it is necessary to analyze subject field. The size of the subject field is an important factor of the logical model's performance.

Secondly, subject field should be independent, though there are interactions, but it must have an independent meaning, and clear border to define data.

Finally, subject field should be of completeness, that is, all the analysis data of this defined subject can be found in this subject field. If it involves new data, new data should be added to the subject. This is an evaluative process of gradual optimization.

2. Partition of granularity level

In logical design of data warehouse it is necessary to determine the granularity levels which directly affect the amount of data and the appropriate query type. By estimating the number of data rows and the number of DASD, we can determine to choose single granularity or multi-granularity, as well as the granularity levels of partition. The appropriate levels of granularity partition contribute to the overall data warehouse performance.

Logical model design is faced with a design problem of the granularity determination that affects the size of the amount of data stored in data warehouse and the query type that the data warehouse can give. Rational determination of the granularity has a direct impact on other aspects of design. Therefore, trade-off should be made between the amount of data the size and level of detail.

3. Determining data partitioning strategy

In order to select the appropriate data partitioning criteria, the general performance factors below should be considered: The amount of data which determine whether the data segmentation should be partitioned and how to partition, the requirements of data analysis and processing which are the data partitioning standards for the analysis and processing of data partition and closely linked to the object, whether partition strategy is simple and being the easiest to implement. At last data partition must also be fitful for the granularity levels.

4. Relational mode definition

Each subject is implemented by a number of tables, and these tables link together to form a complete subject by the public key. In the conceptual model, we have identified the basic subject of the data warehouse, and describe each subject's public key keys and basic contents. In logical model the selected subject must be partitioned to form a number of tables, and the relationship mode between the tables must be defined. The relationship mode's definition effects data warehouse query and analysis performance.

2.3 The physical model's performance factors

Physical model design is to define the data storage structure and the index strategy, determine the location of the data and storage allocation. To achieve high-performance physical model, the designer must explicitly define a table index structures and index optimization, classify and merge the table reasonably, establish hardware-optimized storage structure, model time and space efficiency based on the data environment, frequency of data usage, data scale and response time requirements, optimize time and space efficiency, and make clear external storage devices. External storage equipment design involves the layout of physical storage, partition principle, block size requirements, device I / O characteristics, etc.

3. Fuzzy comprehensive evaluation on performance factor of model design

Target	First-class Indexes	Second-class Indexes
Evaluation of data warehouse model design performance(U)	Conceptual model design (U_1)	Subject-field partition; (U_{11})
		Public key between Fields(U_{12})
		Determination the link between the subject area(U_{13})
		Determination the field attribute group(U_{14})
	Logical model design (U_2)	Analysis of subject area(U_{21})
		Granularity levels classification(U_{22})
		Data partitioning strategy(U_{23})
		Definition the relationship(U_{24})
	Physical model design (U_3)	Determination the data storage structure(U_{31})
		Index strategy(U_{32})
		Location of the data store(U_{33})
		Storage allocation(U_{34})

We will quantify the performance index of data warehouse by quantitative analysis. Taking it into account that all performance factors are dynamic and most factors are difficult to describe quantitatively, but it will reduce the information asymmetry by the use of fuzz evaluation process[4]. Therefore, this paper will use fuzzy comprehensive evaluation method to quantify the various factors and evaluate the performance of the data warehouse model design process.

3.1 Establishing fuzzy evaluation sets of Data model performance index

The evaluation factors of data warehouse model design performance set can be expressed as: $U=\{U_1, U_2, \ldots\ldots, U_n\}$, the weighting set is $A=\{A_1, A_2, \ldots\ldots, A_n\}$, and A_i (i=1,2,…n) is the weighing of Ui :

$$0 \le A_i \le 1, \sum_{i=1}^{n} A_i = 1$$

$U_i=\{U_{i1}, U_{i2}, \ldots\ldots, U_{in}\}$, $(i=1, 2, \ldots\ldots, n.)$.

The weighting set is $A_i=\{a_{i1}, a_{i2}, \ldots\ldots, a_{in}\}$, where $a_{ij}(j=1,2,\ldots\ldots, n)$ is the weighing of u_{ij} ,

$$0 \le a_{ij} \le 1, \sum_{j=1}^{n} a_{ij} = 1$$

The performance indexes in Data Warehouse model design comprise the conceptual model design performance, logic model design performance, and physical model design performance. Therefore, the performance factors set of model design process are given below:

$U=\{U_1, U_2, U_3\}$= (Conceptual model design, Logical model design, Physical model design)

$U1=\{U_{11}, U_{12}, U_{13}, U_{14}, U_{15}\}$= (Determination the subject domain classification , Determination the domain's public key, Determination the link between the subject area , Determination the field attribute group)

$U2=\{U_{21}, U_{22}, U_{23}, U_{24}, U_{25}\}$= (Analysis of subject area, Granularity levels classification, Determination the data partitioning strategy, Definition the relationship between schema definitions)

$U3=\{U_{31}, U_{32}, U_{33}, U_{34}, U_{35}\}$ = (Determination the data storage structure, Determination the index strategy, Determination the location of the data store , Determination the storage allocation)

According to the above-mentioned system analysis, we structure the index system as follows:

Table 1 the index system of performance evaluation

3.2 Multi-stage mathematical model

We adopt the method of comprehensive evaluation combining AHP with FUZZY to carry on the evaluation of enterprise management environment, the procedures are as follows [3]:

(1)Get the set of the first-class Indexes: $U = (U_1, U_2, U_3)$, and their respective weights: $W = (W_1, W2, W3)$.

(2) Get the set of the second-class Indexes: $U_i = (U_{i1}, U_{i2}, \ldots, U_{in})$, and their respective weights: $W_U= (W_{U1}, W_{U2}, \ldots, W_{Un})$.

(3) the set of modal particle: suppose $V = (v_1, v_2, \ldots, v_5)$ as a set of modal particle ,which shows various grades of judgment from below to top, as shown in the following Table 2:

Table2. The distribution of the indexes judgment results

Index Grade	V_1	V_2	V_3	V_4	V_5
Index Score	[90,100]	[80,90)	[60,80)	[30,60)	[0,30)
Index Judgment	Fairly good	Good	average	bad	Quite bad

(4) Give the comprehensive evaluation to the factors of each U_i according to the basic model:

$$R_i=\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{15} \\ r_{21} & r_{22} & \cdots & r_{25} \\ \cdots & \cdots & \cdots & \cdots \\ r_{m1} & r_{m2} & \cdots & r_{m5} \end{pmatrix}$$

Meanwhile, r_{sj} means the subordinate degree of factor index r_{is} to the V_j judgment, which fulfils:

$$0 \le r_{sj} \le 1, \sum_{j=1}^{m} r_{sj} = 1 (s = 1,2,\ldots m; j = 1,2,\ldots 5)$$

(5) Give a comprehensive evaluation to the factors of Ui according to the basic model. Fuzzy matrix R_u is used to represent the fuzzy relation from U to V:

$U_i = W_i \cdot R_u = (U_1, U_2, \ldots, U_5)$

(6) Use the synthetic operation of fuzzy matrix to get A comprehensive evaluation model: $A=W_i \cdot Ru$, and $R_u = (U_1, U_2, U_3)$.

To conclude, (A, V, R) constitutes a mathematical model of fuzzy comprehensive evaluation, fuzzy matrix R converts fuzzy sets U and U_i into fuzzy set A, and then A is the fuzzy comprehensive evaluation for the target.

(7)To suppose $F = (f_1, f_2 \ldots f_5)^T$, F is a set of scores and a vertical vector, which means the standard scores of each factor in the set V.

(8)We can calculate the score Z of performance of data warehouse by the multiplication of the vectors, Z is the algebraic value: $Z=A\cdot F$. Then we can appraise enterprise performance of data warehouse according to the valve of F.

3.3 Data and empirical analysis

Taking a practical data warehouse project as an example, we choose some relevant data warehouse experts, senior engineer from a software company in China to score the performance. We suppose that full marks are '100',then divide and arrange them on five grades according to their scores, every experts gives the score on every item, the evaluation matrix of all factors is as following:

$$R=\begin{pmatrix} 1/6 & 2/3 & 1/6 & 0 & 0 \\ 0 & 0 & 5/6 & 1/6 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 1/6 & 2/3 & 1/6 \end{pmatrix}$$

In the matrix, r_{sj} =the total score of j grade of s item / the number of scoring persons

The weights of the sub factors can be obtained by AHP, the result is as follows:

$A_1 = (0.75, 0.25), A_2 = (0.22, 0.36, 0.09, 0.33)$

The subordinate-degree vector to the modal set V is as follows:

$A = (0.59, 0.41)$

The second layer fuzzy evaluation as follows:

$$B_1 = A_1 \circ R = (0.75, 0.25) \circ \begin{pmatrix} 0.23 & 0.17 & 0.23 & 0.37 & 0 \\ 0.17 & 0.43 & 0.32 & 0.08 & 0 \end{pmatrix} = (0.22, 0.24, 0.25, 0.29, 0)$$

$$B_2 = A_2 \circ R = (0.22, 0.36, 0.09, 0.33) \circ \begin{pmatrix} 1/6 & 2/3 & 1/6 & 0 & 0 \\ 0 & 0 & 5/6 & 1/6 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 1/6 & 2/3 & 1/6 \end{pmatrix} = (0.07, 0.18, 0.42, 0.28, 0.06)$$

And the first layer fuzzy evaluation as follows:

$$B = A \circ R = A \circ \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

$$= (0.59, 0.41) \circ \begin{bmatrix} 0.22 & 0.24 & 0.25 & 0.29 & 0 \\ 0.17 & 0.18 & 0.41 & 0.28 & 0.06 \end{bmatrix}$$

$$= (0.16, 0.22, 0.31, 0.29, 0.02)$$

Then the value of performance evaluation is as follows:

$B = (0.16, 0.22, 0.31, 0.29, 0.02)$

Suppose the score set :F=（100,85,70,55,40）^{T}，then the value of performance evaluation is as follows:

$Z= A \circ F = (0.16, 0.22, 0.31, 0.29, 0.02) \cdot (40,55,70,85,100)^T= 66.85$

Decision-makers can compare the performance results of the project with the critical values of its performance indexes to determine whether they should endorse the model design activities. The above results prove that the management environment of this enterprise is in the average level. At the same time, the evaluation model weight parameter should be adjusted for different types of the data warehouse model design.

4. Conclusions

As data warehouse model design performance evaluation indexes, the evaluation results obtained by fuzzy comprehensive evaluation method can make the performance analysis clear during the model design phase. At the same time, evaluation results may be compared with the actual performance values after data warehouse structures is finished and the expert score weights of the fuzzy comprehensive evaluation can be adjusted.

5. Acknowledgment

The study is funded by key project of Shanghai Scientific Committee, subject number: 08dz1204802 .

6. References

[1] Gollarelli M,Maio D,Rizzi S, Italy, **Applying Vertical Fragmentation Techniques in Logical Design of Multidimensional Database**,2000

[2] Materialized Views, Schrefl M, Thalhammer T, Australia, Austria, **On Making Data Warehouse Acrive**,2000

[3] Paulraj Ponniah , **Data Warehousing Fundamentals**[M] Beijing: Publishing House of Electronics Industry,2004.pp.25-39

[4] Li Hongxing, Wang Peizhuang, **Fuzzy Mathematics** [M].Beijing: National defense industry Press ,1993.pp.85-112

Remediation of Soil Contaminated with Crude Oil using Supercritical CO2

Adel A. Azzam, Ali H. Al-Marzouqi*, Abdulrazag Y. Zekri

Chemical & Petroleum Engineering Department,
U.A.E. University, Al-Ain, P.O. Box: 17555, U.A.E.

* Corresponding author. E-mail address: hassana@uaeu.ac.ae, Fax: +9713-7624262

ABSTRACT

Soil and sediment contamination with hydrocarbons is an environmental concern, which demands for more efficient remediation techniques. Pure and modified supercritical carbon dioxide (SC CO_2) was used for the extraction of petroleum hydrocarbons from soil contaminated with crude oil. Effect of CO_2 flow rate (1 and 4 ml/min), temperature (80 and 160 °C), pressure (250 and 350 bar), and addition of 5% (v/v) organic solvent (heptane or toluene) on the extraction efficiency and on the composition of extracted hydrocarbons were investigated.

The maximum extraction efficiency (92.26%) was obtained at 80 °C and 350 bar corresponding to a modified CO_2 with 5% (v/v) heptane. Extraction efficiency of CO2 increased with pressure and decreased with temperature. Chemical modification of CO2 by adding heptane increased the extraction efficiency. Analysis of the soil after the extraction process shows that pure SC CO_2 was able to remove up to 92.86% of TPH in the contaminated soil. In addition, a significant reduction in PAH level was observed. Supercritical fluid extraction proved to be an efficient method for the remediation of hydrocarbon-contaminated soil.

Keywords: Remediation, Contaminated Soil, Crude Oil, Supercritical CO2

1. INTRODUCTION

Soil contamination with crude oil and petroleum products is often observed at industrial sites, causing environmental pollution, which can be hazardous to the health of plants, animals, and humans [1-4]. The hydrocarbon molecules may contain hazardous complex chemical mixtures such as total petroleum hydrocarbons (TPH), polychlorinated biphenyls (PCBs), and polycyclic aromatic hydrocarbons (PAHs). Removal of such compounds from contaminated sites is an important and challenging problem. The most important and widely used remediation methods are incineration, thermal desorption, biological remediation, chemical treatment and solvent extraction [5]. Conventional techniques such as landfill disposal, thermal desorption, incineration and liquid solvent extraction are expensive and involve risks associated with air and residual pollution. Biological remediation is a rather slow process, with possible logistic and practical disadvantages.

Despite great efforts and expenditure of resources to develop both technically and economically effective cleanup processes for contaminated soils, no widely accepted method has been found and further research is still needed. New methods are therefore being investigated in order to improve the remediation efficiency and lower the costs or the remediation time. Since three decades ago, supercritical fluids (SCFs) have been used as extraction media to remove various types of substances from solid matrices. The unique properties of SCFs that make them technically attractive are their enhanced ability to dissolve organic compounds, an ability, which can be easily tuned by changing temperature and/or pressure, thus changing the fluid properties from gas-like to liquid-like. Such properties allow the SCFs to dissolve and carry away materials like a liquid but also enter very small pores like a gas. The most popular fluid is supercritical carbon dioxide (SC CO_2) because it is non-toxic, non-flammable, chemically stable, readily available, inexpensive, environmentally acceptable, and can easily be separated from the products. Although SCF technology has been successfully realized for environmental remediation in the laboratory, its commercialization still lacks the significant technological improvement required in order to reach economic feasibility. Like other new technologies, SFE technology, utilizing CO_2 as a fluid has its specific problems. One of these problems is the limited ability of SC CO_2 to dissolve and separate polar or high molecular weight organic compounds even at very high densities. To increase the efficiency of the SFE process for such compounds, the selectivity and solubilizing power of SC CO_2 can be enhanced by the addition of polar organic compounds, known as modifiers.

Significant research has been carried out in order to study various aspects of contaminant removal by SC CO_2. Comprehensive presentations of various aspects on the use of this technology for extraction purpose are available in several critical reviews [1,6-8] and hundreds of other scientific articles. Supercritical CO_2 has been successfully used for extracting a variety of organic compounds such as polycyclic aromatic hydrocarbons (PAHs) [9-12], polychlorinated biphenyls (PCBs) [1,11,13-16], pesticides [17-18], and hydrocarbons [11,19-23]. However, data for CO_2 extraction at extremely high pressures and temperatures are scarce in the literature, especially for soil contaminated with crude oil. Al-Marzouqi et al. [23] showed that SC CO_2 at 300 bar and 120 °C is able to extract about 70% of hydrocarbons from a typical UAE soil contaminated with crude oil. The objective of the present study was to investigate the ability of pure and modified CO_2 under supercritical conditions to remediate soil contaminated with crude oil and achieve higher extraction efficiencies.

2. EXPERIMENTAL

Materials

Carbon dioxide (purity of 99.995%) was supplied by Abu Dhabi Oxygen Company. Crude oil (average molecular weight = 281.5 g/mole and density = 0.8634 g/ml) was obtained from Bu Hasa oil field (Abu Dhabi, UAE). The chemical modifiers (n-heptane and toluene) and the organic solvents (dichloromethane and methanol) were of analytical grade with purity ≥99% and were supplied by Sigma Aldrich. Soil samples (bulk density = 1.6 g/ml and average particle size = 150 μm) were collected from Sahel oil field in the UAE. The porosity and permeability of the soil were 35% and 20.15 Darcy, respectively.

Experimental design

Extraction of hydrocarbons with SCFs from contaminated soil was carried out by following the full factorial experimental design with four factors: pressure (250 and 350 bar), temperature (80 and 160 °C), flow rate (1 and 4 ml/min) and fluid type (pure SC CO_2, modified SC CO_2 with 5% (v/v) toluene and modified SC CO_2 with 5% (v/v) n-heptane). Each experiment was repeated twice, resulting in a total number of 48 experiments. Experiments were run in random order to eliminate various types of biases due to uncontrolled nuisance factors. The statistical analysis was performed using the statistical package SPSS (SPSS inc., Version 15.0). All the statistical analyses of the effects of variables on the extraction efficiency were performed using a multi-way analysis of variance (ANOVA) with two replications per cell.

Experimental apparatus

The experimental setup consisted of a 260-ml capacity syringe pump and a controller system (ISCO 260D), a 100-ml stainless steel extraction chamber (DBR-JEFRI 100-10-BE), and a cold trap as described earlier (Al-Marzouqi et al., 2007). The extraction chamber was kept in an air-circulating oven (Memmert ULE 400) with a temperature control ranging from 30-250 °C. Pressure within the extraction chamber was measured and controlled by the ISCO system. A micrometering valve (HIP 15-12AF1-V) was used as the expansion valve at the exit of the extraction chamber to achieve a good control of the flow rate. Circulating methanol at -15 ^oC was used as a cold trap to separate CO_2 from other components of the mixture.

Experimental procedures

Soil samples were spiked with 10 w/w% crude oil and placed in the extraction chamber. The extraction chamber was kept in the oven at the desired temperature until thermal equilibrium was reached (30-60 min). The chamber was then pressurized with CO_2 to the desired pressure and kept for another 30 minutes to reach equilibrium. In the case of modified CO_2, the second syringe pump was used to deliver the cosolvent (heptane or toluene), which was mixed with the CO_2 stream at desired ratio. Pure and modified carbon dioxide at supercritical condition was then added to the ISCO SCF Extraction system (SFX system) and equilibrated for about 15 minutes. The SCF was allowed to flow through the coil of tubing and enter the extraction chamber from the bottom. The fluid was equilibrated with the spiked soil sample for at least 30 minutes. The supercritical solution was then allowed to flow into a vial and the extract was separated from the supercritical fluid by depressurizing the system in the cold trap. The residual hydrocarbons in the soil, after SFE process, were also analyzed for concentration of total petroleum hydrocarbon (TPH) and polycyclic aromatic hydrocarbons (PAHs).

3. RESULTS AND DISCUSSIONS

The CO_2 extraction efficiency (the ratio of extracted hydrocarbons to the initial amount of crude oil in place) is used throughout this study to evaluate the capacity of CO_2 to extract hydrocarbons from the soil. The average extraction efficiencies obtained at each of the investigated operating conditions are tabulated in Table 1. The lowest value of extraction efficiency (68.38% ± 1.99) was obtained for modified SC CO_2 (with an addition of 5 % toluene) at 250 bar and 160 °C, while the maximum efficiency (92.26 % ± 5.40) was found for SC CO_2 (with an addition of 5 % n-heptane) at 350 bar and 80 °C. The highest efficiency obtained by SC CO_2 alone (without modifier) was 78.51 % ±0.46, which was obtained at 350 bar and 160 °C. The complexity of crude oil mixture containing many compounds with significantly different physico-chemical properties that vary with temperature and pressure are believed to cause such a large variation in the extraction capacity of SC CO_2.

Results of the multi-way ANOVA based on the original values of extraction efficiency show that temperature, pressure and fluid type have significant effect on the extraction efficiency, but the flow rate of the CO_2 does not have a significant effect, i.e. Sig. >0.05. Moreover, pressure and fluid type interact. This means that the effect of pressure depends on which fluid is used and vice versa, which is not the case with temperature. However, by checking the validity of the ANOVA model using residual analysis, the normality assumption was found to be satisfied, i.e. the p-value was higher than 5%.

Effect of temperature

Figure 1 illustrates the effect of temperature on the extraction efficiency. Values on the figure (including bars showing the standard error of the mean) represent the mean value of extraction efficiency for 24 experiments at each temperature. Results indicate that temperature has an inverse effect on the extraction efficiency. This might be due to the increase in the kinematic viscosity and interfacial tension due to the decrease in CO_2 density with an increase in temperature.

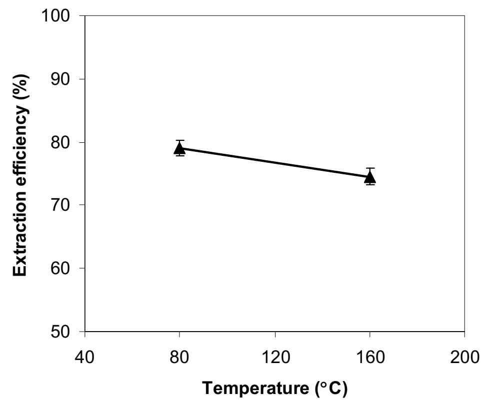


Figure 1. Effect of temperature on the extraction efficiency.

Effect of flow rate

Effect of flow rate (1 and 4 ml/min) on the extraction efficiency is shown on Figure 2. Values on the figure represent the mean value of the extraction efficiency for 24 experiments at each flow rate. Decreasing the flow rate usually ensures more contact time and results in higher extraction efficiencies for a given amount of CO_2 used. However, saturation is achieved at certain flow rates, below which the flow rate does not affect the extraction efficiency of the solvent. Results indicate that flow rate does not affect the extraction efficiency for the conditions used in this study. Therefore, the extraction process should be operated at 4 ml/min in order to reduce the extraction time.

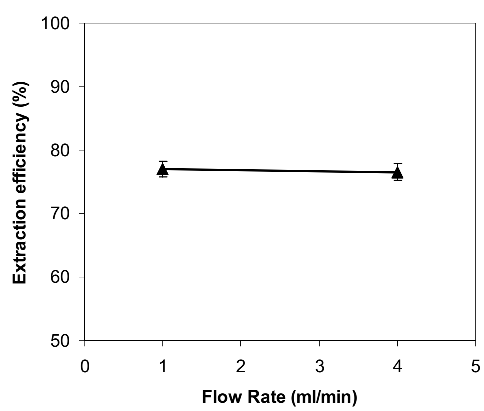


Figure 2. Effect of CO_2 flow rate on the extraction efficiency.

Table 1. Properties and average extraction efficiencies of supercritical fluids for soil samples contaminated with crude oil.

Temperature (°C)	Pressure (bar)	CO_2 flow rate (ml/min)	Modifier 5% (v/v)	CO_2 density (g/ml)	CO_2 viscosity (µPa·s)	CO_2 kinematic viscosity × 10^8 (m^2/s)	Average extraction efficiency (%) ± SEM*
80	250	1	–	0.68622	56.03	8.17	72.32 ± 0.49
80	250	4	–	0.68622	56.03	8.17	75.07 ± 0.92
80	350	1	–	0.78897	70.376	8.92	77.76 ± 0.78
80	350	4	–	0.78897	70.376	8.92	77.40 ± 0.55
160	250	1	–	0.39294	33.905	8.63	68.44 ± 0.43
160	250	4	–	0.39294	33.905	8.63	69.03 ± 1.47
160	350	1	–	0.52948	43.726	8.26	78.51 ± 0.46
160	350	4	–	0.52948	43.726	8.26	77.91 ± 0.37
80	250	1	n-Heptane	–	–	–	80.40 ± 2.96
80	250	4	n-Heptane	–	–	–	79.51 ± 2.99
80	350	1	n-Heptane	–	–	–	92.26 ± 5.40
80	350	4	n-Heptane	–	–	–	87.68 ± 1.20
160	250	1	n-Heptane	–	–	–	73.03 ± 2.18
160	250	4	n-Heptane	–	–	–	78.23 ± 4.66
160	350	1	n-Heptane	–	–	–	85.07 ± 0.55
160	350	4	n-Heptane	–	–	–	82.91 ± 5.02
80	250	1	Toluene	–	–	–	76.05 ± 2.58
80	250	4	Toluene	–	–	–	71.65 ± 1.43
80	350	1	Toluene	–	–	–	77.63 ± 3.30
80	350	4	Toluene	–	–	–	81.01 ± 0.56
160	250	1	Toluene	–	–	–	72.35 ± 3.02
160	250	4	Toluene	–	–	–	68.38 ± 1.99
160	350	1	Toluene	–	–	–	70.38 ± 0.15
160	350	4	Toluene	–	–	–	70.13 ± 2.38

*SEM: Standard Error of the Mean

Effect of pressure and fluid type

Due to the interaction between pressure and fluid type, effect of these parameters cannot be shown separately, therefore, Figure 3 shows the effect of both pressure and the fluid type on the extraction efficiency. Each point on Figure 3 represents the mean value of extraction efficiency for 8 experiments for each fluid type at a given pressure. As shown in the figure, the extraction efficiency of pure and modified SC CO_2 increases as the pressure is increased. This might be due to the decrease in the kinematic viscosity due to the increase in CO_2 density with an increase in pressure. Moreover, the extraction efficiency of the modified SC CO_2 by 5% (v/v) heptane is higher than that of both pure SC CO_2 and modified SC CO_2 with 5% (v/v) toluene. The higher extraction efficiency when utilizing heptane can probably be attributed to the richness of Bu Hasa crude oil in aliphatic non-polar hydrocarbon compounds such as n-alkanes (C_6-C_{22}) as reported by Al-Marzouqi et al. (2007). However, due to the interaction between pressure and fluid type, the extraction efficiency of modified SC CO_2 with 5% (v/v) toluene is found to be higher than that for pure SC CO_2 at the low pressure (250 bar) but lower at the high pressure (350 bar).

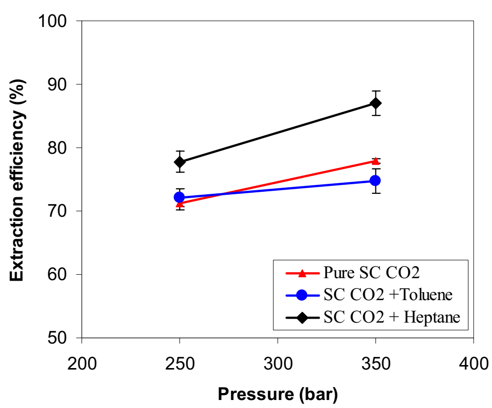


Figure 3. Effect of pressure and fluid type on the extraction efficiency.

Analysis of total petroleum hydrocarbon (TPH)

The capacity of pure SC CO_2 to extract TPH from soil saturated with Bu Hasa crude oil was investigated for selected runs (Table 2). As shown in the table, pure SC CO_2 at high pressure (350 bar) and low temperature (80 °C), is capable of extracting 92.86% of TPH from the polluted soil compared to 90.98% removal of TPH at the same pressure and higher temperature (160 °C). Removal percentage was less at the lower pressure of 250 bar (83.54% and 76.15% at 80 and 160 ^{o}C, respectively), which matches the results obtained from the extraction efficiency of SC CO_2. This study shows that pure SC CO_2 can effectively remediate the contaminated soil and thus reduce the harmful effects of the TPH compounds on the environment.

Analysis of polycyclic aromatic hydrocarbons (PAHs)

The PAHs measurement was conducted for selected runs to investigate the efficiency of SC CO_2 in extracting PAHs from soil samples contaminated with Bu Hasa crude oil. Concentration of 16 PAHs in the selected soil samples after the SFE process is tabulated in Table 3. Results show that the modified SC CO_2 with 5% (v/v) heptane at low temperature (80 °C) and high pressure (350 bar) was not able to completely remove some of the PAHs from the contaminated soil. Also, the extraction by pure SC CO_2 at the same pressure and temperature was the worst among all other conditions. However, pure SC CO_2 at 160 °C and 350 bar resulted in a better extraction of the 16 PAHs. This might be attributed to the effect of high temperature, which increases the volatility of the PAHs and thus increases their solubility in the fluid.

4. CONCLUSIONS

Effects of temperature, pressure, CO_2 flow rate and two modifiers (heptanes and toluene) at 5% (v/v) on the extraction capacity of SC CO_2 were investigated. The results of this study indicate that SC CO_2 is an effective solvent, which leads to high extraction efficiencies when applied at high pressures. Furthermore, the results from this study show that the flow rate does not have a significant effect on the efficiency of SC CO_2. Therefore, it is recommended to use the high flow rate, i.e. 4 ml/min, in order to reduce the time required for the remediation of contaminated soil. Moreover, the temperature, i.e. 80 and 160 °C, has no significant effect on the extraction efficiency of SC CO_2 at the high pressure (350 bar). Therefore, it is recommended to apply the low temperature during the extraction process in order to save energy. Chemical modification of CO_2 by adding 5% heptane was more effective than the same level of modification by toluene. The optimum condition to extract hydrocarbons from soil contaminated with Bu Hasa crude oil was by modified SC CO_2 with 5% heptane at high pressure (350 bar), low temperature (80 °C), and flow rate of 1 ml/min. Supercritical CO_2 was able to remove 92.86% of the TPH present in contaminated soil. Additionally, pure SC CO_2 and SC CO_2 chemically modified with 5% (v/v) heptane were capable of significantly reducing the concentration levels of PAHs in the soil contaminated by Bu Hasa crude oil.

Table 2. TPH analysis of the clean soil, soil spiked with crude oil before SFE and treated soil after the SFE process.

Sample	**SFE Temperature (°C)**	**SFE Pressure (bar)**	**TPH (µg/mg)**	**TPH Removal (%)**	**Extraction Efficiency (%)**
Clean soil	–	–	< 0.23	–	–
Spiked soil with crude oil before SFE	–	–	56875	–	–
Treated soil after SFE	80	350	4057	92.86	78.69
	160	250	13564	76.15	69.22
	160	350	5129	90.98	77.95
	80	250	9361	83.54	71.83

Table 3. PAHs analyses of the clean soil, spiked soil with crude oil before SFE and treated soil after the SFE process. Removal efficiencies (%) are shown in parenthesis. Removal efficiency was assumed 100% for PAH concentration < LOD*.

Sample		Clean soil	Spiked soil with crude oil before SFE	Treated soil after SFE				
Temperature (°C)		–	–	80	160	160	80	80
Pressure (bar)		–	–	350	250	350	250	350
Modifier		–	–	–	–	–	–	Heptane
PAH (µg/kg)	Naphthalene	<7.89	10648	<7.89 **(100%)**	<7.89 **(100%)**	<7.89 **(100%)**	<7.89 **(100%)**	78 **(99.26%)**
	Acenaphthylene	<10.7	<10.7	<10.7	<10.7	<10.7	<10.7	<10.7
	Acenaphthene	<5.12	3260	7.89 **(99.75%)**	16.8 **(99.48%)**	15.4 **(99.52%)**	16 **(99.50%)**	19.5 **(99.40%)**
	Flourene	<5.53	357	<5.53 **(100%)**	<5.53 **(100%)**	<5.53 **(100%)**	<5.53 **(100%)**	<5.53 **(100%)**
	Phenanthrene	<4.85	10417	279 **(97.32%)**	66.8 **(99.35%)**	75.1 **(99.27%)**	553 **(94.69%)**	292 **(97.19%)**
	Anthracene	<4.99	<4.99	<4.99	<4.99	<4.99	<4.99	<4.99
	Fluoranthene	<4.98	947	42.3 **(95.53%)**	32.3 **(96.58%)**	<4.98 **(100%)**	8.19 **(99.13%)**	40.2 **(95.75%)**
	Pyrene	<5.00	3921	924 **(76.43%)**	274 **(93.01%)**	63.1 **(98.39%)**	393 **(89.97%)**	622 **(84.13%)**
	Benzo(a)anthracene	<4.90	1168	<4.90 **(100%)**	9.53 **(99.18%)**	11.4 **(99.02%)**	<4.90 **(100%)**	9.85 **(99.15%)**
	Chrycene	<4.92	1107	9.85 **(99.11%)**	<4.92 **(100%)**	10.3 **(99.06%)**	<4.92 **(100%)**	10.8 **(99.02%)**
	Benzo(b)flouranthene	<4.54	<4.54	<4.54	<4.54	<4.54	<4.54	<4.54
	Benzo(k)flouranthene	<4.61	<4.61	<4.61	<4.61	<4.61	<4.61	<4.61
	Benzo(a)pyrene	<4.99	<4.99	<4.99	<4.99	<4.99	<4.99	<4.99
	Dibenzo(a,h)anthracene	<5.34	283	<5.34 **(100%)**	<5.34 **(100%)**	<5.34 **(100%)**	<5.34 **(100%)**	<5.34 **(100%)**
	Benzo(g,h,i)perylene	<5.45	750	37 **(95.06%)**	<5.45 **(100%)**	<5.45 **(100%)**	<5.45 **(100%)**	13.8 **(98.16%)**
	Indeno(1,2,3-cd)pyrene	<5.42	326	36.2 **(88.89%)**	<5.42 **(100%)**	<5.42 **(100%)**	<5.42 **(100%)**	<5.42 **(100%)**
Extraction efficiency (%)		**–**	**–**	**78.69**	**69.22**	**77.95**	**71.83**	**97.66**

* LOD: limit of detection.

5. REFERENCES

[1] G. Anitescu, and L.L. Tavlarides, "Supercritical extraction of contaminants from soils and sediments", **The Journal of Supercritical Fluids**, Vol. 38, No. 2, 2006, pp. 167–180.

[2] E.W. Liebeg and T.J. Cutright, "The investigation of enhanced bioremediation through the addition of macro and micronutrients in a PAH contaminated soil", **International Biodeterioration and Biodegradation**, Vol. 44, 1999, pp. 55–64.

[3] D. Sarkar, M. Ferguson, R. Datta, S. Birnbaum, "Bioremediation of petroleum hydrocarbons in contaminated soils: comparison of biosolids addition, carbon supplementation, and monitored natural attenuation", **Environmental Pollution**, Vol. 136, No. 1, 2005, pp. 187–195.

[4] N. Vasudevan and P. Rajaram, "Bioremediation of oil sludge contaminated soil", **Environment International**, Vol. 26, No. 5–6, 2001, pp. 409–411.

[5] S. Paria, "Surfactant-enhanced remediation of organic contaminated soil and water", **Advances in Colloid and Interface Science**, Vol. 138, 2008, pp. 24–58.

[6] P.K. Wong and J. Wang, "The accumulation of polycyclic aromatic hydrocarbons in lubricating oil over time - a comparison of supercritical Fluid and liquid-liquid extraction methods", **Environmental Pollution**, Vol. 112, 2001, pp. 407–415.

[7] B.T. Bogolte, G.A.C. Ehlers, R. Braun, A.P. Loibner, "Estimation of PAH bioavailability to Lepidium sativum using sequential supercritical fluid extraction e a case study with industrial contaminated soils", **European Journal of Soil Biology**, Vol. 43, 2007, pp. 242–250.

[8] A.S. Pimenta, B.R. Vital, J.M. Bayona, R. Alzaga, "Characterisation of polycilic aromatic hydrocarbons in liquid products from pyrolysis of Eucalyptus grandis by supercritical fluid extraction and GC/MS determination", **Fuel**, Vol. 77, No. 11, 1998, pp. 1133–1139.

[9] C. Lutermann, W. Dott, J. Hollender, "Combined modifier/in situ derivatization effects on supercritical fluid extraction of polycyclic aromatic hydrocarbons from soil", **Journal of Chromatography A**, Vol. 811, 1998, pp. 151–156.

[10] S.B. Hawthorne, C.B. Grabanski, E. Martin, D.J. Miller, "Comparisons of Soxhlet extraction, pressurized liquid extraction, supercritical fluid extraction and subcritical water extraction for environmental solids: recovery, selectivity and effects on sample matrix", **Journal of Chromatography A**, Vol. 892, No. 1–2, 2000, pp. 421–433.

[11] P. Hallgren, R. Westbom, T. Nilsson, S. Sporring, E. Björklund, "Measuring bioavailability of polychlorinated biphenyls in soil to earthworms using selective supercritical fluid extraction", **Chemosphere**, Vol. 63, No. 9, 2006, pp. 1532–1538.

[12] S. Sporring, S. Bøwadt, B. Svensmark, E. Björklund, "Comprehensive comparison of classic Soxhlet extraction with Soxtec extraction, ultrasonication extraction, supercritical fluid extraction, microwave assisted extraction and accelerated solvent extraction for the determination of polychlorinated biphenyls in soil", **Journal of Chromatography A**, Vol. 1090, No. 1–2, 2005, pp. 1–9.

[13] T. Nilsson and E. Björklund, "Selective supercritical fluid extraction as a tool for determining the PCB fraction accessible for uptake by chironomid larve in a limnic sediment", **Chemosphere**, Vol. 60, No. 1, 2005, pp. 141–146.

[14] W. Zhou, G. Anitescu, L.L. Tavlarides, "Polychlorinated biphenyl (PCB) partition equilibrium between St. Lawrence River sediments and supercritical fluids", **The Journal of Supercritical Fluids**, Vol. 29, No. 1–2, 2004, pp. 31–42.

[15] E. Reverchon and I. De Marco, "Supercritical fluid extraction and fractionation of natural matter", **The Journal of Supercritical Fluids**, Vol. 38, No. 2, 2006, pp. 146–166.

[16] C. Gonçalves, J.J. Carvalho, M.A. Azenha, M.F. Alpendurada, "Optimization of supercritical fluid extraction of pesticide residues in soil by means of central composite design and analysis by gas chromatography–tandem mass spectrometry", **Journal of Chromatography A**, Vol. 1110, No. 1–2, 2006, pp. 6–14.

[17] C. Quan, S. Li, S. Tian, H. Xu, A. Lin, L. Gu, "Supercritical fluid extraction and clean-up of organochlorine pesticides in ginseng", **The Journal of Supercritical Fluids**, Vol. 31, No. 2, 2004, pp. 149–157.

[18] R. Jaffé, D. Diaz, K.G. Furton, E. Lafarg, "High Temperature Supercritical Carbon Dioxide Extractions of Geological Samples: Effects and Contributions from the Sample Matrix", **Applied Geochemistry**, Vol. 15, No. 1, 2000, pp. 79–89.

[19] M.O.P. Crespo and M.A.L. Yusty, "Comparison of supercritical fluid extraction and Soxhlet extraction for the determination of aliphatic hydrocarbons in seaweed samples", **Ecotoxicology and Environmental Safety**, Vol. 64, No. 3, 2006, pp. 400–405.

[20] A. Akinlua, N. Torto, T.R. Ajayi, "Supercritical fluid extraction of aliphatic hydrocarbons from Niger Delta sedimentary rock", **The Journal of Supercritical Fluids**, Vol. 45, No. 1, 2008, pp. 57–63.

[21] I. Okamoto, X. Li, T. Ohsumi, "Effect of supercritical CO2 as the organic solvent on cap rock sealing performance for underground storage", **Energy**, Vol. 30, No. 11–12, 2005, pp. 2344–2351.

[22] R.J. Hwang and J. Ortiz, "Mitigation of Asphaltics Deposition during CO_2 Flood by Enhancing CO_2 Solvency with Chemical Modifiers", **Organic Geochemistry**, Vol. 31, No. 12, 2000, pp. 1451–1462.

[23] A.H. Al-Marzouqi, A.Y. Zekri, B. Jobe, A. Dowaidar, "Supercritical Fluid Extraction for the Determination of Optimum Oil Recovery Conditions", **JPSE**, Vol. 55, 2007, pp. 37–47.

'An Innovative Approach to Interdisciplinary Teaching for Built Environment Students'

Sarah DICKINSON
Built Environment Department, Sheffield Hallam University.
Sheffield, South Yorkshire, S1 1WB, UK

and

Professor Paul WATSON
Built Environment Department, Sheffield Hallam University.
Sheffield, South Yorkshire, S1 1WB, UK

and

Ann FRANKS
Built Environment Department, Sheffield Hallam University.
Sheffield, South Yorkshire, S1 1WB, UK

and

Garry WORKMAN
Built Environment Department, Sheffield Hallam University.
Sheffield, South Yorkshire, S1 1WB, UK

ABSTRACT

Interdisciplinary teaching is essential for students who embark on a course within the Built Environment. This paper explores how the implementation of interdisciplinary teaching activities can enhance the learning experience of students, by preparing them for Industry and equipping them with the necessary skills for improving their integration and efficiency, when working within the construction industry.

The paper further explores the teaching methods used on Built Environment courses at Sheffield Hallam University, and the benefits they bring to student development. It also identifies how interdisciplinary teaching is embedded within courses. Furthermore, the paper outlines the collaboration to promote knowledge transfer between the different professional disciplines within the Built Environment.

The authors draw upon feedback from students, employers and professional bodies to demonstrate the success of student learning experience. Such modules facilitates the student to reflect on their experience and develop personally and professionally. This paper reflects on the experience of both staff and students to provide examples of best practice that can be disseminated to other Universities.

Keyword(s):
Built Environment, Collaboration, Interdisciplinary, Knowledge Transfer, Reflective Practitioner

1. INTRODUCTION

The construction industry has complex processes in which communication is a key activity in achieving a successful project. Therefore, particularly with reference to current economic downturn, employers are now seeking students, who are more 'work-ready' with confidence, have the ability to communicate effectively with all disciplines and can work as part of a coherent effective team. This is why embedding interdisciplinary teaching is a vital imperative for all Built Environment courses.

The success of engineering and construction projects depends very much on the strength of the multi disciplinary teams involved. The diverse ranges of skills and competences required for a construction project have to be moulded into a coherent holistic whole. The Latham (1994) report emphasised the need for the UK construction industry to work together as a team and further advocated a necessity for interdisciplinary working at the design stage of a project. Therefore it is essential that Universities address this critical issue by providing opportunities for interdisciplinary student work on Built Environment courses.

Teaching staff coordinate the modules, provide support for students and facilitate the learning process, but students have to manage their practice groups. As in industry they are required to create a professional working relationship with the group and manage their time and resources efficiently and effectively.

In the development of these modules key issues have to be addressed including; student experience relating to group work, translating 'real' projects into a manageable academic content and managing workloads for academic staff. And embedding Interdisciplinary learning with in the Built Environment Programme.

Sheffield Hallam University along with many other universities offer a range of courses in The Built Environment, such as Building Surveying, Quantity Surveying and Construction Management, however in reality all of these disciplines must be aware of what each discipline does and gain some experience of the challenges that are associated with them working together. The Built Environment Department at Sheffield Hallam University recognises the importance of collaboration between the disciplines and has introduced specific modules that focus on this aspect of interdisciplinary learning.

For many students the prospect of starting their new construction career; liaising and integrating with other disciplines can be daunting. It is therefore imperative that students understand and experience organisational behaviour within the industry that they are to work within, and it would be improvident for their studies to be entirely in one discipline. To develop communication skills and prepare students for this cross discipline working in the work place, students complete two modules, in which they work in interdisciplinary teams, comprising of Building Surveyors, Quantity Surveyors and Construction Managers. These modules are:

Interdisciplinary Project at Level 4 (year 1)

This module introduces students to the complex multidisciplinary nature of the construction sector they have selected to study. In a project driven environment they develop an understanding of their own professional role and that of others within a small multi disciplinary team representing four key professions; Architectural Technology, Building Surveying, Construction Management, and Quantity Surveying. Within that team, the students explore issues relating to Design, Building Appraisal, Construction Process and Measurement, drawing on individual team member's specialist knowledge to solve problems and draw conclusions.

The module provides a coherent and integrated introduction of the four selected professional roles in construction. Interdisciplinary aspects are reinforced through assessment that involves both single profession and multidisciplinary groups.

Integrated Project at Level 6 (year 4)

At final year level students have developed a comprehensive knowledge base and the professional confidence to work in small interdisciplinary practice teams, providing external consultancy services on a client focused project within a professional context. Students are able to apply the knowledge and skills developed in previous modules and draw on the previous work experience from their placement year, to develop solutions on a client driven project. This means that the module has to be dynamic and flexible and can change from year to year. It was designed to meet the requirements of the professional bodies and the industries need for interdisciplinary and problem based working.

2. RATIONALE FOR INTRODUCING INTERDISCIPLIANY PRACTICE IN THE BUILT ENVIRONMENT

Reflective Learning

The two modules are simulated and are problem based, they draw upon the students individual knowledge of their specific discipline by reflecting on their past studies. They have to be able to communicate effectively and understand the consequences of their decisions. The achievement of the modules learning outcomes concurs with 3 of the 5 adult learning principles of theorist Knowles, that:

- "Adults have accumulated experiences which can be a rich resource for learning.
- Adults become ready to learn when they experience a need to know something.
- Adults tend to be less subject centred than children; they are increasingly problem centred".

(Cited by Fry et al 1999, p.25) [1]

Experimental learning such as experience gained through the students' degree course and whilst on work placement plays a central role in the process of interdisciplinary learning. Kolb's (1984) popular theory of experimental learning is pertinent to interdisciplinary learning in that the ideas and knowledge from students' studies in relation to the problems that are encountered in practice can be formed and reformed through reflection of experience. Kolb's Learning Cycle can be adapted to interdisciplinary teaching, as depicted in Figure 1.0.

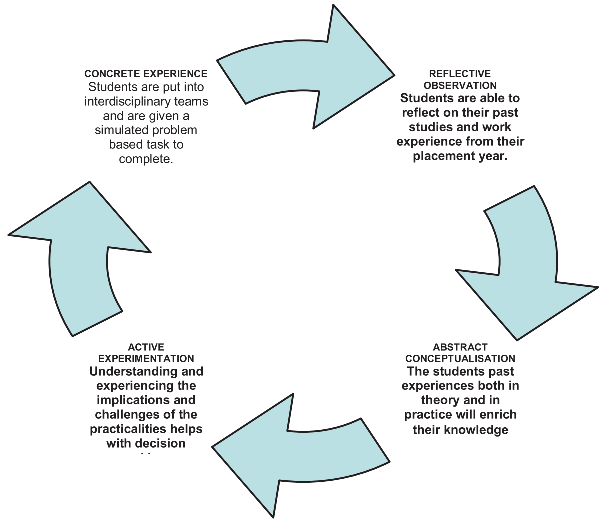


Figure 1.0: Interdisciplinary Learning Cycle adapted from Kolb (1984) 'Learning Cycle'.

Students are asked to provide a reflective report of their learning experiences subsequent to the completion of the module. It also provides valuable feedback to tutors who can then adapt and amend the module to develop the needs of the students both personally and professionally to prepare them for industry.

Constructive reflection is an important part of the interdisciplinary teaching. According to Moss and Hind (2005) [2] reflection can be structured or unstructured. Sheffield Hallam University's Built Environment interdisciplinary teaching at both levels incorporates both types of reflection, which in turn will develop the student's employability skills.

According to Moss and Hind (2005) [2], Unstructured Reflection is when students are able to think about past situations, such as in their work placement year and past study projects. Students are then able to cogitate on the success of their past experiences and recall what they could have done better. This type of reflection allows the student to identify aspects from past experiences that could be developed further.

The process of structured reflection also takes place, which in this case will be the students mark and feedback from completing the exercise, in which the students are able to improve, based on the tutors feedback and suggestions. This feedback can then be reflected upon and aids improved the student performance.

Implementation of Multidiscipline's

Built Environment students need to appreciate that professional multidisciplinary teams need to work together effectively and efficiently to produce successful projects. This is corroborated by Fry et al (1999, p.335) [1] who suggests that "Each team member must understand his or her role and its boundaries, and seek to understand the contribution of other team members... good communication, other skills for team working, and commitment to helping the client through productive collaboration must be developed during educational programmes... appropriate skills and attitudes could be developed within learning experiences confined to one profession, but multidisciplinary and interprofessional learning are often seen as key to enhancing collaborative practice."

Mullins (1990, p.374) [4] recognises that this type of interdisciplinary teaching is classed as informal groups in which "…are based more on personal relationships and agreement of group members than on defined role relationships [of a formal group]". Formation of these interdisciplinary groups will improve the success of projects through improvements of effective teams.

Problem Based Learning

Students tend to have less understanding when new information that is simply provided to them, or by asking them to read relevant material. Students must be allowed to play around with the information, so they can apply it to a situation. (Gibbs and Habeshaw 1989) [5]

It is recognised that by involving the learners in a simulated problem based situation similar to that encountered in the work place, students can reflect on their experiences and encounter problems, which is more effective than just discussing the possibilities of what could happen. By incorporating a problem based scenario students are able to recall and reflect on the information that they have learnt in past modules and their own experiences onsite, relate and discuss information to the interdisciplinary modules. This can be related to the 'Learning Model' of Hind and Moss (2005, p.8) shown in model 2.0.[6]

Model 2.0. The Learning Model. (Hind and Moss, 2005, p.8)

"Although the purpose of using problems in Problem Based Learning is to stimulate learning of information and concepts brought out by the problems (rather than to 'solve' the problems), Problem Based Learning does teach both a method of approaching and an attitude towards problem solving". (Schwartz et al 2001, p.2) [7]

Professional Competences

The list below demonstrates the Chartered Institute of Building (CIOB) professional competences [8] which are achieved on completion of the two interdisciplinary modules. Competences such as these are mapped to the requirements by Professional Bodies such as the Chartered Institute of Builders (CIOB) and the Royal Institution of Chartered Surveyors (RICS).

- Decision Making
- Communication
- Managing Information
- Planning and Organising Work
- Managing Quality
- Managing Health and Safety
- Implementing Sustainable Construction and Development
- Delivering Commercial and Corporate Objectives
- Personal Management at Work

Thus the modules cover the noted softer skills of management required by the professions and the Construction Industry.

3. CHALLENGES TO INTERDICIPLINARY TEACHING

There are many difficulties when implementing interdisciplinary teaching such as large student numbers, different degree routes part time/full time, programme duration; 3, 4 or 5 years. Timetable restraints, allowing all group members to meet at a given time, in a given place. Therefore with this in mind it is the tutor's main role to coordinate the activities and develop a supported student centred learning approach. Staff are selected to lead each specialist area and give students subject centred learning and support. In addition to this the module tutor provides support to the student groups and forms a link between the client and the students.

Part time students were concerned that the group activity would be difficult to undertake due to their other work and study commitments. The staff at Hallam University managed this by creating part time student groups and providing an electronic Blackboard area for group discussions and the sharing of information. The Blackboard site group areas were praised by students who found it supported the group learning experience.

It is important that the activity simulated a real project team and students where asked to create a project file and produce minutes of meetings. This not only enabled the tutors to monitor the professionalism of the group but to provide a clear record of decisions and targets. Students reported that this was a successful tool in creating and focusing the group, most students fully engaged with this activity.

The first year Interdisciplinary project is based on a simulated clients brief, but the final year project is a 'live' community project. The client group is made up of a number of people who are involved with the project including a local councillor, 2 local government officers, 3 offices workers, library staff and local residents. This added to the complexity of the assessment but most students responded well to the project. Feedback from students noted that while this was different from their other studies, which were more theory based, they enjoyed the challenge.

4. CONCLUSION

As with any complex modules this complex there are a number of challenges, both for staff in the management and coordination of the project, and for students undertaking the module. Each year students and staff reflect on the modules. They consider and evaluate aspects of the project that have worked well and those that need further development. Thus ensuring that the modules are continually improved and continue to be current and relevant to the professions and industry.

Interdisciplinary teaching provides the best possible scope for enhancing students understanding of multidisciplinary working within the construction industry. It increases the employability of students through exposure to the context of work. Providing students with the opportunity to reflect critically upon their own experiences in the work place and their knowledge from previous studies. By completing the modules students are able to critically reflect on their overall development of teamwork and problem solving skills. And identify ways of developing skills in relation to working with other disciplines.

Students often find group work challenging and the process must be carefully managed and controlled by staff to create a professional climate in which the individual can express their views. Feedback from Students on the final year module stated that they found the random allocation of students (in alphabetical order) 'fair and transparent' and while they initially found the prospect of working with students they didn't know worrying it added to the professionalism of the module. Some students thought that this provided a 'level playing field' in which they could excel.

Students felt that the module had successfully replicated the interdisciplinary nature of the construction industry. It was also reported by students that they understood the other professional roles more, as a result of undertaking the module. Many of the students said that they felt they had improved their communication skills and confidence. A product of this process is that student's social networks have crossed courses, which is good for the future of the industry.

Overall feedback on this module was extremely good, with students feeling it encapsulated their existing knowledge and allowed them to work in a professional context. Many reported they had grown as an individual during the 12 week project and felt that it had prepared them for work in the industry.

Staff have experienced the support of its External Examiners in the creation of this module. Those External Examiners that work in an academic environment understand the complexity and the problems that can arise in interdisciplinary group working. Those External Examiners who are in industry appreciate the practical nature module as preparation for the workplace.

5. REFERENCES

[1] Fry, H., Ketteridge, S., & Marshall, S (1999). **A Handbook for teaching and learning in Higher Education.** London: Kogan Page

[2] Hind, David and Moss, Stuart. (2005). **Employability Skills.** Sunderland: Business Education Publishers Limited.

[3] Fry, H., Ketteridge, S., & Marshall, S (1999). **A Handbook for teaching and learning in Higher Education.** London: Kogan Page

[4] Mullins, Laurie. (1990). **Management and Organisational Behaviour.** London, Pitman Publishing.

[5] Habeshaw, Graham and Gibbs, Trevor. (1989). **Preparing to teach.** Bristol: Technical and Educational Services Ltd.

[6] Hind, David and Moss, Stuart. (2005). **Employability Skills.** Sunderland: Business Education Publishers Limited.

[7] Schwartz, Peter, Mennin, Stewart and Webb, Graham. (2001). **Problem Based Learning.** London, Kogan Page Limited.

[8] Charted Institute of Building. **Record of Competence Achievements**. [Online]. Last accessed 22 February 2010 at: http://www.ciob.org.uk/membership/routes/pdp

Tri-Axial Inertial Magnetic Tracking in Quiet Standing Analysis using wavelet transform.

A. Martínez-Ramírez [1], P. Lecumberri [1], M. Gómez[1], M. Izquierdo[2]
[1]Mathematics Department, Public University of Navarra, [2] Studies, Research and Sports Medicine Center, Government of Navarra, Pamplona (Spain).

1. INTRODUCTION

Patients who are afflicted with frailty typically exhibit losses of muscle strength, fatigue easily and suffer increased risk (and fear) of falling [1-6]. Real time human motion tracking is an accurate, inexpensive and portable system to obtain kinematic and kinetic measurements and has particular applicability to the monitoring of disability in aging population [7-12]. Using a tri-axial inertial magnetic sensor, suitable for ambulatory measurements, we analyze the output signals of a quiet standing tests with in a healthy and in a frailty population. Time–frequency information based on wavelet decomposition was used for analysing all signals. The aim of this study was to examine the orientation signals and explore orientation patterns from a tri-axial inertial magnetic sensor suitable for ambulatory measurements, of a quiet standing tests with in a healthy and in a frailty population.

2. METHODOLOGY

The sensor we use consists of three tri-axial accelerometers and three gyroscopes integrated in a wearable device suitable for ambulatory measurements. The MTx (Xsens Technologies B.V. Enschede, Netherlands) is a small and accurate 3DOF inertial Orientation Tracker. It provides drift free 3D orientation as well as kinematic data: 3D acceleration, 3D rate of turn (rate gyro) and 3D earth magnetic field. So, the MTx is an excellent measurement unit for orientation measurement of human body segments [13].

We developed a custom software to manage the acquisition and processing the signals provided by the MTx sensor with the graphical programming language LabVIEW. This software offers the possibility to measure some variations of Romberg's tests as well as "up and go" test, gait test, one sit-to-stand movement test and a thirty seconds duration of repeated sit-to-stand movements test. It gives back 3D orientation, acceleration, earth magnetic field, angular velocities measurements, as well as 3D linear velocity, force and power real time calculations [13,14]. The software features real time data display, signal analysis and control with various pre- and postprocessing options including wavelet analysis.

In this research, we are concerned with evaluating four variations of Romberg's test. This simple test offers an important clue to the presence of pathology in the proprioceptive pathway and should be meticulously carried out during the neurological evaluation. The patient stands with feet together, and maintains balance with eyes open and eyes closed in 2 different positions; with feet together and semi-tandem (heel to instep).

Twenty five subjects from a frailty population and twenty four subjects from a healthy population volunteered to participate in this study.

Our signals are non-stationary signals. The wavelet transform is one of the most powerful tools for non-stationary signal processing.

It is well known from Fourier theory that a signal can be expressed as the sum of a possibly infinite series of sines and cosines. The big disadvantage of a Fourier expansion, however, is that it only has frequency resolution and no time resolution. To overcome this problem several solutions have been developed in recent decades that are more or less able to represent a signal in the time and frequency domain at the same time.

The wavelet transform (or wavelet analysis) is probably the most recent solution to overcome the shortcomings of the Fourier transform [15,16,17].

The continuous wavelet transform (CWT) is a transformation into a wavelet basis space. The idea behind this transformation is to project the signal onto different shifted and scaled versions of a so called 'mother wavelet', an oscillating signal that only exists over a finite period of time.

The CWT requires a great computational cost. This is the reason why the Discrete Wavelet Transform (DWT) is the most popular approach applied.
The DWT implementation can be performed by repeatedly filtering the signal with a pair of filters. Specifically, the DWT decomposes a signal into an approximation signal using a low pass filter and a detail signal using a high pass filter. The approximation signal is subsequently divided into new approximation and detail signals. This process is carried out iteratively producing a set of approximation signals at different detail levels (scales) and a final gross approximation of the signal.

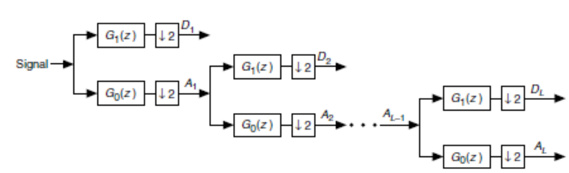


Fig 1: Discrete wavelet transform.

We tried different mother wavelets to calculate the wavelet indices and finally we chose the Coiflet 5 which experimentally yielded the best results.

Coiflets are mother variations on Daubechies wavelets with vanishing moments for wavelet function w(t) and scaling function £(t). This wavelet allows a very good approximation of polynomial function at different resolution [18].

Approximation and detail coefficients of sample Orientation signals taken from a frail subject and a healthy subject are provided in Figs. 2 and 3 respectively.

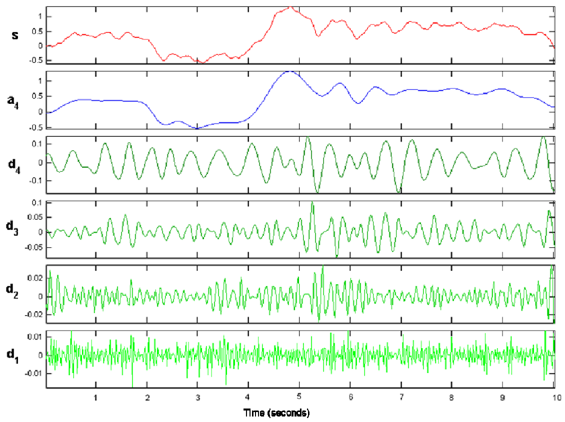


Fig 2: Wavelet decomposition of a sample Orientation signal from a frail subject during semi-tandem quiet standing balance tests with eyes closed.

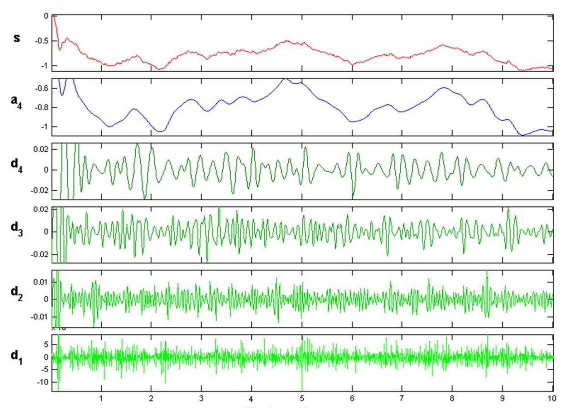


Fig 3: Wavelet decomposition of a sample Orientation signal from a healthy subject during semi-tandem quiet standing balance tests with eyes closed.

3. RESULTS

Orientation signals were recorded in the form of Euler angles and decomposed into an approximation signal at level 4, and details signals at levels 1, 2, 3 and 4.

The approximation gives us information about the shape of the signal, whereas each detail corresponds to a frequency interval or timescale, reflecting vibrations in different frequency range [16,19]. The details 4 correspond to low frequencies and the details 1 correspond to high frequencies.

The absolute sum of the coefficients of the details 1, 2, 3 and 4 were calculated using the DWT.

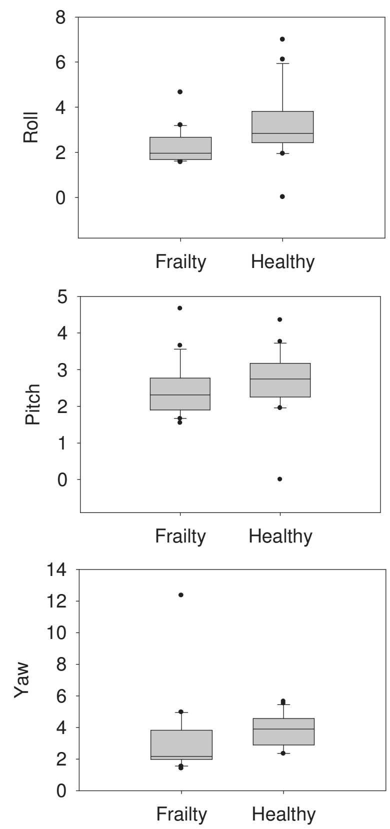


Fig 4: Boxplot of details signals at levels 1 of the 3D Orientation (roll, pitch, yaw) when the subjects performed semi-tandem quiet standing balance tests with eyes closed.

Figure 4 and Figure 5 corresponds to ten seconds duration semi-tandem quiet standing balance tests with eyes closed.

The box indicates the lower and upper quartiles with the central line showing the median. The top and bottom lines of the box represent, respectively, the medians for the upper and lower halves of the data and the ‘cat’s whiskers’ represent the highest and lowest values of the distribution, excluding outliers and extreme values. Outliers and extreme values are also presented.

Figure 4 shows the sum of the coefficients of the details 1. Healthy group showed significant greater values than frailty group in the orientation around each axe. Figure 5 represents the sum of the coefficients of the details 4. In this case, frailty group showed significant greater values than healthy group in the orientation around each axe.

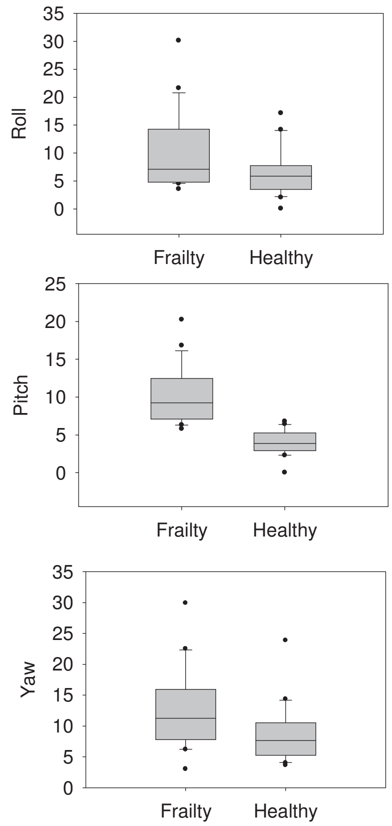


Fig 5: Boxplot of details signals at levels 4 of the 3D Orientation (roll, pitch, yaw) when the subjects performed semi-tandem quiet standing balance tests with eyes closed.

4. CONCLUSION

Accelerometry is an adequate method of monitoring subjects during their daily living activities because it makes possible to obtain objective and reliable measurements with no supervision at low cost. A wide range of measurements are possible, including balance control during quiet standing. Combining accelerometers and gyroscopes gives us information related to orientation and postural changes.

The system developed may be used for an occasional clinical assessment and, alternatively, is suitable for long-term ambulatory monitoring with no supervision, which makes it very useful for monitoring elderly people [20].

These results provide evidence for choosing orientation and wavelet details from a body fixed sensor (i.e including tri-axis accelerometer, magnetometer and inclinometer) to identify conditions associated with stability losses. Differences in stability can be identified with orientation sensors, using wavelet decomposition.

5. REFERENCES

[1] W. M. Bortz II, "A conceptual framework of frailty: A review," Journals of Gerontology - Series A Biological Sciences and Medical Sciences, vol. 57, pp. M283-M288, 2002.

[2] L. P. Fried, C. M. Tangen, J. Walston, A. B. Newman, C. Hirsch, J. Gottdiener, T. Seeman, R. Tracy, W. J. Kop, G. Burke and M. A. McBurnie, "Frailty in older adults: Evidence for a phenotype," Journals of Gerontology - Series A Biological Sciences and Medical Sciences, vol. 56, pp. M146-M156, 2001.

[3] J. E. Morley, "Mobility performance: A high-tech test for geriatricians," Journals of Gerontology - Series A Biological Sciences and Medical Sciences, vol. 58, pp. 712-714, 2003.

[4] K. Rockwood, D. B. Hogan and C. MacKnight, "Conceptualisation and measurement of frailty in elderly people," Drugs and Aging, vol. 17, pp. 295-302, 2000.

[5] D. M. Buchner and E. H. Wagner, "Preventing frail health," Clinics in Geriatric Medicine, vol. 8, pp. 1-17, 1992.

[6] L. Ferrucci, J. M. Guralnik, S. Studenski, L. P. Fried, G. B. Cutler Jr. and J. D. Walston, "Designing Randomized, Controlled Trials Aimed at Preventing or Delaying Functional Decline and Disability in Frail, Older Persons: A Consensus Report," Journal of the American Geriatrics Society, vol. 52, pp. 625-634, 2004.

[7] R. Moe-Nilssen, "A new method for evaluating motor control in gait under real-life environmental conditions. Part 1: The instrument," Clinical Biomechanics, vol. 13, pp. 320-327, 1998.

[8] R. Moe-Nilssen, "A new method for evaluating motor control in gait under real-life environmental conditions. Part 2: Gait analysis," Clinical Biomechanics, vol. 13, pp. 328-335, 1998.

[9] R. Zhu and Z. Zhou, "A real-time articulated human motion tracking using tri-axis inertial/magnetic sensors package," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 12, pp. 295-302, 2004.

[10] R. Takeda, S. Tadano, M. Todoh, M. Morikawa, M. Nakayasu and S. Yoshinari, "Gait analysis using gravitational acceleration measured by wearable sensors," *J. Biomech.,* vol. 42, pp. 223-233, 2009.

[11] R. Takeda, S. Tadano, A. Natorigawa, M. Todoh and S. Yoshinari, "Gait posture estimation by wearable acceleration and gyro sensor," in 2009, pp. 111-114.

[12] A. Weiss, T. Herman, M. Plotnik, M. Brozgol, I. Maidan, N. Giladi, T. Gurevich and J. M. Hausdorff, "Can an accelerometer enhance the utility of the Timed Up & Go Test when evaluating patients with Parkinson's disease?" *Med. Eng. Phys.,* vol. 32, pp. 119-125, 3, 2010.

[13] A. M. Sabatini, "Quaternion-based strap-down integration method for applications of inertial sensing to gait analysis," Medical and Biological Engineering and Computing, vol. 43, pp. 94-101, 2005.

[14] H. M. Schepers, D. Roetenberg and P. H. Veltink, "Ambulatory human motion tracking by fusion of inertial and magnetic sensing with adaptive actuation," *Medical and Biological Engineering and Computing,* vol. 48, pp. 27-37, 2010.

[15] Mallat, S. & Zhong, S., "Characterization of signals from multiscale edges", IEEE Transactions on patiern analysis and machine intelligence 14 (7), 710–732, 2006.

[16] Mallat, S.G., "Theory for multiresolution signal decomposition: the wavelet representation", IEEE Transactions

on Pattern Analysis and Machine Intelligence, vol. 11, no. 7, pp. 674-693, 1989.
[17] N. Bidargaddi, L. Klingbeil, A. Sarela, J. Boyle, V. Cheung, C. Yelland, M. Karunanithi and L. Gray, "Wavelet based approach for posture transition estimation using a waist worn accelerometer," in 2007, pp. 1884-1887.
[18] S. -. Huang and C. -. Hsieh, "Coiflet wavelet transform applied to inspect power system disturbance-generated signals," *IEEE Trans. Aerospace Electron. Syst.,* vol. 38, pp. 204-210, 2002.
[19] Rioul, O., Vetterli, M., "Wavelets and signal processing", EEE Signal Processing Magazine 8 (4), 14–38, 1991.
[20] B. Najafi, K. Aminian, F. Loew, Y. Blanc and P. A. Robert, "Measurement of stand-sit and sit-stand transitions using a miniature gyroscope and its application in fall risk evaluation in the elderly," IEEE Transactions on Biomedical Engineering, vol. 49, pp. 843-851, 2002.

6. ACKNOWLEDGMENT

Supported in part by a grant from the Ministry of Health, Spain (Network of Aging and Frailty RD06/0013/1003)

Complex Event Processing in Power Distribution Systems: A Case Study

Debnath Mukherjee Deepti Shakya Prateep Misra
Tata Consultancy Services Limited
Plot A2, M2 & N2, Sector V, Block GP
Salt Lake Electronic Complex, Kolkata, India 700091
{debnath.mukherjee, deepti.shakya, prateep.misra}@tcs.com

Abstract

Complex Event Processing (CEP) is an emerging discipline. This paper focuses on the application of CEP for fault detection and classification in 11kV radial distribution system using data collected from a Phasor Measurement Unit (PMU). The analysis has been done by monitoring the electrical quantities in the 11kV radial distribution system simulated using Matlab Simulink. The PMU is placed at the substation and transmits data to the Command and Control Center. The data is analyzed to identify the signatures of different types of faults and based on that rules have been designed to categorize them. An architecture stack has been designed based on a commercial CEP product (Tibco Business Events) to implement the fault detection. In this paper we present the architecture, the 11kV distribution system simulation and share our experience about fault detection. The paper outlines the categorization and analysis of various types of faults like Single Line to Ground (SLG) fault, Double Line to Ground (DLG) fault, and Three phase (3Φ) fault using CEP software. Thus it shows a real life application of CEP software for fault classification with low computational time.

Keywords

Complex Event Processing, Phasor Measurement Units, Architecture, Simulation, fault detection, fault classification

1 Introduction

Complex Event Processing (CEP) is an emerging technology which can be used to detect interesting patterns amongst events being received in real time. There are two related fields: Event Stream Processing is concerned with time-ordered sequence of events. CEP is of wider scope which includes partially ordered set of events known as “event clouds” [1].

In this paper we describe how CEP can be used for detection and classification of faults in 11kV distribution system using PMU. Fault detection and clearing is important for any network from safety and reliability perspective. In particular, we describe techniques that can be applied to distribution systems that monitor phasors using Phasor Measurement Units (PMU). The unique contribution of this work is efficient fault detection and categorization using CEP technology.

2 Related Work

Current fault detection technology is based on current transformer based relaying. The relay co-ordination ensures that the impacted zone is minimum. The shortcoming of this is relay setting and relaying accuracy may differ, sometimes not providing optimum fault clearance.

The authors are not aware of any reported fault categorization and detection work based on CEP technology. CEP platforms for phasor data concentrators and stream processing have been reported in [2]. However it does not mention specific use of CEP for fault detection and classification. Analysis of cause of faults and fault location is being considered in future work.

3 Phasor Measurement Units

PMUs are widely used in transmission networks for applications on Wide area monitoring and protection, Voltage instability analysis and prediction. Suitably engineered PMUs can also be integrated in to distribution level IEDs (Intelligent Electronic Devices). The paper examines such distribution level PMUs for fault detection and classification.

4 Problem Statement

The problem under consideration is detection of occurrence and type of faults in the 11kV radial distribution system. This distribution system is simulated using Matlab Simulink. Three phase, 13 bus balanced radial distribution system is simulated in which a PMU is placed at bus 1. The PMU is simulated using the three phase measurement block of Simulink library. Appendix A shows the system details. Table 4 mentioned in appendix A shows the line and load data for the test system. The system is simulated at frequency of 60Hz. In this simulation, different types of faults are created at different locations. At bus 1, three phase voltage and current waveforms are sampled with sampling frequency of 1kHz, in form of magnitude and phase angles. The PMU sends this data to Command and Control station where it is processed for fault detection and classification using CEP software. The format of the transmitted synchronized data follows the standard mentioned in IEEE Standard C37.118 [3]. PMU aggregates the signals i.e three phase voltage and current in terms of magnitude and phase angles. These data are then sent to a Command and Control station where the data is processed for fault detection and classification using CEP software. The format of the transmitted synchronized data is similar as mentioned in [3]. The captured data is then used to detect the type of fault occurred in the distribution system.

5 Using PMUs to detect faults

Let A, B, C and G represent phase a, b, c and ground. Different types of fault that can occur in the distribution system are:

Single Line to Ground Fault	(SLG)	i.e.	AG, BG and CG.
Line to Line fault	(L-L)	i.e.	AB, BC and CA.
Double Line to Ground Fault	(DLG)	i.e.	ABG, BCG and CAG.
Three phase Fault	(3L)	i.e.	ABC.

Whenever any one of these faults occur in the system, the voltage and current of a phase will show change in terms of decrease and increase in magnitude from its normal values respectively. If the fault occurs the phase angle for that particular voltage also show variation from its normal value. Similarly, phase angle for that phase's current can also show changes from its normal value. These changes can be used as the key parameters for fault classification.

As per [4], thresholds have been decided for these parameters for fault classification. For voltage,

if 1.1pu >|V|> 0.8pu then the system is in normal state

else if |V| < 0.8pu then there is a Fault in System.

else if |V| > 1.1pu then capacitor switching or transients occurs in system.

where |V| represents the phase voltage magnitude and pu stands for per unit.

So, when the system is running under normal condition, the voltage will be less than 1.1pu and will be greater than 0.8pu. When a fault occurs, the voltage magnitude will be less than 0.8pu. Voltage magnitude show changes for low fault resistances, but if there is a high resistance fault in the system then the voltage magnitude may remain in the normal band i.e. greater than 0.8pu; but the current for that particular phase will show noticeable change from its normal value. Similarly, the phase angle change for voltage is considered to be normal for a change of ±1 degree while for current, phase angle change of ±10 degrees is acceptable.

Two case studies for SLG fault on phase A, are discussed in this paper. When SLG fault on phase A occurs in the system, the magnitude and phase angle for phase A voltage and phase A current will show changes from their normal values. This change depends on the severity of the occurrence of the fault. The other two phases i.e. phase B and C will also suffer changes but the values will remain within a normal range.

Two cases for SLG fault i.e. AG are considered. These cases discuss the change in voltage as well as current waveform monitored by PMU placed at bus-1 with respect to change in fault resistance as well change in fault location.

Table.1. Magnitude and phase angle variation in phase A Voltage and Current with respect to fault resistance when SLG occurs at bus 25

Fault Resistance (ohm)	Voltage (V_A) (pu)	Phase voltage (Φ_{VA}) (degrees)	Current (I_A) (pu)	Phase current (Φ_{IA}) (degrees)
1	0.73	-8.73	5709.4	-59.98
10	0.86	-9.96	4100.5	-37.92
50	0.95	-5.54	2090	-25.14
100	0.95	-4.32	1654.9	-25.48
150	0.96	-3.87	1500.6	-26.28

The first case shows the variation in voltage and current parameter for phase A with respect to change in fault resistance from low to high value at one particular location i.e. at bus-25. While the second case shows changes in voltage and current parameter for phase A measured at bus 1 for change in fault location i.e at bus 21 and bus 31.

Case I: Study of the effect of variation in fault resistance at end of bus-25 of the system shown in Appendix A. This case considers that AG fault occurs at the end of bus 25. Faults with different fault resistances are created at this location and their effects are analyzed at bus-1 using PMU.

Table.1. shows the variation in three phase voltage and current data monitored by PMU at bus 1, with the change in fault resistances. It is observed that for low fault resistance the voltage magnitude is lower than 0.8p.u.but as the fault resistance is increased it has reached above this value. Thus for high resistance faults, the voltage magnitude remains within the normal range. Table 1 also show changes observed in phase angle (Φ_{VA}) for phase A for low as well as high resistance faults. The change observed is dependent on the instant of occurrence of fault as well as on the fault resistance. Thus, Table-1 shows that phase angle for phase A voltage changes even for high fault resistances but stays within normal limits as this parameter is more dependent on the instant of occurrence of fault rather than the severity of fault. Fig.1. show the variation in three phase voltage magnitude when high resistance faults occur at end of bus 25.

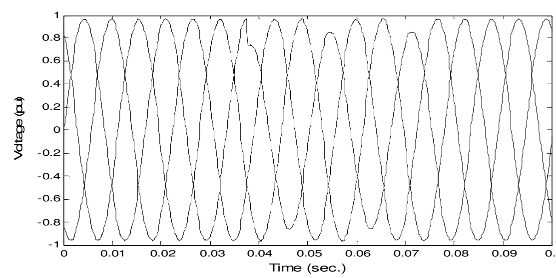


Fig.1. Three phase voltage waveform when AG fault occurs at instant 0.0376sec and is cleared at 0.0708sec with fault resistance of 50Ω at end of bus25.

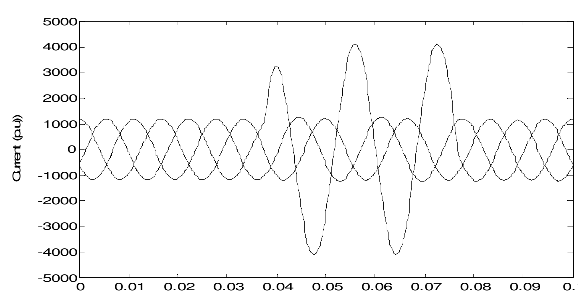


Fig.2. Three phase current waveform when AG fault occurs at instant 0.0376sec and is cleared at 0.0708sec with fault resistance of 50Ω at end of bus25.

Similarly, three phase current is also monitored at bus 1 using PMU. Table-1, show changes in magnitude as well as change in its phase angle for phase A current observed at bus 1, with respect to change in fault resistance. It is observed that magnitude of phase A current shows significant change even for high resistance faults. Similarly, its phase angle is also showing changes from its normal value for low as well as high resistance faults. In Fig.2. Phase A current is showing abrupt change from its normal value for AG fault, while the currents for other two phases remains within normal range.

Similar studies can be done for Double line to ground fault, Line to line fault and three phase fault.

Case II: Study of variation in fault location as well as fault resistance at end of bus-21 and bus-31.

AG fault is simulated at two different locations. First a fault was simulated at distance of 2.5km from monitoring end i.e. at the end of bus 21 then cleared. Then a second fault was simulated at 11km i.e. at the end of bus 31. Table.2 and Table 3. show the change in voltage and current parameters for phase A monitored at bus 1.

Table2. Magnitude and phase angle variation in phase A Voltage with respect to change in fault resistance as well as fault location when SLG occurs at bus 21&31

Fault Resistance (ohm)	Voltage (V_{A1}) (pu)	Phase voltage (Φ_{VA1}) (degrees)	Voltage (V_{A2}) (pu)	Phase voltage (Φ_{VA2}) (degrees)
1	0.51	-20.99	0.79	-6.83
10	0.86	-14.87	0.87	-8.23
50	0.95	-5.54	0.94	-5.38
100	0.95	-4.32	0.95	-4.28
150	0.96	-3.87	0.96	-3.85

Table.2. shows the comparison between the change in voltage magnitude and its phase angle at bus 21 and 31 with respect to increase in fault resistance for SLG fault at phase A. Let V_{A1} and V_{A2} represent the voltage magnitude monitored when the fault occurs at bus end 21 & 31 respectively. Similarly, Φ_{VA1} and Φ_{VA2} represents the phase angle for phase A voltage respectively. As it is AG fault, the other two phases may suffer change from its normal values in terms of magnitude and phase angle but within normal range. It is observed that for low fault resistances the voltage dip is more severe (as measured by PMU) when the fault occurs at bus end 21 in comparison to fault occurring at bus end 31. However if a high resistance fault occurs at these two locations, the voltage magnitude for phase A remains higher than 0.8p.u as measured by the PMU. Considering that each fault condition occurs at the same instant at both locations, change in phase angle for voltage A decreases with increase in fault resistance. It is also observed that for same fault resistance, change in phase angle (Φ_{VA}) decreases as the fault location increases from monitoring bus 1.

Table3. Magnitude and phase angle variation in phase A current with respect to change in fault resistance as well as fault location when SLG occurs at bus 21&31

Fault Resistance (ohm)	Current (I_{A1}) (pu)	Phase current (Φ_{IA1}) (degrees)	Current (I_{A2}) (pu)	Phase current (Φ_{IA2}) (degrees)
1	10646.6	-62.71	4510.6	-58.34
10	5334	-29.54	3613.8	-40.72
50	2175.7	-22.18	2057.6	-25.97
100	1682.1	-24.1	1648.7	-26.42
150	1515.5	-26.53	1498.5	-26.53

Table.3. shows the current magnitude and its phase angle variation for phase A only when AG fault occurs at two different locations. Let I_{A1} and I_{A2} represent the current magnitude monitored by PMU at bus 1 due to faults located at bus end 21 & 31 respectively. Similarly, Φ_{IA1} and Φ_{IA2} represents the change in phase angle for phase A current respectively. It shows that I_{A1} & I_{A2} show severe changes for low resistance faults at both locations. However, it is observed that for same resistance faults, I_{A1} will be of higher value than $I_{A2.}$ Thus, change in current magnitude is more prominent when fault occurs near to monitoring location and its magnitude gets reduced as the distance of occurrence of fault increases from the monitoring end. It was observed that at two locations i.e. at bus 21 & bus 31, the PMU placed at bus 1 recorded comparatively less change in magnitude for I_{A1} & I_{A2} when the high resistance fault occurred compared to low resistance fault.

Similarly, change in phase angle (Φ_{IA1} & Φ_{IA2} for current show variation with respect to change in fault resistance as well as fault location. The severity decreases with the increase in distance of fault location with respect to bus 1 as well with the variation in fault resistances.

Similarly, cases can be studied for double line to ground fault, Line to line fault and three phase fault.

For high resistance faults, three phase voltage magnitude may remain in normal limit. In that case three phase current magnitude can be observed for fault classification. Thus, if the current magnitude for a phase is higher than defined normal limit then it indicates a SLG fault for that particular phase. Similarly, if two phase or three phase current magnitude are higher than normal limit its DLG fault or three phase fault respectively.

The phase angle change for three phase voltage and current can be examined for additional information. These parameters can be observed to know the instant of occurrence of fault. If the fault occurs at zero crossing, the phase change will be near to zero on the other hand this change will be maximum if the fault occurs at the peak value. These values can also be used to distinguish among the DLG fault and L-L faults.

6 Solution

This section is organized as follows: Section 6.1 gives the solution overview and Section 6.2 onwards gives the details about the different modules of the solution.

6.1 Solution Overview

In this section we discuss the solution architecture for implementing fault detection using PMUs. The architecture diagram of the solution is shown in Figure 3. The PMUs are placed at the sub-stations from where its messages are communicated to the command and control station (CCS). The latter receives messages from multiple PMUs and processes them to determine if there is any fault.

The message sent from the PMU has the following fields (only those required for fault detection are shown):

1. Voltage magnitude and phase angle for each of the phases,
2. Current magnitude and phase angle for each of the phases.
3. Timestamp of measurement

From these message attributes, the pattern that a fault has occurred has to be detected. The pattern for a Single Line to Ground Fault (phase A to ground) is shown below

(IF Line to Neutral Voltage for Phase A < 0.8 pu

OR

IF Phase A current exceeds the current threshold)

AND

IF Other phase (B,C) voltage and current magnitudes are within normal limits

THEN

IT IS A SINGLE LINE TO GROUND FAULT IN PHASE A

The logic consists of two parts: Part 1 implements the logic that either the line A voltage is below normal threshold OR the line A current is above normal threshold. Part 2 implements the logic that other phase (B and C) current and voltage related parameters are within normal limits. The two parts are joined by an AND operator.

Similar rules hold for the double line to ground fault and three phase faults: for double line to ground faults, two of the phases will show abnormal characteristics and for three phase faults, all the three phases show abnormal characteristics.

The logic above can be mapped to "rule-based" systems. A *rule* consists of two parts: an IF part (known also LHS) and a THEN part (known as action). The IF part contains a *condition*, which when satisfied fires an *action* which is contained in the THEN part. Rules are evaluated in cycles: a number of input conditions may be satisfied in the first cycle which then fires actions which may change the value of some variables that are considered in the IF part of some rules. In the next cycle, some more rules may get triggered because some variables have been changed as mentioned before (in some cases same rules already fired before may be fired again if the variables in the condition part of the rule have different values than before as a result of executing some actions). In this way evaluation is done in cycles until no new rules are fired. This is known as "Run to Completion" in Tibco Business Events.

One class of CEP software, such as Tibco Business Events [5,6], includes rule-based inference engines which are enhanced to support event streams. Tibco Business Events receive event streams through the Tibco Rendezvous bus software which enables event producers to send events. Some CEP software such as Esper, Streambase and others use "continuous queries" to detect complex events. Continuous queries are SQL-like queries which are registered with the CEP server before the real-time data comes through, and act upon the real time data producing output with low latency. Some software such as Tibco Business Events offer both queries and rules.

The Tibco Business Events CEP Engine was used for this problem. Each type of fault was mapped to a Business Event rule: There were separate rules for line A to ground fault, line B to ground fault, line C to ground fault., Double line to ground faults and three phase fault.

The architecture diagram of the solution is shown in Figure 3.

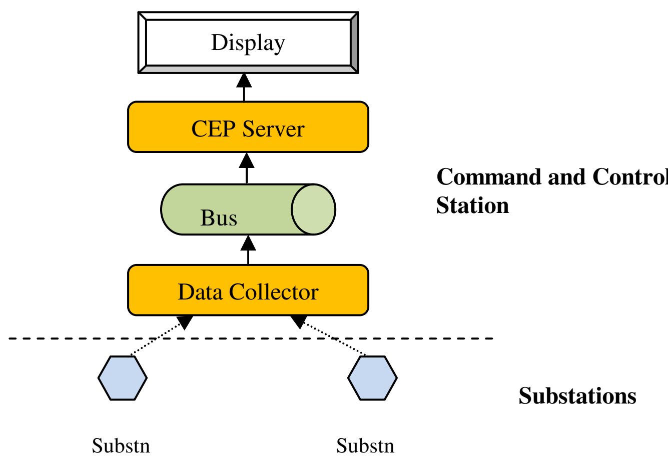


Fig.3: Architecture of fault detection

The architecture of the system consists of the following building blocks:

- At the substation level, a PMU sends a message to the command and control center
- At the command and control station, the message from the PMU are sent to a data collector which is an independent server program
- The data collector packages the PMU message as an event and sends to the CEP server (Tibco Business Events in this case) via the message bus (Tibco Rendezvous).
- The CEP Server executes the rules and sends to the display and also to pager/mobile alerts.

- It also generates events that are sent to external systems that can take automated action based on the event.

In the following sub-sections we explain the various components of the architecture.

6.2 Data Collector

The Data Collector is a Java program that adapts the input signal from the PMU to a format suitable for transmitting to the CEP Server via Rendezvous.

The Data Collector takes the input signal from the PMU and converts it to the event format. The event format is the same as that of the PMU message: consisting of the voltage, voltage phase angle, current and current phase angle for each of the phases and the timestamp.

The formatted data is sent to the Tibco Rendezvous bus in the Tibco-specific format. The Rendezvous bus is a daemon that runs on each machine participating in distributed computation. While in the prototype both Data Collector and CEP Server were hosted on the same computer, if the Data Collector and CEP Server are hosted on separate machines, the Rendezvous daemon need to be running on both the machines.

6.3 CEP Server

The CEP server hosts the rules for detecting faults. There are three rules for line to ground faults (for the three phases), three rules for double line to ground faults (for AB to Ground, BC to Ground and CA to Ground faults) and one rule for the Three Phase Fault case. Please see Section 6.1 for the structure of the rules.

One of the requirements of the CEP Server was that the display should not show the same type of fault if it occurs in consecutive cycles; and the display should indicate when a fault is cleared. We found that this required a *Concept* object to be used. We now explain Business Events *Concepts* and how it was used in our prototype.

A *Concept* in Business Events is similar to a class in object-orientation theory except that methods cannot be defined on it. A *Concept* can have attributes and it can inherit from other *Concepts*. *Concepts* can be instantiated just as Classes can be instantiated similar to objects in Object Oriented Design theory.

A *Concept* called PMU with two fields was defined: a Boolean field FAULT (which is true if a fault occurred in the last cycle) and a String field FAULTTYPE (indicates the type of fault observed in the last evaluation). An instance of this *Concept* was initialized for each PMU during application startup. Using this concept and its two fields (which hold the state of the PMU) the requirement that the display should not show the same type of fault if it occurs in consecutive cycles, was satisfied. Additionally if no fault occurred, a “no fault rule” was designed to reset the fields of the *Concept* instance and also indicate when a previous fault is cleared.

6.4 Display

The design can support both command line output and SCADA (Supervisory Control and Data Acquisition) human machine interface.

6.5 Prototype of Solution

A prototype of the solution has been built. A MATLAB simulation was used to generate data required for fault classification. An 11kV balanced distribution system with one radial feeder is simulated using Matlab Simulink. In the simulation, a PMU is placed at “bus 1” to monitor the three phase voltage and current waveforms. The PMU is simulated using Simulink library. Three types of fault were simulated (Line A to Ground, AB to Ground and Three Phase) with the change in fault resistance as well as fault location. These fault conditions cause change in phase voltage and current in terms of magnitude and phase angle. This change is captured by PMU at “bus 1”. The data is sampled at frequency of 1000Hz. Then average per cycle is computed for voltage as well as current waveform, such that one data sample per 60 Hz cycle is output. The files of data generated by MATLAB were read by the data collector and sent to Rendezvous from where Tibco Business Events picked up the events and generated the faults and fault outputs. The prototype has been designed keeping in mind that network communication will be used to transfer data between the PMU and the data collector.

7 Results and Experience

The experience of using MATLAB to simulate faults is explained in Section 5. The key results and experience of using CEP software are explained below:

How using rules simplified the development and enabled us to develop a prototype in a short time

The CEP software demonstrated two features that simplified development:

1) Definition of events and *Concepts* using a user-friendly IDE made it easy to develop the solution.
2) Presence of rule definition capabilities made developing the solution easy. Rules can be defined by non-technical business users also.

Some amount of state maintenance is required unlike traditional rule approaches

Simply using rule engines is not sufficient for this problem. Some amount of state has to be maintained about each PMU preferably in memory as explained earlier (see Concepts in Section 6.3).

Simplicity of messaging framework during connection setup phase

The CEP server was integrated with the Rendezvous messaging infrastructure. We found the Rendezvous messaging API simpler than standardized API like JMS (Java Message Service) during the setup phase, because as compared to obtaining a Connection, ConnectionFactory and Session, there was only one call to Tibrv.open().

The performance of Fault Detection using CEP Server

We carried out an analysis of the performance of the fault detection system. Fault data was fed in at a high rate to the CEP Server by the data collector. Each fault was detected in less than a millisecond. 3000 faults were detected in 1422 milliseconds, giving an average fault detection time of 0.474 millisecond per fault. The actual fault detection rule execution time is expected to be less than 0.474 ms because the latter includes the time for the client program to post to the Business Events engine via Rendezvous. The hardware used was Intel Core 2 Duo 2.33 GHz CPU with 2 GB RAM.

8 Future Work

In this work, only one PMU is placed in the distribution system. This work can be extended by placing more PMUs optimally in the distribution system for increasing the reliability of the solution. This fault classification algorithm can further be extended to fault location algorithm. Lines to line faults are not covered in this paper and will be considered in future work.

The current work is based on a set Voltage and Current level for fault determination. However, features can be incorporated to detect continuous Under voltage conditions and Overloaded conditions in the network and distinguished from high resistive faults. Radial network is assumed. Algorithm could be modified to adapt to in-feeds from local generation, alternate switching routes.

9 Acknowledgments

The authors would like to acknowledge the contributions of Sumit Kumar Ray, Narayanan Rajagopal, and Ranjeet Vaishnav of TCS for their valuable contributions in this work.

10 REFERENCES

[1] David Luckham, "The Power Of Events: An Introduction to Complex Event Processing In Distributed Enterprise Systems", Addison Wesley

[2] Downloaded from http://openpdc.codeplex.com/wikipage?version=63 on 24th February, 2010

[3] IEEE Standard for Synchrophasors for Power Systems, IEEE Power Engineering Society, pp 1-57, March,2006

[4] IEEE Recommended Practice for Monitoring Electric Power Quality, IEEE Standards Coordinating Committee Report, pp. 1-70, June,1995.

[5] TIBCO Software Inc., "Tibco Business Events User Guide"

[6] TIBCO Software Inc., "Tibco Business Events Language Reference"

11 APPENDIX-A

The data corresponding to the practical 13 Bus Distribution

System is tabulated in Table.4. Base kVA: 1000; Base kV: 11

Conductor Type: ACSR Line Resistance: 0.0086 p.u/km

Line Reactance: 0.0037 p.u/km

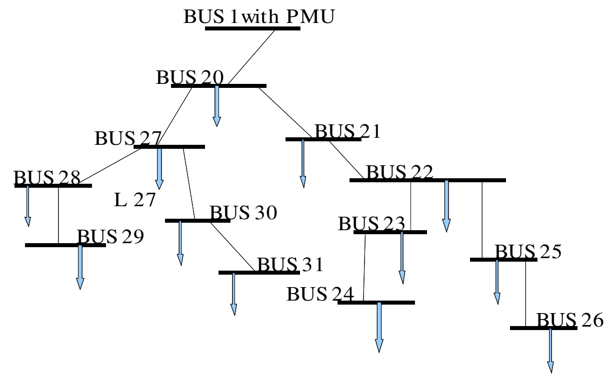


Fig.4. 13 bus 11kV balanced distribution system.

Table.4. Line and load data for the distribution system

End Buses of Lines		Line Length (kms)	Load at Bus y	
Bus x	Bus y		Real Power (kW)	Reactive Power (kvar)
1	20	1	108.0	52.0
20	21	1.5	94.0	46.0
21	22	3	81.0	39.0
22	23	5	108.0	52.0
23	24	2.5	108.0	52.0
22	25	3	102.0	50.0
25	26	4	41.0	20.0
20	27	1	108.0	52.0
27	28	1.5	162.0	79.0
28	29	2.5	68.0	33.0
27	30	4	68.0	33.0
30	31	5	95.0	46.0

A Conceptual Framework for Project Engineering Success

Rodger Oren, PhD
Bureau of TennCare
Nashville, TN, USA

ABSTRACT

This article discusses the framework of project engineering and project management success in the new millennia. The model evolves from a project-centric view toward a contemporary, holistic, integrated collection of factors in the context of the organization, the individual and the project. The conceptual framework provides the context for future research as it describes the evolution of project success constructs from the last century to the present. The work contributes to the field by looking at the organization through the lenses of the firm's culture, its politics, and its organizational makeup. In addition, the framework presents the variable of presencing, or the ability to anticipate the future, as an item necessary in the successful project of the 21st century.

Keywords: Project Management, Culture, Politics, Leadership, Triple Constraint, and Project Success.

1. INTRODUCTION

The practitioners in the field of project engineering/project management continue to be plagued with a majority of their projects encountering failures, even as the understanding and use of project engineering management techniques increases (Johnson, [6]; Kerzner, [8]). Projects are late, over budget, or not able to meet original scope requirements. The century may be new, but the project engineering/project management song remains the same.

The Wall Street Journal [2] reported on a natural gas pipeline project, from Colorado to Ohio, which cost $6.7B, a 50% overrun of costs, affecting the return on investment. In addition, the scope has not been met; the amount of natural gas to be delivered through the pipeline appears to be less than initially estimated, due to recent discoveries of natural gas in the Northeast US (Davis, [2]). With the current worldwide recession, it appears that project management failures are on the increase. About 68% of projects are encountering some form of issue creating an unsuccessful outcome (Levison, [9]). Success in projects in the new millennium appears to be as elusive as ever.

Organizations and their employees are encountering considerable levels of change, which can be characterized as chaotic, creating a situation where the firm must learn and adapt in order to survive (Lichtenthaler, [10]). Scharmer [16] stated the need for a person to anticipate the future as it emerges, as a technique to lead the change necessary in contemporary society. Challenges to the organization requires the culture to make sense of the situation, in order to provide some shelter for the participants in this environmental storm the firm encounters (Ravasi & Schultz, [15]).

The academic and practitioner communities have attempted to understand how success is achieved within the project context. Academics have published books and papers on many project engineering and management topics, tools and techniques. Practitioners have sought to learn from published materials and associations. However, the field continues to deal with more failures than successes. Are we in the project engineering and management community missing something?

2. WHAT ARE PROJECTS?

Projects are recognized as systems, with interacting components that require attention in order to succeed (Kerzner, [8]). This thinking is a paradigm shift, moving from an analytical mechanistic approach to a holistic mode, where interacting, interdependent variables are the name of the game (Gharadejaghi, [3]). As the system increases in complexity its interdependencies increase, which describes the situation of the project manager in today's environment. Projects are more complex, interdependent, with many variables that define success.

The environment is not just the project by itself, but an endeavor within the organization, stated by Carden and Egan, [1], “Project management needs to be viewed within the context of the organization” (p. 23). Work in other fields is being reviewed and evaluated for use within the profession, such as how leadership research may be applied to the project profession (Gehring, [4]; Turner and Muller, [18]). Projects are systems that are in the social arena, thereby defining themselves as a multi-minded social model (Gharajedaghi, [3]).

Leadership research from Prabhakar [14] sought to review qualitatively, the transformational behavior of project managers as an item of project success. Neuhauser [13] found that "there are no conclusive findings on effective leadership styles in either men or women in the project environment" (p. 22). While many in the field find leadership to be a factor of success, we find the debate of leader styles, methods and interactions, continues making the topic of leadership within the project management profession an item of continued study. As leadership factors are of debate, more work in leadership as a standalone subject, and leadership pertaining to project management will be required in order to find some agreement within the research community.

As project managers attempt to achieve favorable outcomes in their endeavors they are required to appreciate and consider external factors that will influence or moderate their results. Perceptive managers understand that the organization will contribute, either in a positive or negative manner, to the result that is sought. How the project managers, themselves, are skilled in handling their conduct will add to the mix of items influencing the success of the project. Finally, the basic factors of the project itself will contribute to the success or failure of the tasks itemized in the plan the project manager is executing.

3. THE CONCEPTUAL FRAMEWORK

Projects now have organizational impact, to a degree that was not present in the 20th century. Companies have engaged in re-engineering projects, mergers and acquisitions, and other organizational change activities with low probabilities of success (Mourier and Smith, [12]). Project management has changed from a single project being delivered within the organization, to multiple projects running concurrently, requiring the manager to understand organizational and team dynamics (Lientz and Rea, [11]).

Project management requires an understanding of systems theory, as a number of factors interact and influence the production of the good or service that is expected from the effort (Kerzner, [8]). Systems are defined as a collection of interacting parts that together are greater than the sum of their parts (Gharajedaghi, [3]). Project managers must be systems thinkers, as "the key to success then becomes managing the interaction between the different parts and not the parts themselves" (Gray and Larson, [5]). This statement is an exemplar of the triple constraint where one needs to understand how cost, scope and schedule interact in a manner where each is modified by a change in the other parameter.

However, the triple constraint is not enough in the new millennia. Business cycles are shorter, new threats emerge unforeseen, and businesses see challengers where none previously existed. Gharajedaghi [3] discussed how systems have evolved from older mechanistic concepts to interdependent variations, now to the multi-minded social models consisting of purposefulness and change in their outcomes. Projects are reflective of this archetype, a blending of project parameters, individual and organizational behaviors and actions.

The conceptual model, Figure 1 – Conceptual Framework of Project Success, describes the changes from the interactive mode of thinking to the purposeful, actor-driven representation. The framework depicts how increased complexity of projects, increased stresses on the organization and the need for transformational change drive organizational, individual and project factors into an interrelated mix that is necessary for project management success. This mix is a change from the quantitative project management approach to an integrated holistic approach, which Kerzner [7] defined as consisting of behavioral components necessary for success in contemporary settings. Sensing the organization's factors, the project manager uses their experience and leadership skills with the project factors to produce an outcome that is desired.

The framework visually defines the evolution of the framework from the 20th century to the 21st century. Where factors once stood alone in the 20th century, today the interaction between the organizational, project and individual factors occurs to a high degree. This interaction is documented as a grey overlapping area of the three factor boundaries. The intersection of the variables describes the mixing of the factors into a “project stew” where each parameter within the factor boundary can contribute to the project result.

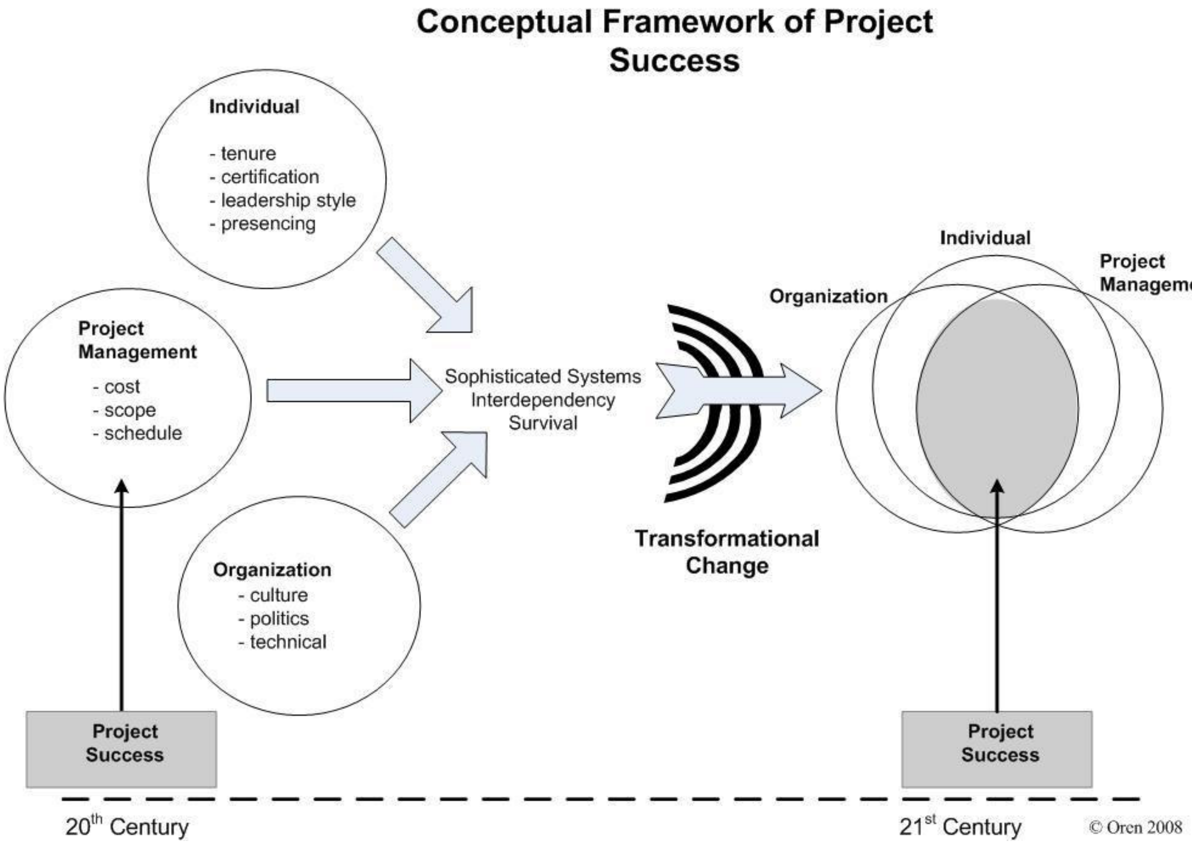


Figure 1 - Conceptual Framework of Project Success

4. CONCLUSION

While project success factors have been studied in previous situations, this work presents a different view of the project result process. The work seeks to investigate how integrative factors combine in a manner that produces a desirable outcome. Some project work has started to look into individual factors; however, no work has sought to integrate some of the presented individual variables with the presencing variable discussed by Scharmer [16]. In addition, no work to date considers an organizational view blending culture with political and structural views presented in the work of Tichy [17].

The framework presented will be researched in a variety of methods, using quantitative and qualitative approaches, in order to determine its value for the profession. Whether or not the model proves its merit as a picture of reality, the development of a new view, with integrated variables will assist the research community in other items to consider when looking at the world of the project manager. Research findings will help the practitioner community leverage their efforts into the variables and factors that produce the greatest benefit. Future work should consider the project as a holistic system, living in the organization composed of people, guided by the leadership of a project manager.

5. REFERENCES

[1] L. Carden and T. Egan, “Does our literature support sectors newer to project management? The search for quality publications relevant to nontraditional industries”, **Project Management Journal**, Vol. 39, No. 3, 2008, pp. 6-27.

[2] A. Davis, “Huge pipeline delivers billions to towns on its route”, **The Wall Street Journal**, Vol. 254, No. 80, 2009, p. A3.

[3] J. Gharajedaghi, **Systems Thinking: Managing Chaos and Complexity**, New York: Butterworth-Heinemann, 2006.

[4] D. R. Gehring, “Applying traits theory of leadership to project management”, **Project Management Journal**, Vol. 38, No. 1, 2007, pp. 44-54.

[5] C. F. Gray and E. W. Larson, **Project Management: The managerial process** (4th ed.), New York: McGraw-Hill Irwin, 2008.

[6] J. Johnson, **My Life is a Failure**, West Yarmouth, Massachusetts: The Standish Group International, 2006.

[7] H. Kerzner, **Advanced project management** (2nd ed.), Hoboken, New Jersey: John Wiley and Sons, 2004.

[8] H. Kerzner, **Project management, a systems approach to planning, scheduling, and controlling** (9th ed.), Hoboken, New Jersey: John-Wiley and Sons, 2006.

[9] M. Levison, “Recession causes rising IT project failure rates”, **Computerworld**. Retrieved August 29, 2009 from http://www.computerworld.com/s/article/9134547/Recession_Causes_Rising_IT_Project_Failure_Rates

[10] U. Lichtenthaler, “Absorptive capacity, environmental turbulence, and the complementarity of organizational learning processes”, **Academy of Management Journal**, Vol. 52, No. 4, 2009, pp 822-846.

[11] B. P. Lientz and K. P. Rea, **Project Management for the 21st Century** (3rd ed.), 2002, Retrieved October 7, 2009, from http://pmi.books24x7.com/toc.asp?bookid=5685

[12] P. Mourier and M. Smith, **Conquering organizational change: How to succeed where most companies fail**, 2001, Retrieved October, 7, 2009, from http://pmi.books24x7.com/viewer.asp?bookid=2485&chunkid=970039997

[13] C. Neuhauser, “Project management leadership behaviors and frequency of use by female project managers”, **Project Management Journal**, Vol. 38, No. 1, 2007, pp. 21-31.

[14] G. P. Prabhakar, “Switch Leadership in Projects an Empirical Study Reflecting the Importance of Transformational Leadership on Project Success Across Twenty-Eight Nations”, **Project Management Journal**, Vol. 36, No. 4, 2005, pp. 53-60.

[15] D. Ravasi and M. Schultz, “Responding to organizational identity threats: exploring the role of organizational culture”, **Academy of Management Journal**, Vol. 49, No. 3, 2006, pp. 433-458.

[16] C. O. Scharmer, “Presencing: Learning the future as it emerges: On the tacit dimension of leading revolutionary change”, **Conference on Knowledge and Innovation**, Helsinki, Finland, 2000, pp. 1-47.

[17] N. M. Tichy, **Managing Strategic Change: Technical, Political and Cultural Dynamics**, New York: John Wiley & Sons, 1983

[18] J. R. Turner and R. Muller, “The Project Manger's Leadership Style as a Success Factor on Projects: A Literature Review”, **Project Management Journal**, Vol. 36, No. 2, 2005, pp. 49-61.

Mathematical Relationship Between Particle Reynolds Number and Ripple Factor using Tapi River Data, India.

Dr.S.M.Yadav,
Associate Professor,CED,
SVNIT,Surat

and

Dr.B.K.Samtani,
Professor,CED,
SVNIT,Surat

and

Dr.K.A.Chauhan,
Associate Professor,CED,
SVNIT,Surat

ABSTRACT

The computation of bed load allows for the fact that only part of the shear stress is used for transport of sediments and some of the shear stress is wasted in overcoming the resistance due to bed forms therefore the total shear stress developed in the open channel requires correction in the form of correction factor called ripple factor. Different methods have been followed for correcting the actual shear stress in order to compute the sediment load. Correction factors are based on particular characteristics grain size of particle. In the present paper the ripple factor has been obtained for non uniform bed material considering the various variables like discharge, hydraulic mean depth, flow velocity, bed slope, average diameter of particle etc. by collecting the field data of Tapi river for 15 years for a particular gauging station. The ripple factor is obtained using Meyer Peter and Muller formula, Einstein Formula, Kalinske's formula, Du Boy's formula, Shield's formula, Bagnold's formula, average of six formulae and multiple regression analysis. The variation of ripple factor with particle Reynolds Number is studied. The ripple factor obtained by different approaches are further analyzed using Origin software and carrying out multiple regression on the 15 years of data with more than 10 parameters, ripple factor by multiple regression has been obtained. These values are further analysed and giving statistical mean to the parameters a relationship of power form has been developed. The ripple factor increases with the increase in the value of Particle Reynolds number. The large deviation is observed in case of Kalinske's approach when compare with other approaches

1. INTRODUCTION

Sarangkheda is one of the gauging stations on river Tapi. In the present paper the last 15 year data collected from this gauging station is used to compute ripple factor. The ripple factor is computed for monsoon season. There are number of approaches used to compute ripple factor but in this paper six approaches are used to compute ripple factor. The field data of 15 years has been analyzed and computer programming in Ms-excel and origin has been used to carry out analysis of the data. The relationship between Ripple factor and Particle Reynolds number has been established. The graphs are plotted for the above parameters using origin software and on the obtained results statistical analysis is carried out.

2. OBJECTIVES

The main objectives of this paper are:
(i) Using various measured parameters determine Ripple factor.
(ii) To develop mathematical model relating Particle Reynolds's number and Ripple factor.

3. STUDY AREA AND DATA COLLECTION

Tapi is the second largest westward flowing river of peninsular India. Total length of the river is 724 kms from origin to Arabian sea. The Tapi basin (fig.2)is situated between latitudes 200 N to 220 N, 80% of the basin lies in Maharashtra and the balance in the state of Madhya Pradesh and Gujarat. Central Water Commission, Tapi Division, Surat is regularly collecting daily data of discharge and sediment at gauging site Sarangkheda on river Tapi (fig.3). Sarangkheda is situated at a distance of about 488 kms from origin. The daily data during monsoon are
collected for 15 years period from 1981 to 1995 and 2000-2005. Bed load data (seasonal) from 1981 - 95 are collected for study from the central water commission data books (1980-1995, 2000- 2005)

4. DISCHARGE AND SEDIMENT OBSERVATIONS

Discharges are observed once in a day at 08:00 hours at all the sites and calculated by area-velocity methods. Cross-section is divided into 15 to 25 segments as per S1192:1981.Depths are measured by sounding rods as per IS 3912:1966.Necessary air and wet line corrections are done as per IS 1192:1981.Velocity is measured by cup-type current meter as per 3910:1966 suspended sediment samples are collected in Punjab Bottle Samplers at a depth of 0.6 D from the water surface.

Particle Reynolds Number

The particle Reynolds number is related to the fall velocity of the particle.

$$= \frac{wD_S}{\nu} \qquad (1)$$

Where, w is the fall velocity of the solid particle with diameter D_S. This dimensionless number is extensively used in the study of the dynamic properties of submerged solid particles and in resistance relations.

Ripple Factor

The computation of bed load allows for the fact that only part of the shear stress is used for transport of sediments and some of the shear stress is wasted in overcoming the resistance due to bed forms therefore the total shear stress developed in the open channel requires correction in the form of correction factor called Ripple factor.

Meyer – Peter and Muller formulae (1984)

In the Meyer – Peter and Muller formulae can be expressed as shown below in its basic form Basic form

$$X = 13.3 (Y^{-1} – 0.047)^{3/2} \qquad (2)$$

The Ripple factor as suggested by Meyer – Peter and Muller formulae is given below

Ripple factor $\mu = (C/C')^{3/2}$ C' based on D_{90}

The characteristics of above equation is discussed as below Characteristics grain size $D = D_i = \Sigma (p_i/D_i) \Sigma p$

Characteristics: The formula is mainly of an experimental nature. The experiments were carried out for D >= 0.4 mm for cases I which suspended load was absent. The formula has been extensively tested and used for rivers with coarse bed material. The table 1.1 presents the approach, concept and features of Meyer Peter's formulae.

Relationship Based on Estimated and Computed Parameters

The method used for estimation of ripple factor depends upon many variables. The seasonal average values of these parameters calculated by all the six methods are based on different approaches.

It is observed that all the six methods used for the study are well known, widely used and involves all-important hydraulic parameters and sediment parameters. Hence the comparisons of Ripple Factor calculated for developing the sediment transport characteristics of this river are compatible. Looking to the above fact average of six methods can be used as base without any loss of accuracy.Therefore average values of ripple factor, estimated by six methods is considered as a reliable base for the comparison of data collected and relation between ripple factor and various non dimensional parameters are established.

Multiple Regression Analysis

A multiple regression analysis is carried out between the measured and calculated basic data, viz., bed width discharge per unit width, flow area, hydraulic mean depth, velocity of flow, bed slope, avg. diameter of sediments with calculated average values of q_{bv}, , μ.

Finally equations are derived of the following form using Tables 4.1 to 4.12 for each river.

$$\mu_{\cdot multi} = C + C_1 \text{ x } q + C_2 A + C_3 \text{ x } S + C_4 \text{ x } B + C_5 \text{ x } V + C_6 \, HMD + C_7 \, Da. \qquad (3)$$

$\mu_{\cdot multi}$ = bed load discharge in terms of volume on the basis of submerged weight

q_{bw} = is bed load discharge for river under consideration on the basis of submerged weight.

q	=	discharge per meter width
A	=	Cross section area of flow
S	=	Bed slope
B	=	bed width of stream
V	=	Velocity of flow
HMD	=	Hydraulic Mean Depth
Da	=	Diameter of sediment

C, C1, C2, C3,C4,C5, C6, C7 are multiplying constants for – q, A,S,B,V,HMD, & Da.

The statistical analysis is carried out between Ripple factor, bed load discharge and various variables by using multiple regressions and the non linear square fitter of Micro cal Origin 7.5 has been used to obtain the base fit curve.

5. DATA ANALYSIS

In this paper 15 years field data of Sarangkheda gauging station of river Tapi River is analyzed.

Step: 1 The daily discharge data is converted in to monthly data.

Step: 2 The monthly data is then converted to seasonal data by taking average of monthly data so obtained is converted in to seasonal data. i.e. monsoon, post monsoon and pre monsoon seasons.

Step: 3 The seasonal data is converted in to yearly data.

Step: 4 The value of ripple factor obtained using six approaches.

Step: 5 After carrying out multiple regressions on this data results are obtained for three seasons.

Step: 6 Origin software is used to develop mathematical model to co relate Particle Reynolds number and Ripple Factor for each approach, average and multiple regression ripple factor

6. RESULT ANALYSIS

Almost for all the methods the variation of ripple factor calculated at particular station is almost uniform. An attempt is made under this study to develop simple equations which are best suited for the river under consideration the values used on the basis of average of six different methods can be used without mush loss of accuracy. The pattern of variation of hydraulic parameters or variables follows particular path. Following this path a multiple regression equation is developed for Ripple factor. The variation of hydraulic parameters is such that ripple factor shows very large variation during monsoon. Looking to such distribution, the best-fit curve based on least square method does not give the satisfactory results or very good value of co-efficient of determination on yearly basis. Therefore, an attempt is made to study seasonal variation of ripple factor. Ripple factor takes into account the effect of non-uniformity of flow. This non-uniformity can be correlated with bed conditions, flow conditions, dynamic conditions etc. and the relation between ripple factor and various non dimensional parameters can be established for given conditions of specific weight of sediment, diameter and sediment characteristics. However, it is very difficult to correlate all the parameters affecting the ripple factor. But depending upon local conditions the correlation can be made between ripple factor and the variables affecting the sediment transportation. Origin software is used to develop mathematical model to co relate Particle Reynolds number and Ripple Factor for each approach, average and multiple regression ripple factor. Table 1.2 represents the developed mathematical models. The statistical analysis of curve plotted between ripple factor and Particle Reynolds number is done by using non-linear-square-fitter to obtain the best fit curve.From the figure no.4 depicting Particle Reynolds number vs. Ripple factor, it is observed that the pattern of variation is same for all methods except Meyer Peter and Kalinske's. Comparison of all the six methods shows that Meyer Peter and Kalinske's equations deviate maximum. The pattern of variation for averaged and multiple regression Ripple factor curves appear similar. Comparison of all the six methods with corrected averaged Ripple factor curve and multiple regression curves shows that Ripple factor by Meyer Peter and Kalinske's Ripple factor gives more deviation. An extremely large deviation is observed in case of Kalinske's Ripple factor curve.

Sr.	Scientist	Approach	Concept	Features
1	Duboy's	Empirical	Excess Shear Stress	(1) Effect of bed forms is not considered
				(2) Bed Load moves in series of parallel layers and velocity of lowest layer =0
				(3) Assume linear variation of velocity
				(4) For critical condition entire bed moves as a single layer.
				(5) Developed for uniform material with different densities of sediments.
2	Shield's	Dimensionless Considerations	Excess Shear Stress	(1) Based on sediment site varying between 1.56mm to 2.47mm and sp. Gravity varying between 1.6 to 4.2
				(2) Variation of results up to 200%
3	Meyer- Peter Muller	Empirical	Excess Shear Stress	(1) Effect of Bed form considered
				(2) Total Shear Stress is partially utilize for over coming the firm resistance of undulation
				(3)Bed load transport is function of shear stress due to grains
				(4) Channel slope is divided into two parts (a) S' (slope required to overcome the resistance of grains) (b) S" (Slope required to overcome resistance of bed regularities)
				(5) Values of constants are different for uniform and non uniform material.
				(6) Used for rivers carrying coarse bed material
				(7) Results by this equation almost coincides with Einstein equation which is complicated
4	Einstein Brown	Semi Theoretical	Fall Velocity Criteria	(1) Contains no explicit correction for the shear stress, using fall velocity of particles but temperature effects are taken into kinematics viscosity.
5	Kalinske	Semi theoretical	Turbulence modern theory	(1) Bed Load transported is linked with characteristics of turbulent flow and uniform bed materials
				(2) Assume Gaussian distribution for flow velocity near bed
6	Bagnold	Semi Theoretical	Stream power concept	(1) Used concept of dispersion of solid particles under shear
				(2) Total Shear Stress= Shear stress at boundaries of shear stress due to collision of particles
				(3) Shear stress due to collision of particles depends upon normal force on particles and angle of internal friction.

7. DISCUSSION ON RESULTS

Following findings can be summarized from above study.

1. The extremely large deviation observed in case of Kalinske's approach as his formula consider uniform bed material which is not possible in case of river flow.
2. The pattern of variation of Particle Reynold's number with Ripple factor is same for all methods except Meyer Peter and Kalinske's.
3. The variation of Ripple factor with Particle Reynolds number for averaged and multiple regression Ripple factor curves appear similar.

8. REFERENCE

[1] Chow V.T., **"Open channel Hydrulics",** Mc Graw Hill Book Co, 1959.

[2] Einstein H.A. & F.M. Abdel-Aal, **"Einstein Bed – Load function at High sediment Rates",** JHD, Prose. ASCE, Vol. 98, No. Hy-1,Jan 1972.

[3] Garde R.J. & K.G. Ranga Raju, "**Resistance Relations for Alluvial Channel flow",** HJE, Proc. ASCE, Vol. 92, No. Hy-4 July 1966.

[4] Graf W.H. 1971, **"Hydraulics of sediment transport",** Chapter II, Bed form mechanics, Megraw – Hill Book Company.

[5] Simons D.B. & Richardson, D.V., 1971, **"Flow in alluvial channels, "River Mechanics"** edited by H.W. Shen, Vol.1, Fort colins, Colorado.

[6] Meyer-Peter, E & R Muller, **"Formulas for Bed load Transport",** Proc. IAHR, 2nd Congress, Stockholm, 1948.

[7] Vanoni V.A. & G.N.Nomicos, **"Resistance properties of Sedimen – ladenstreams"**, Trans.ASCE, Vol. 125, 1960

Table.1.2 Mathematical models for Particle Reynolds Number
Tapi-Sarangkheda-Monsoon Season

Sr.No	NAME OF SCIENTIST	MODEL	EQUATION	a	b	c	d	e	Chi^2	R^2
1	E.BROWN	Asymptotic1	y = a-b*c^x	0.933	0.017	0.99529	-	-	4.3188E-9	0.99898
2	DUBOY'S	Allometric1	y = a*x^b	0.430	0.1705				0.89291	0.94374
3	SHIELD'S	Cubic	y = a + b*x + c*x^2 + d*x^3	3.057	-0.018	0.00004	-2.19E-08	-	0.04348	0.97888
4	MEYER	Poly4	y = a + b*x + c*x^2 + d*x^3 + e*x^4	-10.89	0.108	-0.00035	4.81E-07	-2.36E-10	0.08997	0.86823
5	BAGNOLD'S	Cubic	y = a + b*x + c*x^2 + d*x^3	0.246	-0.0037	7.45E-07	-4.77E-10	-	5.5092E-10	0.94379
6	KALINSKE'S	Allometric1	y = a*x^b	1.0258	0.19466			-	0.00362	0.81133
7	AVERAGE	Poly4	y = a + b*x + c*x^2 + d*x^3 + e*x^4	4.299	-0.03	0.0001	-1.41E-07	7.08E-11	0.00741	0.81235
8	MU MULTIPLE	Poly4	y = a + b*x + c*x^2 + d*x^3 + e*x^4	3.536	-0.022	0.00008	-1.07E-07	5.36E-11	0.00836	0.81746

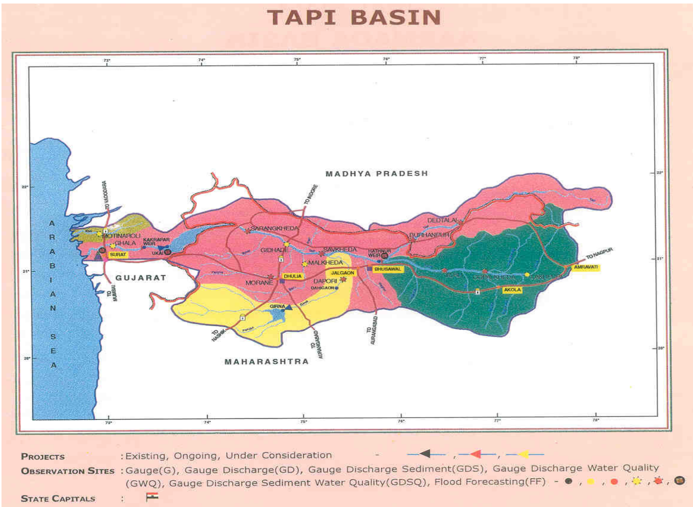


Fig.2 Tapi Basin.

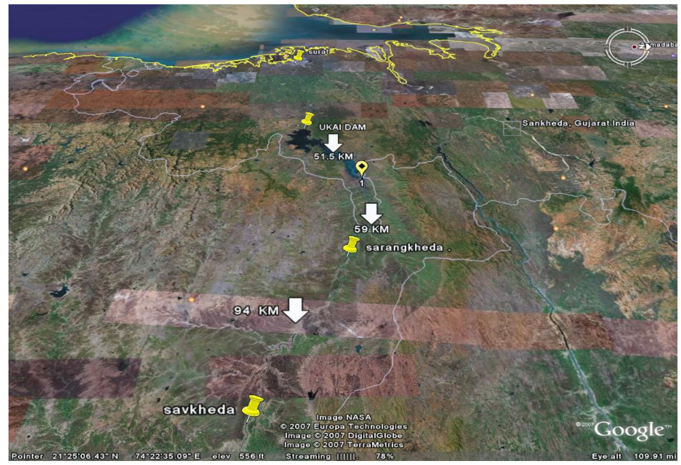


Fig.3 Savkheda, Sarangkheda gauging stations and Ukai dam.

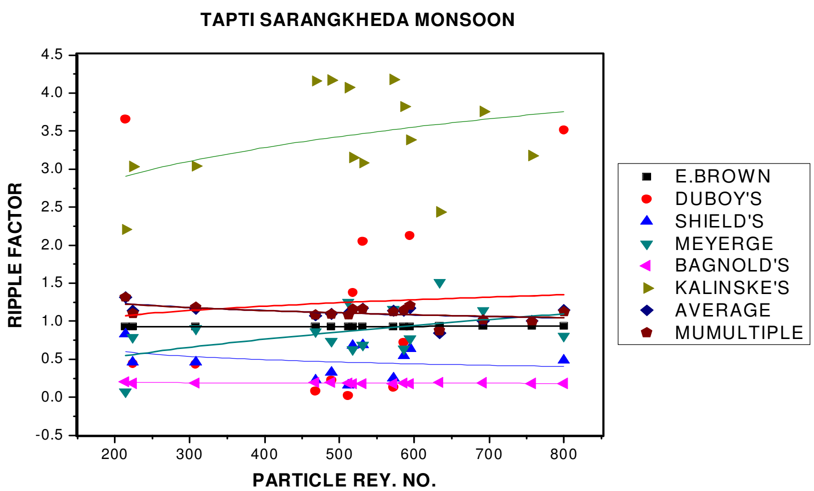


Figure.4 Particle Reynolds Number versus Ripple Factor

Non-invasive method for pre-hospitalization treatment of heart attack patients

Syed Ammar Zaidi(1), Marcin Marzencki(2), Carlo Menon(1) and Bozena Kaminska(2)

(1) MENRVA group, Simon Fraser University, Burnaby, BC
Tel.: +1-778-782-6860, Fax: +1-778-782-4951, e-mail: cmenon@sfu.ca
(2) CIBER lab, Simon Fraser University, Burnaby, BC
Tel.: +1-778-782-7129, Fax: +1-778-782-4951, e-mail: mjm11@sfu.ca

***Abstract*—We propose a novel method of heart attack treatment by low frequency diastolic timed vibrations. It can be applied quickly after the onset of symptoms by unspecialized personnel drastically improving the survival chances of the patient. The method is based on applying low frequency mechanical vibrations in synchronization with the heart cycle of the patient to facilitate disruption and clearance of acute coronary thrombosis. We present an analysis of the proposed methodology and provide experimental results obtained with a first prototype of diastolic timed vibrator. We show that vibrations of required frequency can be successfully synchronized with a real time ECG signal.**

I. Introduction

Diseases of the heart are the leading cause of death in the United States, with higher mortality rate than cancer (malignant neoplasms) [1]. Over 7 million men and 6 million women are living with some form of coronary heart disease. Over a million people suffer a (new or recurrent) coronary attack every year, and about 40% of them die as a result of the attack [2]. This means that roughly every 65 seconds, an American dies of a coronary event.

Myocardial Infarction (MI) or heart attack is most often caused by a blood clot, also known as thrombus, in the arterial vasculature surrounding the heart. MI refers to myocardial cell death and occurs due to a complete coronary obstruction which results in a profound blood flow impairment causing inadequate oxygen delivery to the heart muscle. Once such an obstruction begins, cell death can occur in as little as 20 minutes. Complete death of all myocardial cells at risk can occur in, at the earliest, 2 to 4 hours [3]. Various methods have been developed to treat thrombus before MI occurs. The techniques vary from surgical procedures, such as coronary artery bypass grafting to minimally invasive procedures such as angioplasty, atherectomy, thrombectomy, and intra-arterial thrombolysis [4]. Procedures such as angioplasty involve pressing the thrombus against the walls of the vessel using a balloon catheter or drawing it out of the vessel. Alternative invasive procedures such as intra-arterial thrombolysis involve direct insertion of thrombolytic agents i.e. tissue plasminogen activator (TPA) into the artery through a process known as catheterization. These agents are capable of dissolving the thrombus and are inserted after the location of the thrombus is determined by another catheterization process known as coronary angiography [5]. Other methods include exposing the blood clots to low frequencies of continuous wave ultrasound. In these methods, known as ultrasound induced clot dissolution, the treatment is dependent on the intensity and duration of the ultrasound [6]. Preferred invasive methods such as angioplasty require significant setup time and resources in order to be successfully used for treatment. Before the treatment method itself can begin, patients are required to undergo a number of clinical diagnosis tests which include, but are not limited to, electrocardiogram, blood tests, coronary catheterization and the like [3]. Incidentally, the most effective treatment occurs during the first 60 minutes of the symptoms known as the golden hour. However, by the time an average patient reaches a hospital (approximately 2.7 hours after the onset of symptoms) most deaths have already occurred [5]. This is worsened by the fact that those who manage to survive this deadly period of illness have to spend additional time in the hospital undergoing examinations or receiving transport to a cardiac cathlab before the treatment can begin. As a result, speed of intervention is the biggest factor in saving a patient's life and is the key to effective heart attack treatment. It has been suggested [7] that instead of being transported to regional revascularisation centers, heart attack patients should receive immediate care at their nearby hospitals or other facilities. In addition, if treatment could begin during transportation to the hospitals, it would play a key role in ensuring the survival of the patient.

In this paper, we propose a method that could be safely applied by unspecialized personnel on-site or during patient transportation to the hospital. We believe that this method could drastically improve the survival rate of heart attack patients. It consists in applying low frequency mechanical vibrations synchronized with heart cycle of the patient, preferably along with injection of thrombus dissolving drugs.

Our paper is organized as follows. After the introduction, we present our method in details, including review of state of the art and analysis of the underlying idea. Subsequently, we present the architecture of the proposed system and we discuss various aspects of its implementation. In the following part, we describe a prototype system and the experimental results obtained. We close this paper by providing conclusion and proposition of future works.

II. Proposed Method

We present a novel, non-invasive method suitable for treatment of heart attack and other states of low coronary blood flow in a human. It is based on applying low level vibrations in the chest area along with application of clot dissolving drugs. By performing vibrations during the diastolic period of the cardiac cycle (the relaxation of the heart) it is expected that coronary flow is increased and thrombus dissolution is achieved. Instead of continuous vibrations, diastolic timed vibrations must be performed in order to ensure the heart is not interrupted during the systolic period of cardiac cycle (contraction of the heart), which can have very negative effects, especially on a weakened heart. In this study, we first aim at developing a vibrating system that is independently controlled by a real-time ECG signal. The triggering of the vibrating system should be synchronized with the ECG signal in such a way so as to remain vibrating in the diastole and cease all vibrations in the systole of the ECG signal.

Our goal is to create a device for field use - a Diastolic Timed Vibrator (DTV) to be used as medical emergency system to remediate acute states of low coronary blood flow, such as those exhibited in angina pectoris (chest discomfort secondary to coronary artery narrowing) or heart attack (an acute blockage of a coronary artery, usually by a blood clot). The DTV will impose mechanical vibrations to the chest of the patient in order to improve coronary blood flow. We aim at creating an inexpensive and portable system requiring minimal intervention of specialized personnel.

A. Diastolic mechanical vibrations

There is strong experimental evidence that diastolic mechanical vibrations on the chest wall increase human coronary blood flow (CBF). In past studies, diastolic vibrations performed on patients with coronary arterial disease (CAD) and on normal subjects resulted in an immediate increase of CBF as measured by both transesophageal doppler and coronary flow wire. The CBF increase in CAD patients was significantly larger than those of normal subjects [8]. In addition, clinical studies performed on humans and canines have shown that external diastolic vibrations can release incomplete relaxation (IR) and improve the systolic function of the heart [9], [10]. Similar studies consisting of external vibrations performed on human patients with aortic regurgitation (AR) and ischemic heart disease (IHD) resulted in a decrease of left ventricle systole pressure; proving that vibration induced depression does occur in humans [11].

Clinical studies have shown that diastolic timed mechanical vibrations around 50 Hz improve coronary blood flow and left ventricular (heart muscle) performance in human volunteers, with and without coronary artery disease [8], [9]. Low frequency vibration is a known potent vasodilator, especially for arteries with a degree of active tension or spasm [12], which is often the case in heart attack [13], and it has further been shown to significantly enhance clot dissolution with or without a thrombolytic agent both in-vitro and in commercially available catheter systems [4]. Low frequency external tapping has also been documented to lead to reliable and immediate clearance of acute coronary thrombosis in animal models presumably by enhancing dis-adherence of clot from a narrowed intraluminal surface [14]. We suggest that the efficacy of disrupting and clearing thrombosis could be maximized by providing vibrations at different frequencies (via frequency sweep or random frequency variation) as this would facilitate rupture of the different chemical bondings of the clot and add turbulence in the vascular system. It would improve mixing of clot dissolving agent and enhance erosion of clot surface. Vibrations in the 40-60Hz range fall within the resonance frequency spectrum of the heart muscle [15] which would thereby ensure a maximized therapeutic effect.

B. ECG synchronization

Our method provides a new technique for disrupting and clearing the thrombus present in a patient's arterial vasculature surrounding the heart. During systole the heart is contracting and pressure needed for driving the blood is being generated within the chambers of the heart. As a result, vibration, which can interfere with the heart's contractile process, cannot be performed on the heart during this phase and should only be done when the heart is in relaxation phase - diastole [11]. Furthermore, it has been demonstrated in clinical studies that vibrations timed exclusively to the diastole of the cardiac cycle advantageously facilitate heart muscle relaxation and paradoxically improve the strength of the heart contractions and hence can be utilized safely [10], [16]. We propose to develop a device that would apply mechanical vibrations to the chest to augment coronary perfusion, disrupt blood clots and generally improve blood circulation. The therapy could be performed by a paramedic in an ambulance or by a trained person in a clinic or an emergency room.

C. System architecture

The proposed system is composed of four main parts: a vibrator, accelerometer, ECG system and a LabView VI containing signal processing and control. Fig. 1 presents a schematic of the system architecture.

1) Vibrator: We use a standard electromagnetic motor driven by a 50 Hz source to generate rotary movement translated into linear movement of a plate. Furthermore, a variable damping stage is added to adjust amplitude of generated vibrations. In order to be able to generate vibrations only in the desired periods of the heart cycle, a fast electromagnetic relay[1] is introduced on the power line of the motor. The relay is driven from a DAQ[2] connected to the LabView interface.

2) Accelerometer: A MEMS accelerometer[3] has been integrated into the vibrating plate to provide a feedback on the generated vibration amplitude. All three axes can be monitored for added reliability. Signal from the accelerometer is digitalized using the DAQ and sent to the LabView interface for further processing.

[1]Panasonic APE30106

[2]A NI9205 with analog inputs and a USB6008 with analog outputs are used in our system.

[3]STMicroelectronics LIS3L02AL

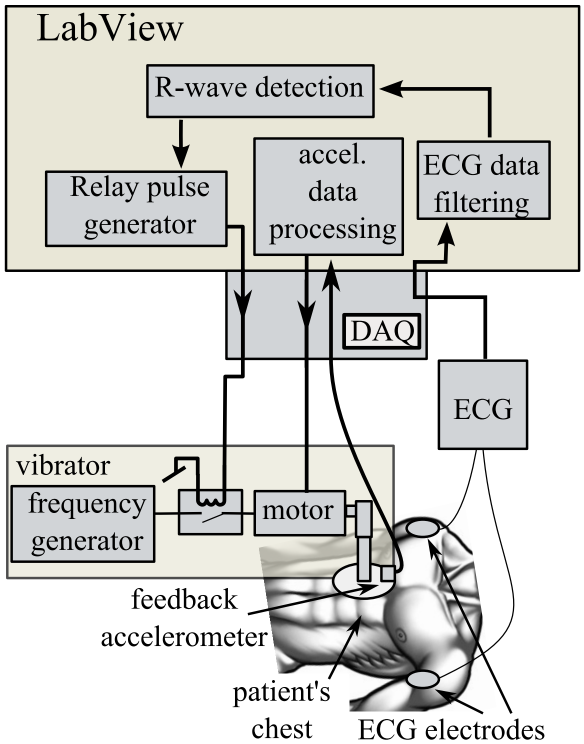


Fig. 1. Block diagram of the proposed diastolic vibration system.

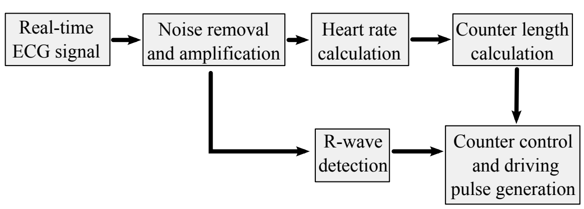


Fig. 2. Block diagram of the ECG signal processing used to generate driving signal for starting mechanical vibrations in synchronization with diastolic period of the cardiac cycle.

3) ECG: An ECG acquisition system Burdick EK10 is employed to amplify and filter human ECG signal. The resulting signal is digitalized by the DAQ and further processed in LabView.

4) LabView VI: An algorithm has been developed to detect the systole and diastole in a real time ECG signal in order to allow diastolic timed vibration. We created a virtual instrument using National Instruments LabView. As shown in Fig. 2, a real-time ECG signal is first filtered to eliminate noise. We use a low pass filer to eliminate unwanted high frequency components. Subsequently, it is high-pass filtered to detect the QRS complex. The resulting data is then used for heart rate calculation and detection of the R peak in the QRS complex.

The QRS complex is used to calculate the heart beat rate by finding the ratio between the number of QRS complexes over the elapsed time for a specific interval. After the heart rate is determined, and therefore the period of the ECG signal, lengths of two counters are calculated for running the vibrating system. These counters, in coordination with the R peak detection, are time controlled to stop the vibrating system during systole and enable it during diastole. After detecting the R peak, systole counter is reset disabling the vibrating system until the systole cycle is complete. Once the systole counter reaches its limit (which is set to match with the completion of the systole cycle), the vibrating system is enabled again for a duration determined by the diastole counter. The diastole counter is accordingly set to reach its limit before the beginning of the systole cycle. In case where the two counters overlap due to incorrect counter length calculation, the systole counter has priority over the diastole counter; thus ensuring that any detection of R peak would disable the vibrating system. The systole counter duration was approximated based on the QT interval calculations performed during past clinical studies of heart disease patients. Based on the data collected in these studies, a 30 BPM (2 second period) heart rate would have an approximate QT interval of 0.5 seconds. As a result the systole counter was set to be 1/4 length of the period [17]. This approximate calculation was used during the early phase of the testing and will be replaced with a more efficient regression based algorithm.

The resulting system ensures that vibrations stop before the QRS complex begins. The counters are updated in real time to adapt to varying heart rate value.

III. Experimental Results

In order to verify the accuracy of our predictions concerning the effectiveness of diastole timed vibrations, we started with building a model system. We concentrated on proper synchronization of mechanical vibrations with ECG signal.

In order to determine if an external vibrating system can be controlled by an ECG signal, we used a 5V DC gear motor as the vibrator and the ECG signal was generated by a Multiparameter Patient Simulator Fluke PS420. The 50 Hz vibration was simply generated by attaching an eccentric weight to the shaft of the DC motor and adjusting the driving voltage. The ECG signal was kept at a heart rate of 30 BPM.

Fig. 3 shows the generated mechanical vibrations synchronized with a real-time ECG signal. The overall period of the heart beat is 2 seconds in which the PQRST region of the ECG lasts for $0.55 \pm 0.05s$. Although the systole (QRST region) only lasts for 0.41 ± 0.04 seconds, the vibrating system is turned off before the PQRST region begins in order to ensure that vibrations occur only when the heart is in its relaxation state. During the diastolic cycle, the DC motor is allowed to vibrate for 1.40 ± 0.04 seconds. There was some overlap of vibrations onto the outer edges of the P and T waves of the ECG signal; however it was ensured that vibrations never occurred during the critical QT interval.

After having successfully synchronized a DC motor with a real-time ECG signal, a commercially available massager device Human Touch HT-1280 (which has frequency and stroke amplitude suitable for clinical use) was tested as the vibrator. Fig. 4 shows the modified massager with mounter MEMS accelerometer on the vibrating plate. In this setup, a relay was placed between the power supply and the massager,

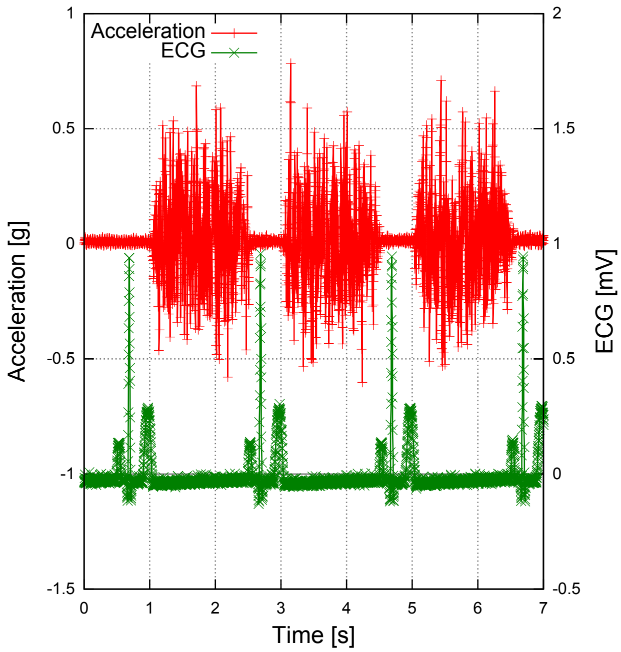


Fig. 3. Experimental results for mechanical vibrations (represented by acceleration amplitude) synchronized with diastolic period of the cardiac cycle (represented by an ECG signal at 30 BPM).

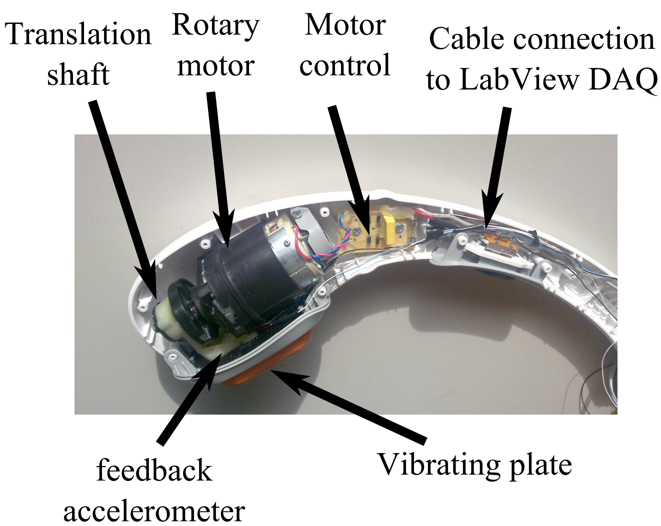


Fig. 4. Human Touch HT-1280 massager with modified motor control and a feedback MEMS accelerometer mounted on the vibrating plate.

allowing the triggering of the massager through the Labview VI. This device has much higher rotary inertia, therefore a special damping system had to be introduced in order to allow more precise timing of the generated vibrations. Fig. 5 shows the experimental results obtained with the vibrating system synchronized with the ECG signal. The driving signal is used to indicate when the vibrations are triggered based on the counter algorithm described earlier in the system architecture section.

It can be seen that although there is a delay between deactivating the driving signal and the actual stopping of the vibrations, all vibrations stop at the beginning of the QRS complex. The delay was measured to be approximately $17ms$ between the end of the driving pulse and the actual termination of the mechanical vibrations. The delay caused the vibrations to overlap onto the beginning of the PR interval and the last portion of the T wave. For the most part, the QT interval was free of vibrations and with a more precise damping system and ECG processing algorithm the delay will also be eliminated. The approximated counter based algorithm will be replaced with a more concise regression based algorithm; thus ensuring precise calculations of the QT interval and precise triggering of the vibrations.

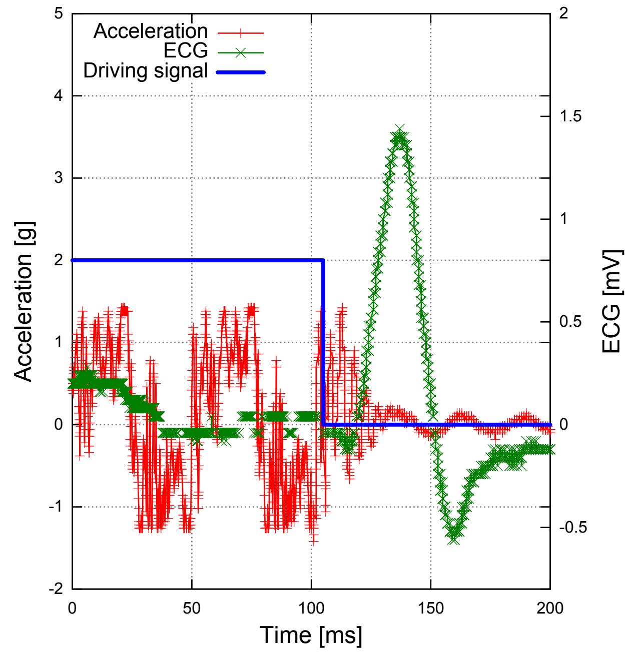


Fig. 5. Experimental results of vibration signal generated by the modified Human Touch HT-1280 massager synchronized with and ECG signal at 30 BPM showing delay between termination of the driving signal and the actual termination of vibrations.

IV. Conclusion

The presented diastolic timed vibrating system is a novel and innovative method for rapid treatment of heart strokes and other low blood flow cases. It has been shown through clinical studies that mechanical vibrations are instrumental in increasing the coronary blood flow, as well as aiding in the improvement of the systolic function of the heart. In case of heart attack, mechanical vibrations along with application of thrombolytic agents can improve clot dissolution and thus increase chances of patient's survival. We presented the first prototype of a diastolic timed vibrator driven by a LabView VI and synchronized with a commercial ECG system. We first demonstrated successful ECG signal synchronization using a simple DC motor. Subsequently, we modified a massager device and successfully verified its functionality as a diastolic timed vibrator as well. The algorithm used in the control Labview VI relied on the triggering of counters based on the length of the systolic and diastolic cycles of a real-time ECG signal. Results showed that we can accurately synchronize mechanical vibrations with real time systolic and diastolic ECG cycles.

V. Future work

Our next goal is to utilize our test system on a live subject to determine the effectiveness of the diastolic timed vibrations on an actual thrombus. Initially, these tests will be performed on animal subjects. We also plan to build a diastolic timed vibrator with integrated miniature ECG acquisition system and a microcontroller with timing algorithms implemented. It would enable us to have a stand-alone device, suitable for clinical use. A further technical challenge will be to reduce the time delay between inactivation command for vibration and vibration to actually stop, such as to enable termination of vibration upon initial sensing of an R wave, to prevent systolic vibration for irregular rhythms.

Acknowledgment

This project was supported by NRC-IRAP Pacific Region in collaboration with Ahof Biophysical Systems Inc. and Simon Fraser University. The authors would like to thank Andrew Hoffman for his suggestions and inputs to the project.

References

[1] National Center of Health Statistics, “Deaths and percentage of total death for the 10 leading causes of death: United states, 2002-2003.” [Online]. Available: http://www.cdc.gov/nchs/data/hestat/leadingdeaths03_tables.pdf

[2] American Heart Association. (2003) Heart attack and angina statistics. Retrieved December 7 2006. [Online]. Available: http://www.americanheart.org/presenter.jhtml?identifier=4591

[3] W. J. Kostuk, “Coronary artery disease - angina, unstable angina, myocardial infarction,” October 2008, discussion paper prepared for The Workplace Safety and Insurance Appeals Tribunal.

[4] M. A. Evans, D. M. Demarais, C. S. Eversull, and S. A. Leeflang, “System and methods for clot dissolution,” U.S. Patent US 6,663,613 B1, December 16, 2003.

[5] G. O. Turner and M. B. Rosin, *Recognizing and Surviving Heart Attacks and Strokes: Lifesaving Advice You Need Now*. University of Missouri Press, 2008.

[6] M. Nedelmann, C. Brandt, F. Schneider, B. M. Eicke, O. Kempski, F. Krummenauer, and M. Dieterich, “Ultrasound-induced blood clot dissolution without a thrombolytic drug is more effective with lower frequencies.” Cardiovascular Diseases, March 2005.

[7] M. Larkin, “Speed is the key to effective heart-attack treatment,” *The Lancet*, vol. Vol 355, p. 472, 2000.

[8] N. Taihei, X. Yoshiro, T. Takehiko, H. Hideyuki, H. Nobuo, K. Hideichi, S. Kunio, K. Hiroshi, and C. Noriyoshi, “Diastolic mechanical vibration on the chest wall increases human coronary blood flow,” *Japanese Circulation Journal*, vol. 58(7), p. 476, 1994.

[9] Y. Koiwa, H. Honda, T. Takagi, J. Kikuchi, N. Hoshi, and T. Takishima, “Modification of human left ventricular relaxation by small-amplitude, phase-controlled mechanical vibration on the chest wall.” *Circulation*, vol. 95, no. 1, pp. 156–162, Jan 1997.

[10] Y. Koiwa, T. Takagi, J. Kikuchi, H. Honda, N. Hoshi, and T. Takishima, “The improvement of systolic function of depressed left ventricle by external vibration at diastole.” *Tohoku J Exp Med*, vol. 159, no. 2, pp. 169–170, Oct 1989.

[11] Y. Koiwa, T. Ohyama, T. Takagi, J. Kikuchi, H. Honda, N. Hoshi, and T. Takishima, “Clinical demonstration of vibration-induced depression of left ventricular function.” *Tohoku J Exp Med*, vol. 159, no. 3, pp. 247–248, Nov 1989.

[12] L. E. Lindblad, R. R. Lorenz, J. T. Shepherd, and P. M. Vanhoutte, “Effect of vibration on a canine cutaneous artery.” *Am J Physiol*, vol. 250, no. 3 Pt 2, pp. H519–H523, Mar 1986.

[13] P. B. Oliva and J. C. Breckinridge, “Arteriographic evidence of coronary arterial spasm in acute myocardial infarction.” *Circulation*, vol. 56, no. 3, pp. 366–374, Sep 1977.

[14] J. Folts, “An in vivo model of experimental arterial stenosis, intimal damage, and periodic thrombosis.” *Circulation*, vol. 83, no. 6 Suppl, pp. IV3–I14, Jun 1991.

[15] Y. Koiwa, R. Hashiguchi, T. Ohyama, S. Isoyama, S. Satoh, H. Suzuki, and T. Takishima, “Measurement of instantaneous viscoelastic properties by impedance-frequency curve of the ventricle.” *Am J Physiol*, vol. 250, no. 4 Pt 2, pp. H672–H684, Apr 1986.

[16] Y. Koiwa, T. Naya, T. Takagi, H. Hond, N. Hoshi, H. Kamada, K. Shirato, H. Kanai, and N. Cyubachi, “Diastolic mechanical vibration on the chest wall increases human coronary blood flow,” *Japanese Circulation Journal*, vol. 58, no. 7, p. 476, 1994.

[17] J. K. Alexander, M. I. Ferrer, R. M. Harvey, and A. Cournand, “The q-t interval in chronic cor pulmonale.” *Circulation*, vol. 3, no. 5, pp. 733–737, May 1951.

Value Engineering for Miane 400/230/63 KV Transformer Station in order to improve quality, optimize costs and project commissioning period.

Sara AMINAEE

Manager director - Better Solution value Engineering consultant Company

Toronto, Ontario, Canada

and

Seyed Ataollah RAZIEI

Project Engineering group, Pad Pay Sazeh Consulting Engineers Company.

Tehran, Tehran, Iran

ABSTRACT

Value Engineering is an identified valuable management procedure, which is used to implement and increase the system output. It is a systematic way to make an effective use of the allotted budget to the projects and to identify the products and services.

Keeping in view the definition, in order to secure the area requirements and the requirements of the Miane Foolad Co, the approval of establishing a 400/230/63 KV substation was obtained. This project was allotted to Azarbaijan Regional Electricity Company as D.B.F.

Based on the estimated credit which was dedicated to this project, it was decided that , the bid documents for the equipment procurement , installation and implementation, and commissioning in 2 ways which are getting finance by the bid winner and using onshore and general process prowling , the bid would be conducted.

In order to implement the project more effectively, reduce the costs and reduce the project commissioning, Azarbayjan Regional Electricity Co. proposed Tavanir Co. (Iran`s power Generation, Transmission and Distribution Management Co.) to use the value engineering processes and with the confirmation of the Value Engineering Committee, a value engineering workstation was established.

In this paper with a brief introduction of Value Engineering, by introducing the project and Value Engineering workstation, we would check in detail the process of this workstation, pointing out the results of the workstation with saving of 27% of required budget of the project.

1-VALUE ENGINEERING

1-1-Definition

According to the American society of Value Engineering, it is a systematic manner with specific techniques which identify the yields of the products and services and defines financial value for lower cost production and secures the risk factors and the required quality. It can be said that value Engineering is an organized effort for analyzing the operation of systems, equipment, services and institute to encourage the actual operation with the lowest cost in the project commissioning period with the relevant quality and suitable safety measures.

1-2- Value Engineering

In this paper to study the executive process of the value engineering, the presented way of the American Society of Value Engineering will be used. This is classified as below:

Pre study workshop:

This phase includes:

1- Recognition and Gathering the information

2- Expansion of a cost model

3- Selecting team members.

4- Preparing study schedule

5- Preparing the work schedule

Work shop job plan:

This phase includes six sub phases:

1- Study and review phase

2- The function analysis phase: Its aim is to expand the division where studying plays an important rule

3- Creativity phase: The purpose of this phase is to expand ideas for the base operation.

4-Evaluation phase: The purpose of this phase is to assess suggested manners and dis-regard the unsuitable ideas.

5- Development phase: The purpose is to prepare the best choice to improve the best idea from the previous phase.

6- Presentation phase: The purpose is to reach to an agreement and define responsibilities for the designers and the project client in following the proposed manner.

The post study phase:

The purpose of this phase is to accomplish the suggestions, which has been confirmed by Value Engineering.

2- PROJECT SCOPE OF WORK

The establishing of the 400/230/63 KV Mianeh substation according to the requirement of the electrical network and providing the necessity of the Folad Mianeh Company by the ministry of planning with the purpose of interchanging electrical energy of Ardebil power plant and neighboring countries and the entire network, resolving the under voltage problem of the region, increasing the stability and network reliability, providing electricity for industrial customers, water pumps for agriculture purposes, reducing the casualties of reactive power and providing the growth of city`s electricity and in order to commissioning in D.B.F has been assigned to the design and development assistant. After receiving the project assignment the act of choosing the forth factor started and the agreement was put in to action.

While choosing the forth factor the land of the substation location was selected near Foolad company. After the base study the land with 320*350 was chosen and was approved for purchase by the relevant committee and after the legal formation the land was given to the project contractor.

By using the forth factor, the consultant of the project was chosen and contract was signed.

- The plan of the Mianeh 400/230/63 KV substation is as below:
- The 400 KV bus bar with the 1.5 switch arrangement.
- Two 400 KV feeders for over head line of the Sahand and Shahid Ghayati power plant
- Two 400 KV feeders for supplying the 35 MVA transformers.
- Two 400/230 KV auto transformers with 315 MVA capacity.
- One 400 KV reactors with 50 MVAR capacity.
- One feeder for supplying 400 KV reactor.

Characteristics of the control room was steel frame, Dech section Slabs, stone flooring, Aluminum windows, 4-millimeter glass, split cooling and heating system.

Accomplishment of the project as assigned has to be done as D.B.F but based on the consultant conditions the tender documents were prepared on D.B.F and cash basis. One tender for wall up activities and surfacing the land and the other tender for buying, installing, testing, commissioning the equipment and also building activities were finalized.

According the high price of the transformers another bid was issued for the purchase of the transformers.

3- VALUE ENGINEERING JOB PLAN

3-1- Pre studying phase:

The first Value Engineering meeting took place with representatives of the responsible units. At that work shop the context and the process of the Value Engineering, the history of the project by consultant and client, an introduce about the base plan, an introduction about the information phase, determining the limits and the limitation of the project, determining the limitations of the Value Engineering and the necessaries of the project (holy cows) , determining the value standards and description of the base plan project by the consultant and the execution of the 400 KV over head network has been discussed and analyzed. Furthermore at that kick of meeting schedule of the workshop has been confirmed.

The research team included the executives, owners and the relevant unit which consist of:

1- Executives: The assistant of planning and development, the executive of the over head lines, the executive of substations and the executive of building.

2-Beneficiaries: The operations assistant of and the technical office of transmission.

Needless to mention that the members were chosen indirectly by the correspondence of the Value Engineering committee with the relevant unit and introducing the representatives by the unit mangers and relevant companies of the workshop.

At that meeting the context and the principles of the Value Engineering were presented nuncupative in order to concentrate on value engineering and get released from their workplaces.

3-2- Work shop phase:

The main study phase is considered to be as the main phase of the Value Engineering and all the process of problem recognition, decision making and taking are done which would be referred to later in this phase.

3-2-1- Information review and study

The shortage of information and usage of incomplete or incorrect information are the main causes of dropping the index value, so in Value Engineering analyzing takes place for having more useful and better information. As result of this, the quality of Value Engineering would increase.

The next step in information phase is defining the aim of the project, the aim of the workstation, the limits of studies, the existing limitations, the standards for choosing and recognition of beneficiaries are as below:

The plans aim:

- The exchange of the produced energy of Ardebil power plant, the neighbor countries and the entire electric network.
- Resolving the under voltage problem
- Increasing the network stability and reliability
- Providing electricity for industrial customers
- Providing electricity for agricultural water sector
- Reducing the casualties of reactive power
- Providing the growth of urban sector electricity
- The aims of workstation:
- Increasing the quality and improvement of plan
- Reducing costs
- Reducing the commissioning period
- Limitations:
- The size of the land according to the location and natural resources was unchangeable.
- The limitation of estimated budget
- The limitation of project commissioning period

Beneficiaries:

- The regional electricity company of Azarbayjan
- Electrical Distribution Company
- The personnel of beneficiary, repairing and maintenance of substations.
- Local industries and farmers
- Project consultant.
- Project contractor.
- Natural resources and environment ministry and the other governmental organizations.
- The nearby residents of substation
- The Value Engineering group after analyzing in whole defined the value standards which would as below:
- Substation reliability
- The correctly functioning of operating system
- The easiness of manufacturing and construction
- The easiness of maintenance and reducing the repairs
- Easiness in maneuver
- Reducing the primary investment

3-2-2- Analyzing the functions (FAST)

Analyzing the functions and drawing the FAST[1] chart was defined as the core of Value Engineering and doing it completely and accurately has great influence in exploring innovation and analyzing the expense of changing new ideas.

The core of Value Engineering after making the different aspects of the project, based on these different sides of the project defined the operation of each part. In this phase, once again, an exploration took place about project and the actions which has to manage and finally the operations were defined in schedule number 1(as per attached)

Drawing the FAST diagram must be done after preparing the detail operation system schedule. Two ways are defined for drawing the FAST diagram, the first one is standard drawing in classic way and the other one drawing with depending on costumers. As seeing in picture number one the final purpose of the project is electricity feeding, and the main operation of the Mianeh substation were recognized as an improvement in connection of 400 KV network, feeding the 230 KV Azarbayjan substation and feeding the 63 KV local substations. (See diagram number 1 as per attached)

The focus points of the group have defined according to two bases. First costly operations in FAST diagram and second the operations which have the wide potential of change. As result of the creativity field of workstation for the idea making phase were defined as below:

- Power transformers
- 400 KV reactor
- 230 KV switchgear
- 400 KV switchgear
- 63 KV switchgear
- The operations

3-2-3-Creativity

After analyzing the operation, the workshop was ready for making ideas.

At first preparing all the team members, providing explanation about parts of the project, operation and preferences were given. Then members of the workstation were asked to brainstorm their ideas for further operation keeping in view the priorities developed ideas and wrote them in their papers. This phase took 120 minutes long and at last about 79 ideas were finalized. After

[1] Function Analysis System Technique

analyzing the ideas which were not relevant the repeated ideas were omitted, and some ideas were merged and finally 24 ideas were chosen to be analyzed.

3-2-4-Development

After identifying the ideas, the member analyzed the remaining ideas and rejected the ideas which were not technically enforceable and also were not relevant to the workstation. There were several cases which had been focused by the group members as below:

The enforceability and the way of executing the changing possibilities in the layout or other cases in implementation of any item.

Detailed cost based on change in ideas

Advantages and disadvantages of each option

The effect of the implementation of each option on value standards and beneficiaries

Finally 7 ideas were chosen. The final ideas were ideas that the Value Engineering team had to analyze according to the defined standards in the information phase. The final ideas were as below:

A: The option which would reduce the cost and increase the value quality standards of and all the group members approved them for executing.

- Omitting the bypass disconnector switch of 230 KV side
- Changing the position of the 230/63 KV transformer to the higher position
- A foresight of further extension for 5, 400 KV overhead line feeders
- Omitting one 230/63 KV transformer with 80 MVA power with its 230 and 63 KV feeder

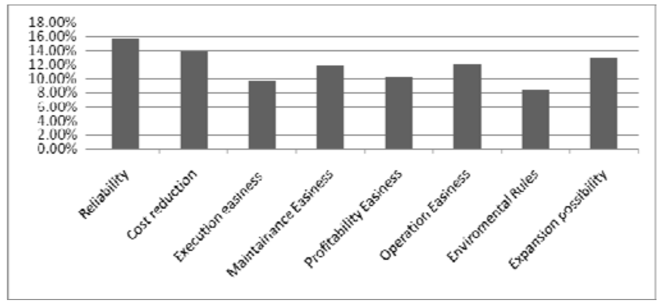
Entrance and exit of 63 KV overhead Foolad line in new substation and omitting one of the 230/63 KV transformer with 80 MVA and non entrance and exit of 63 KV over head line because changing in the route of the overhead line with fed the Foolad Mianeh company

Changing the steep for surfacing the land of the substation in order to reduce the cost

B: The option which would economically, economize but simultaneously reduced the value standards in some cases and the group did not come to a full agreement, so the group presented a briefly result for deciding by the client

- Omitting the reactor
- The plan for changing the position of the reactor to the transformer coil
- Connecting the 400 KV reactor to the over head line by using the disconnector switch
- Changing the 1.5 switch arrangement to the complete 4 switch ring system with the ability to extend to the 1.5 complete switch system in 400 KV side.

For analyzing the final ideas the selection ideas by using the members suggestions and even matrix reviewed for the second time and the priorities of each standards was chosen.



The opinion`s average as regards to the standards

At last the final idea and the analyzing standards scored by the members and the complementary questionnaires were provided to the group members and the performance of each standard was finalized. For each idea according to each standards the variable were multiplied to reach a final variable.

The schedule number 2 shows the score of each option according to value standards.

3-2-5- The extension phase

The schedule number 3 shows the calculated costs of each option.

4- RESULTS

According to the result of the workstations the operations of having special manner in project were easy. Because the preferred options were operative at the same time, thus the group presented their plan which was omitting the bypass disconnector switch of 230 KV side, changing the position of the 230/63 KV transformer to the higher position ,a foresight of further extension for 5, 400 KV overhead line feeders, omitting one 230/63 KV, transformer with 80 MVA power with its 230 and 63 KV feeder, entrance and exit of 63 KV overhead Foolad line in new substation and omitting one of the 230/63 KV transformer with 80 MVA and non entrance and exit of 63 KV over head line because changing in the route of the overhead line with fed the Foolad Mianeh company and changing the steep for surfacing the land of the substation in order to reduce the cost.

References:

[1] S.S.Iyer, Value Engineering usage method, Jebel Ameli, Farat publication , 2004

[2] Tavakoli Reza, Shekari Amir, Value Engineering as a powerful tool, Tadbir journal, Number 132

[3] Teri, Michel, Value management, Mahab Ghods publication, 2005

[4] ARTHURE E MUDGE ."VLUE ENGINEERING " MCGRAW HIL1981

[5]ARTHURE E MUDGE ."VLUE ENGINEERING A SYSTEMATIC APPROACH" MCGRAW HIL1971

S.No	System	Usage	
		Work	Name
1	Outgoing 63 KV feeder	Feeder	Outlet
2	Incoming 63 KV feeder	Feeder	63 KV bus bar
3	Power transformer	Switch	Voltage
4	230 KV transformer feeder	Feeder	230 KV bus bar
5	230 KV outgoing feeder	Feeder	230 KV outgoings
6	Earthing and auxiliary transformer	Switch create	Electrical voltage earth
7	400 KV incoming feeder	Feeder	400 KV bus bar
8	400 KV feeder transformer	Feeder	400 KV transformer
9	Bus bar	Connector	Feeders
10	Neutral reactor	Compensator	Reactive load
11	Protection system	Diagnose	Fault
12	Local control system	Control	Local
13	Remote control system	Control	Long distance
14	DC system	Feeder	DC current
15	AC system	Feeder	AC current
16	Steel structures	Distancing	-
17	Insulator	Formation	Isolation
18	Earth system	Potentialling	Area
19	Guard wires	Protection	Lightning
20	Land	Installing	Equipment
21	Foundations	Resistance	Weight
22	Control Building	Protecting	AC,DC equipments
23	Administrative dept	Office	Personnel
24	Lighting system	Lighting	Area
25	Wall	Physical protection	Necessities

Schedule No.1 – Work analysis

Scoring the options according to the value standards	-	The base plan	Omitting the bypass disconnector switch, Changing the position of the transformer foresight of further extension for 5 feeders	Omitting one transformer with its feeder, entrance and exit of Canan and Foolad over head lines	Imperfect 1.5 switch arrangement in 400 KV side	Changing the angle of entrance and exit of 230 KV over head line	Changing the arrangement of 400 KV side to double bus bar	Omitting the reactor	Land surfacing
Reliability	16.49	88.46	84.62	93.85	39.23	86.92	23.08	17.69	88.46
Cost reduction	14.85	66.15	94.62	97.69	64.62	86.92	50.00	66.92	66.15
Execution easiness	10.09	75.38	93.08	94.62	63.08	82.31	43.85	66.15	75.38
Maintenance Easiness	12.95	80.00	74.62	91.54	54.62	84.62	33.08	59.23	80.00
Profitability Easiness	11.43	76.92	85.38	86.15	66.15	73.85	46.92	63.08	76.92
Operation Easiness	13.04	79.23	83.08	86.92	39.23	83.08	30.00	35.38	79.23
Environmental Rules	9.63	73.86	79.23	83.08	67.69	76.15	50.00	52.31	73.85
Expansion possibility	11.73	60.00	89.23	84.62	51.54	73.85	45.38	53.85	40.00

The comparison done based on the score on value indicators, lifetime expense, primary investment and value final indicator	The base plan	Omitting the bypass disconnector switch, Changing the position of the transformer foresight of further extension for 5 feeders	Omitting one transformer with its feeder, entrance and exit of Canan and Foolad over head lines	Imperfect 1.5 switch arrangement in 400 KV side	Changing the angle of entrance and exit of 230 KV over head line	Changing the arrangement of 400 KV side to double bus bar	Omitting the reactor	Land surfacing
Option`s score according to the standards	*75.50*	*85.59*	*90.33*	*54.61*	*81.62*	*39.25*	*50.27*	*75.50*
Cost of options (Life time expense)	*362.62*	*358.53*	*332.08*	*355.60*	*360.62*	*254.26*	*318.44*	*352.78*
Cost of options (Primary investment)	*263.10*	*260.09*	*240.64*	*257.40*	*261.10*	*350.59*	*230.62*	*260.10*
The value indicator of options	*8.33*	*9.55*	*10.88*	*6.15*	*9.05*	*6.18*	*6.31*	*8.56*

Schedule No 2- Scoring according to Value standards and Cost model of final ideas

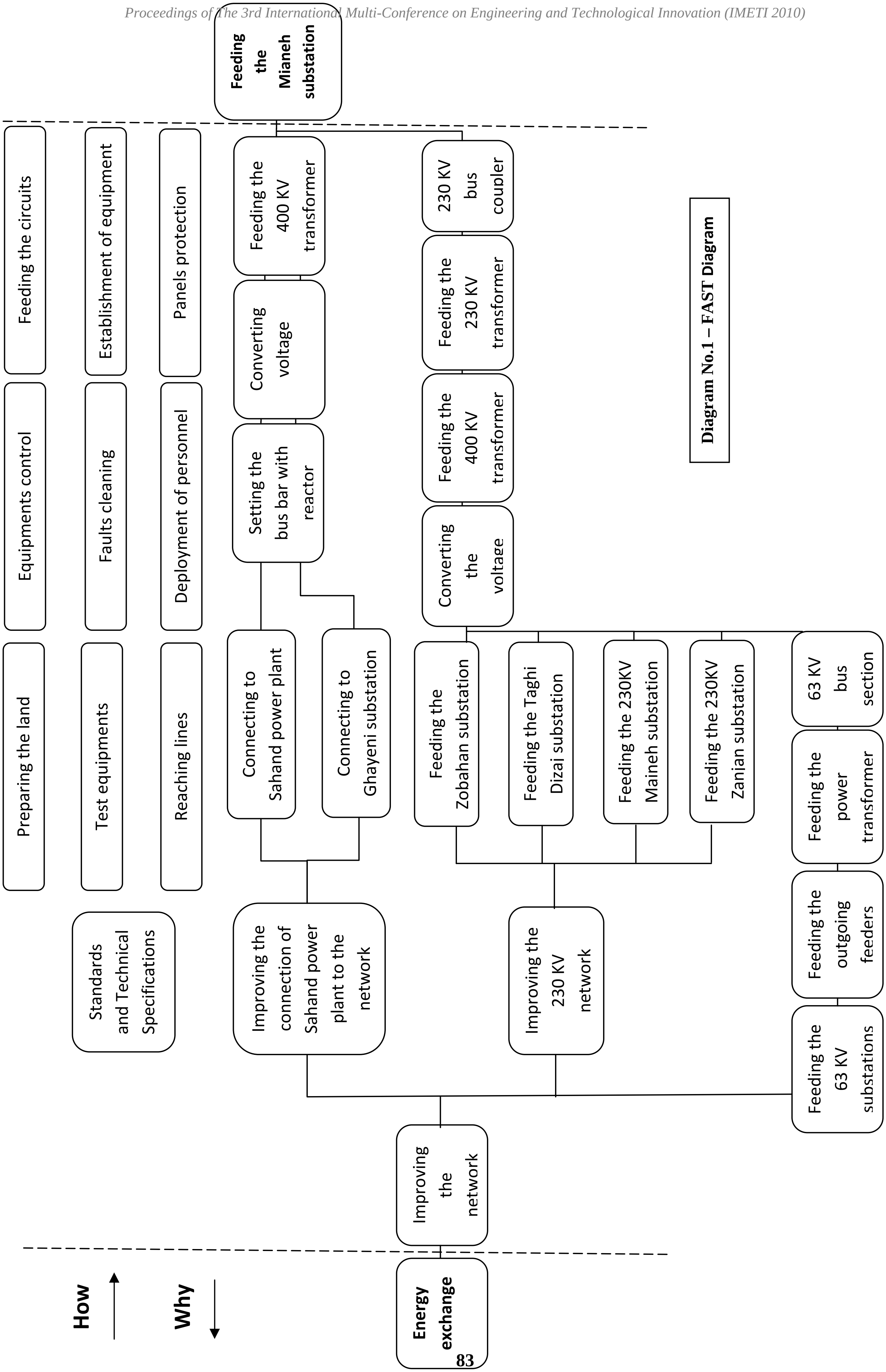


Diagram No.1 – FAST Diagram

Assessment of alternatives and choosing the optimized solution in a 400KV Tabas-Bafgh transmission line project as a good experience for eliminating unnecessary cost.

Sara AMINAEE

Manager director - Better Solution value Engineering consultant Company

Toronto, Ontario, Canada

and

Seyed Ataollah RAZIEI

Project Engineering group, Pad Pay Sazeh Consulting Engineers Company.

Tehran, Tehran, Iran

ABSTRACT

Value Engineering is a systematic procedure with specific techniques, which identifies the product, services, and creates financial value for it, So that the project is commissioned in the most cost effective and quality considering way.

In other words, we can say Value engineering is well-organized method for evaluating the system performance, instruments, services and organizations in order to meet their requirements with lower cost factors throughout the project commissioning coherent with the project quality and safety measures.

During all these years, from the invention of value engineering by Miles, the above method is used in numerous countries for designing and building various products and has been used in many projects to increase productivity and economize solutions for the clients.

This process is introduced and used in Iran for utilities Company. The value-engineering department of Tavanir Co. (Iran`s power Generation, Transmission and Distribution Management Co.) has played an important role as the main core in implementing, planning and using the said method in Electricity sector.

One of the value engineering workshops in this sector is the station used in a 400 KV Tabas–Bafgh transmission line that is designed to reduce costs and increase effectiveness.

1- INTRODUCTION

Value Engineering is provided as a valuable management technical tools which is used for improving systems quality. Value Engineering is a systematic manner in order to optimize the consumption of the allocated budget of the project which recognizes the operation of products and services.

According to the definition after approving the construction of 400 KV over head line of Tabas-Bafgh in order to connect 400 KV network of Khorasan and the entire 400 KV network, improving the voltage quality of Bafgh region, the stability and reliability of entire network, and also transmission of produced energy of Tabas coal power plant to the entire network of Yazd regional electricity in order to improve the plan, reducing cost and project commissioning period, the Value engineering workshop has been suggested for this project, the value engineering workstation was organized.

In this paper, after introducing the Value Engineering in brief we will introduce the project and its workstation, then we will describe the processes and the actions of workstation, at last the result of the workstation will notice.

2- VALUE ENGINEEIRNG

2-1- Definition of Value Engineering technique

According to the American society of Value Engineering, it is a systematic tools with specific techniques which identify the yields of the products and services and defines financial value for lower cost production and secures the risk factors and the required quality. Value Engineering technique is an organized effort for analyzing the operation of systems, equipment, services and institute to encourage the actual operation with the lowest cost in the project commissioning period with the relevant quality and suitable safety measures.

2-2- Value Engineering work shop job plan

In this paper to study the executive process of the value engineering, the presented way of the American Society of Value Engineering will be used. This is classified as below:

Pre workshop phase:

Pre workshop phase includes:

1- Recognition and Gathering the information

2-Create

3- Define scope, purpose, stockholder of project cost model

4- Selecting team members according to project requirement

5- Define criteria

6- Preparing study schedule

7- Define limitations and Holy cows

8- Specify Terminology logy for the project

9- Preparing the workshop schedule

Main study phase:

This phase includes six sub phases:

1- Review pre workshop

2- The Function Analysis system technique phase: Its aim is to expand the division where studying plays an important rule

3- Creativity phase: The purpose of this phase is to expand ideas for the base operation.

4-Evaluation phase: The purpose of this phase is to evaluate and filter and the unsuitable ideas.

5- Development phase: The purpose is to prepare the best choice to improve the best idea from the previous phase.

6- Presentation phase: The purpose is to present and decide for the best value engineering team for the selected design in the following manner.

The ultra studying phase:

The purpose of this phase is to accomplish the suggestions, which has been confirmed by Value Engineering team.

3- PROJECT INTRODUCTION

According to building of Tabas power plant and in order to transmit the produced energy of this power plant to the entire network and also connecting the 400 KV network of Khorasan and 400 KV network of Yazd the 400 KV over head line of Tabas-Bafgh will be built. The mentioned over head line with approximately 295 KM length was analyzed as the base plan (15% of Detailed design) in Value Engineering workstation.

The base plan for building the 400 KV over head line as below:

1- Single copper conductor with double bundle CERLO

2- Two guard wires

As per the consultant`s suggestion the accessories of the project were provided by contractor.

4- TIME SHEDULE OF VALUE ENGINEERING WORKSHOP

4-1- Pre studying phase:

The pre workshop value Engineering took place with representatives of the responsible units. At that workstation the context and the process of the Value Engineering, the history of the project, an introducing the base plan, an introducing the information phase, determining the limits and the limitation of the project, determining the limitations of the Value Engineering and the necessaries of the project (Holy cows), determining facilities and discuss within the value engineering team. the value standards and description of the base plan project has been designed by and discussed within the Value Engineering team. Furthermore at that kick of meeting schedule of the workshop has been confirmed.

Value Engineering team:

1- Executives: The assistant of planning and development, the executive of the over head lines, the executive of substations and the executive of building.

2-Stackholder: The operations assistant of and the technical office of transmission.

Needless to mention that the members were chosen indirectly by the correspondence of the Value Engineering committee with the relevant unit and introducing the representatives by the unit mangers and relevant companies of the workshop.

At that meeting the context and the principles of the Value Engineering were presented nuncupative in order to concentrate on value engineering and get released from their workplaces.

4-2- Workshop phase:

The main study phase is considered to be as the main phase of the Value Engineering workshop and all the process of problem recognition, decision making and taking are done which would be referred to later in this phase.

4-2-1- Review information gathering phase and introduction of the project base plan

The shortage of information and usage of incomplete or incorrect information are the main causes of dropping the index value, so in Value Engineering analyzing takes place for having more useful and better information. As result of this, the quality of Value Engineering would increase.

The next step reviewing information phase is defining the aim of the project, the aim of the workshop, the existing limitations, the standards for choosing and recognition of beneficiaries are as below:

The aims of the project:

Connection the 400 KV network of Khorasan and the 400 KV entire network.

- The voltage quality improvement of Bafgh network.
- Increasing the network stability and reliability
- Transmission of provided energy of Tabas coal power plant to the entire network.

Project necessaries (Holy cows):

Building of 400 KV single line over head line with at least double bundle between Tabas and Bafgh.

The project`s limitations:

- Deserted area of some part of over head line
- Coal mine near the Tabas power plant
- Transmission from near the power plant
- Salt marsh of some part of over head line
- Protected areas

Stakeholder:

The stakeholder of the project are defined as below:

- Yazd regional electricity company
- Khorasan regional electricity company
- Repairing and maintenance personnel of substations
- Local industries and farmers
- project`s consultant
- The contractor of accomplishing the over head line
- Ministry of national resources and environment
- Nearby residents of over head line
- Central core of Value Engineering group of Tavanir Company
- Network management company

The aims of Value Engineering:

- Reducing the cost
- Plan improvement

Limitations of Value Engineering:

- Direction of over head line
- Building cost of the project
- Designing of towers and over head line

After analyzing in whole, the value engineering group selected the value standards which would be the base for assessment as below:

- The physical reliability of project
- Reducing the cost
- Reducing the project commissioning period
- The easiness in maintenance and repairing
- The easiness in preparing the equipment
- Wear out period
- Considering the environmental hazards

4-2-2- Function Analyzing System Technique

Analyzing the functions and drawing the FAST[1] diagram was defined as the core of Value Engineering technique and doing it completely and accurately has great influence in exploring creativity and analyzing the expense of changing new ideas.

The different aspects of the design, based on these different sides of the project defined the operation of each part. In this phase, once again, an exploration took place about project and the actions which has to manage and finally the operations were defined in schedule number 1.

In order to draw out the schedule of analyzing the operation there are several ways (Like from the whole to detail or from detail to the whole and random) The classic way of detail to whole in spite of it being time consuming would assure about having all the parts and operations in FAST diagram. Thus in this workshop the group used this way and the operation schedule which gathered the definition of main, secondary and supporting function were presented.

Drawing the FAST diagram must be done after preparing the detail operation system schedule. There are many techniques for FAST diagram as shown in diagram number 1.

The value engineering team has expanded FAST diagram to define main, secondary and supporting in order to specify the mostly cost operation on FAST diagram (Diagram 1 as per attached)

Thus the creativity phase of workshop for idea making were defined as below:

- To provide the towers and its attachment
- To provide accessories
- To provide phase wire
- To provide guard phase
- To provide insulator

[1] Function Analysis System Technique

- To provide OPGW wire

4-2-3- Creativity

The creativity phase is the most enjoyable stage in Value Engineering workshop. In this phase according to the FAST diagram which was derived from the previous phase the potential point of project has been assigned. The group in whole focused on areas which had the highest potential for increasing the value, thinking and creating new ideas.

There was on emphasize in this phase not to make any evaluation or not to criticize the possibility, efficiency or advantages and disadvantages of each idea and the group should propose any new idea independently avoiding any mental disturbance or self censoring.

The brain group storming technique applied in this phase.

At first preparing all the team members, providing explanation about parts of the project, operation and preferences were given. Then members of the work shop, were asked to brainstorm their ideas for further operation keeping in view the priorities developed ideas and wrote them in their papers. This phase took 240 minutes long and at last about 152 ideas were produced. After finishing the creativity phase each idea were preliminary evaluated. Each idea explained by its proposer in brief and reviewed:

The creativity of the project

- Cost factors
- Cost effectiveness on the project
- Advantages and disadvantages of value standards.
- Client`s agreement.

After explaining each idea and discussing, the group has asked to express their opinion and reaching to an agreement. Finally those ideas which has been filtered and selected for evaluation phase were written in the excel sheet.

After evaluating 152 ideas the opinion of the group was reviewed.

4-2-4- Evaluation phase

In this phase those idea which has been selected by the team were divided among the team for development

4-2-5- Development

During a two week gap in workstation the group analyzed their own ideas and went over the advantages and disadvantages and their costs. Different groups provided their progress reports including the advantages and disadvantages and economized according to Value Engineering system. Then they discussed each idea in detail.

The group categorized and summarized in whole the result of the workshop according to development meeting which was announced in progress form.

A) The options for Tabas-Bafgh overhead line, which as to economize save and increase the value standards and the operations, were agreed in whole.

1) Crossing from near the power plant.

2) Reducing the length by crossing the salty lands.

3) Using the hanging glassy insulators (Fog type)

4) Using the 7NO8 ground wire

B) The options which were economized but reduced the value standards in some situations and the group did not agreed in whole, thus they send a brief report to the owner for their comments.

1) Using the Squap conductor

2) Using the hulter towers

C) The options which had some advantages but due to reduction in value standards, using them were rejected, although they had some value in other projects.

1) Using different kinds of hulter towers

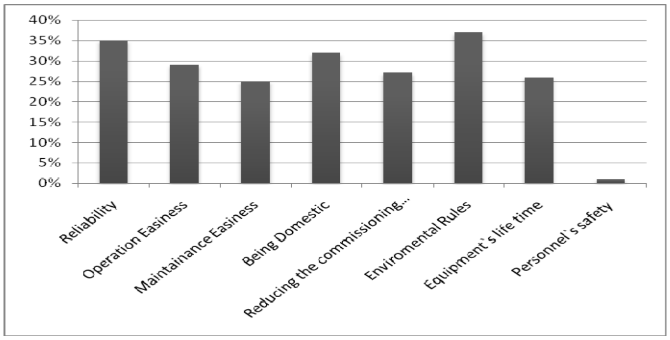
2) Using different towers according to different angles

3) Reducing the foundation volume

For the final analyzing, the selected standards by using the even matrix was reviewed by the members and the priority of each standard was selected.

At last the final idea and the analyzing standards scored by the members and the complementary questionnaires were provided to the group members and the performance of each standard was finalized. For each idea according to each standards the variable were multiplied to reach a final variable.

The schedule number 2 shows the score of each option according to value standards.



4-2-6- The extension phase

The schedule number 3 shows the calculated costs of each option.

5- RESULTS

According to the result of the workstation the operation of having special manner in project was easy. Because the preferred options were operative at the same time, thus the group presented their plan which was passing from near the power plant, reducing the length of the line by passing through the salty lands, using hulty glass isolator and FOG type and reducing the usage of accessories, and also using the 7NO8 system. The plan about changing the cerlo wire to squap and using the hulted towers were presented to the client because of not reaching to an agreement by the group in general.

References:

[1] S.S.Iyer, Value Engineering usage method, Jebel Ameli, Farat publication , 2004

[2] Tavakoli Reza, Shekari Amir, Value Engineering as a powerful tool, Tadbir journal, Number 132

[3] Teri, Michel, Value management, Mahab Ghods publication, 2005

[4] ARTHURE E MUDGE ."VLUE ENGINEERING " MCGRAW HIL1981

[5]ARTHURE E MUDGE ."VLUE ENGINEERING A SYSTEMATIC APPROACH" MCGRAW HIL191

Components	The main operation	The supporting operations			
line course	Creating the foundation	Access to the over head line	Create privacy	-	-
Foundation	Transferring the load to the earth	-	-	-	-
Tower	Transferring the load to the foundation	Wire keeping	Keeping the other components	Maintain the allowed gap	Leading the lightning energy
The conduct wire	Transferring the electrical load	Transferring the information	Tolerance the mechanical loads	Creating the casualties	-
The guard wire	Transferring the lightning energy	Tolerance the short circuit current	Alarming	Tolerance the mechanical loads	-
Insulator	Creating the electrical isolator	Tolerance the mechanical loads	-	-	-
Accessories	Tolerance the mechanical loads	Tolerance the electrical loads	-	-	-
The ground system	Transferring the lightning energy to the earth	Transferring the short circuit current to the earth	-	-	-

Schedule No.1 – Work analysis

Scoring the options according to value standards	The base plan	Using different kinds of halter towers and base plan (1)	Using the towers with special angles according to existed angles	Passing from near the power plant	Reducing the line length (Passing throw the salty lands as possible)	The scope conductor with three bundle	The glass ,foggy ,hang insulator reducing the accessories as possible	Reducing the volume of foundation in hard lands	Using the NO87 wire
reliability	*8.43*	*5.14*	*7.07*	*8.57*	*6.88*	*7.00*	*8.36*	*4.71*	*8.79*
Cost reduction	*5.43*	*7.21*	*3.14*	*9.43*	*9.11*	*6.29*	*9.07*	*1.93*	*8.57*
Reducing the project commissioning period	*6.21*	*4.50*	*2.86*	*9.07*	*7.59*	*6.29*	*7.79*	*4.36*	*7.71*
The easiness in maintenance	*8.14*	*4.21*	*4.93*	*8.43*	*5.27*	*6.43*	*6.71*	*4.43*	*8.29*
Profitability Easiness	*7.57*	*4.64*	*3.14*	*8.79*	*7.41*	*6.79*	*7.43*	*3.36*	*7.79*
Equipment`s life time	*8.14*	*5.79*	*6.57*	*9.00*	*7.77*	*7.64*	*8.14*	*5.29*	*8.93*
Environmental Rules	*6.43*	*5.93*	*7.29*	*8.71*	*8.57*	*8.07*	*8.00*	*7.43*	*8.36*

Schedule No2-Scoring according to value standards

The comparison done based on the score on value indicators, lifetime expense, primary investment and value final indicator	The base plan	Using different kinds of halter towers and base plan (1)	Using the towers with special angles according to existed angles	Passing from near the power plant	Reducing the line length (Passing throw the salty lands as possible)	The scope conductor with three bundle	The glass, foggy, hang insulator reducing the accessories as possible	Reducing the volume of foundation in hard lands	Using the NO87 wire	Passing from near the power plant, reducing the length of the line by passing through the salty lands, using hulty glass isolator and FOG type and reducing the usage of accessories, and also using the 7NO8 system
Option`s score according to the standards	*7.3*	*5.35*	*5.14*	*8.85*	*7.46*	*6.95*	*7.94*	*4.5*	*8.4*	-
Cost of options (Life time expense)	*512,35*	*504,09*	*512,350*	*500,03*	*506,39*	*511,43*	*502,13*	*512,35*	*512,35*	*483,864*
Cost of options (Primary investment)	*423,85*	*413,49*	*423,850*	*411,80*	*418,17*	*415,93*	*413,63*	*423,85*	*423,15*	*395,214*
The value indicator of options	*57.03*	*42.48*	*40.1*	*70.77*	*58.9*	*54.33*	*63.23*	*35.15*	*65.54*	-
Life duration saving	-	*8,258*	-	*12,319*	*5,955*	*915*	*10,212*	-	-	*28,486*
Reducing the investment	-	*10,358*	-	*12,044*	*5,680*	*7,915*	*10,212*	-	*700*	*28,636*

Schedule No3-Cost model of final ideas

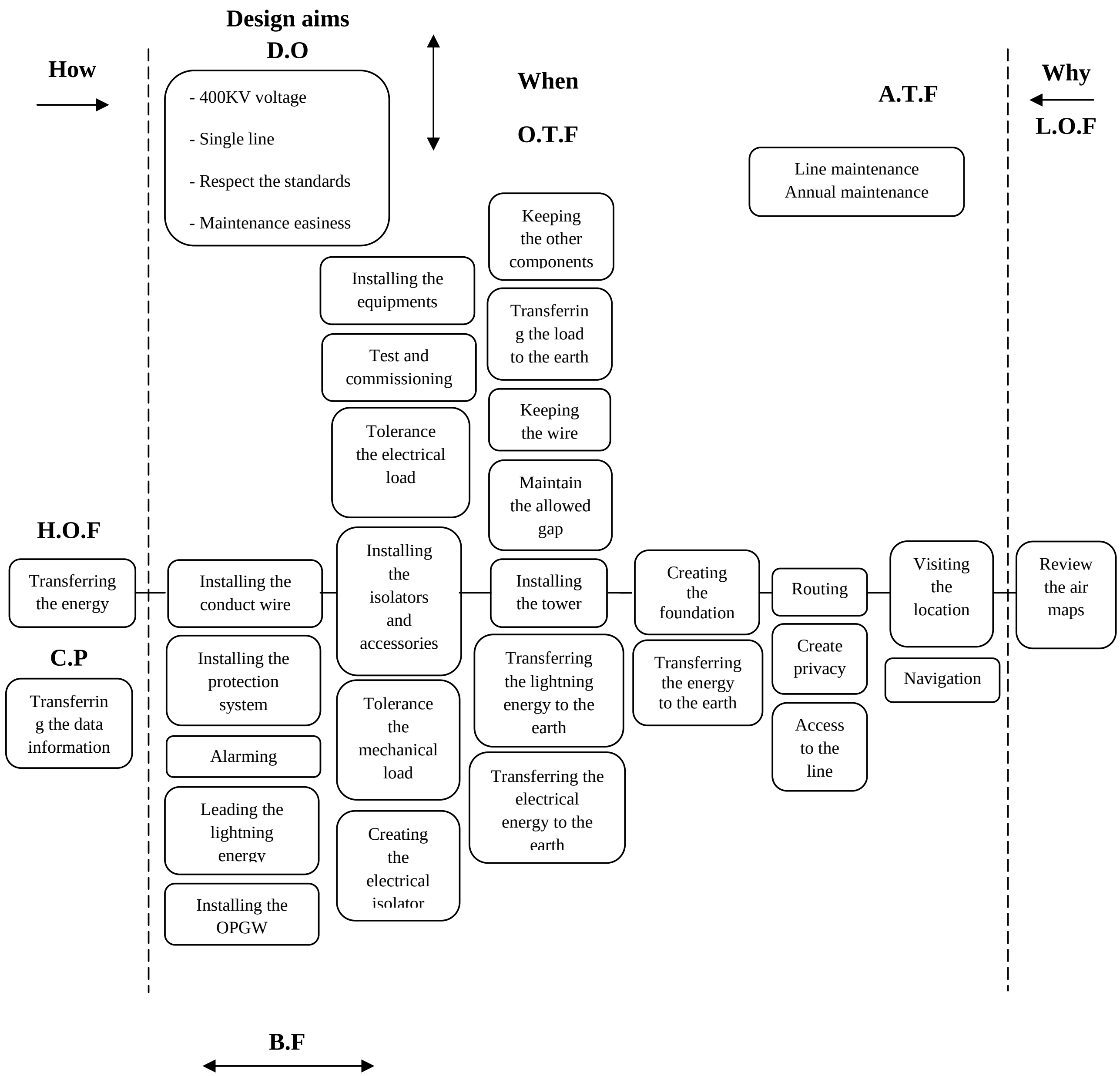


Diagram No.1- FAST diagram

Using Data Mining to Optimize Commercial and Operational Interests: a Case Study of Duke Energy Brazil

Clayton Baltazar, Duke Energy, Av. Presidente Wilson, 118 – Centro, Rio de Janeiro, Brazil

and

Tathiane Ribeiro, Duke Energy, Av. Presidente Wilson, 118 – Centro, Rio de Janeiro, Brazil

and

André Machado Caldeira, DSc, IBMEC/RJ, Av. Presidente Wilson, 118 – Centro, Rio de Janeiro, Brazil

and

Maria Augusta Soares Machado, DSc, IBMEC/RJ, Av. Presidente Wilson, 118 – Centro, mmachado@ibmecrj.br, Rio de Janeiro, Brazil

and

Mihail Lermontov, PhD, UFF, Av. Presidente Wilson, 118 – Centro, Rio de Janeiro, Brazil

and

Rômulo Martins França, MSc, Av. Presidente Wilson, 118 – Centro, Rio de Janeiro, Brazil

and

Thiago Drummond, MSc, IBMEC/RJ, Av. Presidente Wilson, 118 – Centro, Rio de Janeiro, Brazil

and

Walter Gassenferth, MSc, IBMEC/RJ, Av. Presidente Wilson, 118 – Centro, Rio de Janeiro, Brazil

ABSTRACT

This paper presents partial results of a Research Project of data mining for Duke Energy Brazil that is being developed by Ibmec-RJ. The collected data supports a series of tests and studies used in the daily life of the company. It presents a prototype of the software for data mining that is being implemented to optimize Commercial and Operational interests. Through the use of algorithms or search engines, this work try to discover patterns, trends and inference rules in the data to improve optimized decisions by the user. Future papers will present the modules of these tool using neural networks and cluster analysis with the database of Duke Energy Brazil.

Key words: data mining, statistics, association rules, prototype, optimize interests

1. INTRODUCTION

Duke Energy Brazil today has a huge data bank of information covering both the operation, as part of the trade.

This information supports a series of tests and studies used in the daily life of the company.

But as these data were used and manipulated, Duke has realized that there are variables difficult to predictability, since there are inherent relationships between these and other variables, but which are unknown.

The idea of this research is to use data mining techniques to discover these relations above, and not only improve the forecast and actual operation of Duke, but work to improve the development of rules of marketing and operation of the Brazilian electric sector.

The computational tool extraction system of knowledge database using operational and commercial data of Duke is the end product of this research project. Based on the information aspects, the tool will assist in decision making processes of the same.

2. METHODOLOGY

Through the use of algorithms or search engines, try to discover patterns, trends and inference rules in the data. With these rules or functions, the user can make optimized decisions. For this research the files of load, generation, flow and precipitation were prepared and validated.

The validated data were processed in a database using Microsoft Access and created several tables in order to analyze by means of association rules.

The process of extraction of association rules is semi-automatic because it requires the participation of the user defining the data to be analyzed and verification of knowledge discovered, stating whether it is useful or not.

This process aims to extract from large databases, without any prior formulation of hypotheses or unknown data, valid and useful information for decision making.

Support and Confidence are the parameters used in this methodology and must be determined by the user to limit the amount and importance of the extracted rules.

The patterns described by association rules are of the type:

If X then Y (X ==> Y)

Support:

support (X ==> Y) = $P(X \cap Y) \approx$ proportion of cases that occurred in X and Y (both).

Support selects ALL possible rules of type X ==> Y , from the database.

Confidence:

confidence (X ==> Y) = $P(Y / X) = P(X \cap Y / P(X) \approx$ proportion of cases that occurred X and Y,divided by the proportion of times that X occured.

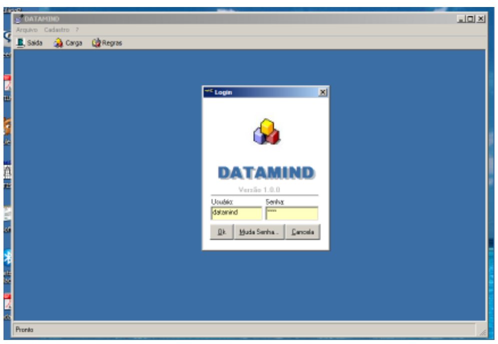
Confidence selects among all possible rules of type X ==> Y, those that occurred Y since X occurred in database.

May be considered interesting with the rules, for example, with support> 10% and confidence > 50%.

3. THE PROTOTYPE

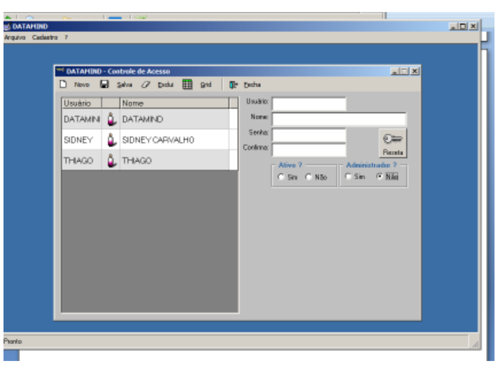
Initial Screen

In this screen the user enters the user name and password as shown below.

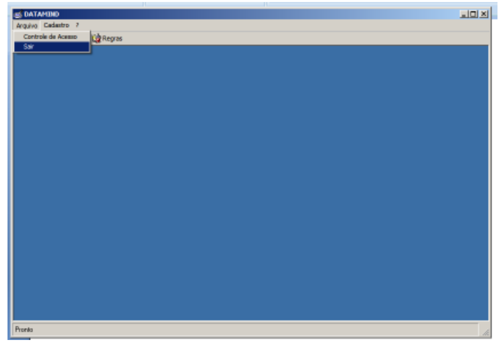


After the enters the name and password, see the following screen:

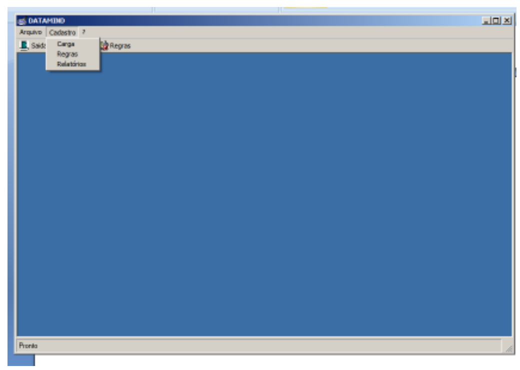
On Tab file, the user has control of his access:



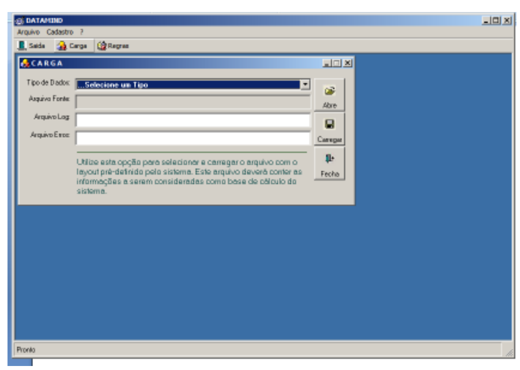
Or exit the system:



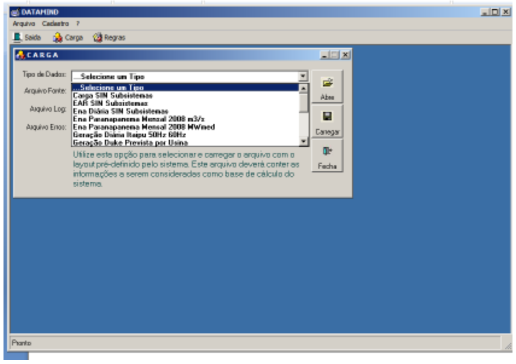
On register tab the user has options to load files, see the association rules and all the reports:



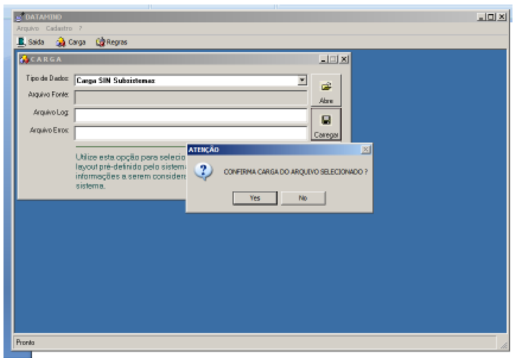
On load, the user selects the data to be analyzed:



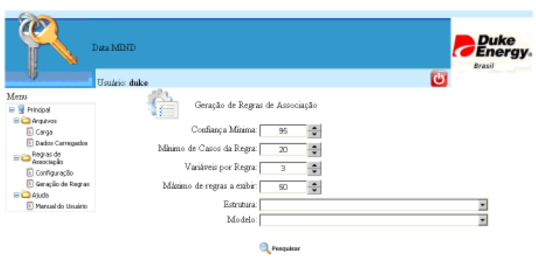
That are:



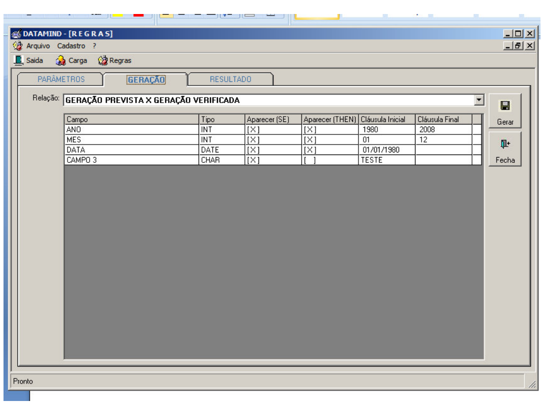
Once selected the data, we have:



On rules ,the user inserts the parameters to generate the association rules in tab parameters:



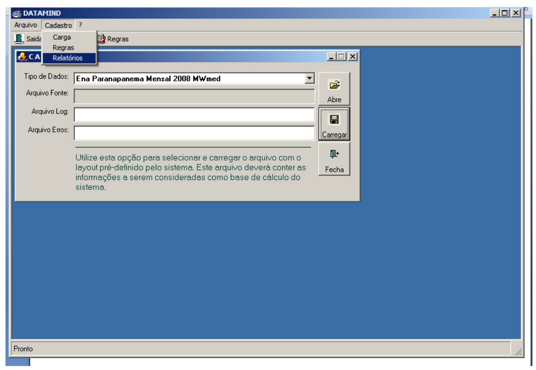
On tab rules the user generates the rules:



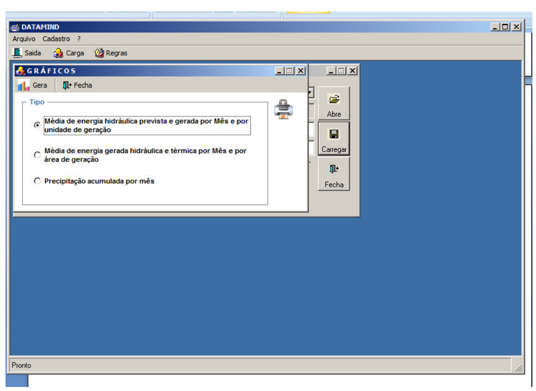
On tab result ,the user has the generated rules:

Regra	Probabilidade	Importância	Suporte
Se INTERCAMBIO ITAIPU 60 SUL **é entre** 143197 e 288061.5 **Então** TOTG **é entre** 1862471 e 3257142	1	0,669006780958576	1
Se TOTGP **é entre** 1715380 e 3017828 **Então** TOTG **é entre** 1862471 e 3257142	1	0,669006780958576	1
Se TOTG **é entre** 1862471 e 3257142 **e** INTERCAMBIO ITAIPU 60 **é entre** 251487.4 e 473114.8 **Então** TOTGP **é entre** 1715380 e 3017828	1	0,669006780958576	1
Se INTERCAMBIO ITAIPU 60 SUL **é entre** 143197 e 288061.5 **e** INTERCAMBIO ITAIPU 60 **é entre** 251487.4 e 473114.8 **Então** TOTGP **é entre** 1715380 e 3017828	1	0,669006780958576	1
Se INTERCAMBIO ITAIPU 60 SUL **é entre** 143197 e 288061.5 **e** INTERCAMBIO ITAIPU 50 **é entre** 202709.4 e 398996.8 **Então** TOTGP **é entre** 1715380 e 3017828	1	0,669006780958576	1
Se INTERCAMBIO ITAIPU 60 SUDESTE < 105279.8 **Então** TOTGP **é entre** 1715380 e 3017828	1	0,669006780958576	1
Se INTERCAMBIO ITAIPU 60 SUDESTE < 105279.8 **e** INTERCAMBIO ITAIPU 50 **é entre** 202709.4 e 398996.8 **Então** TOTG **é entre** 1862471 e 3257142	1	0,669006780958576	1
Se INTERCAMBIO ITAIPU 60 **é entre** 251487.4 e 473114.8 **Então** TOTG **é entre** 1862471 e 3257142	1	0,669006780958576	1

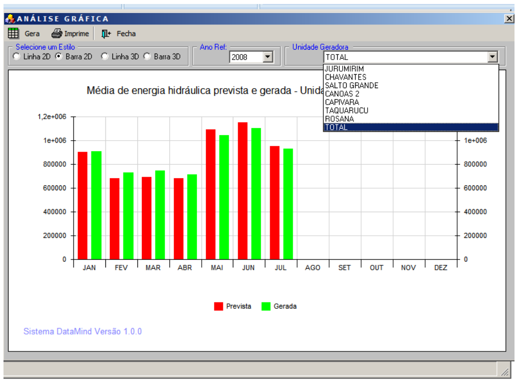
On tab reports the user has:



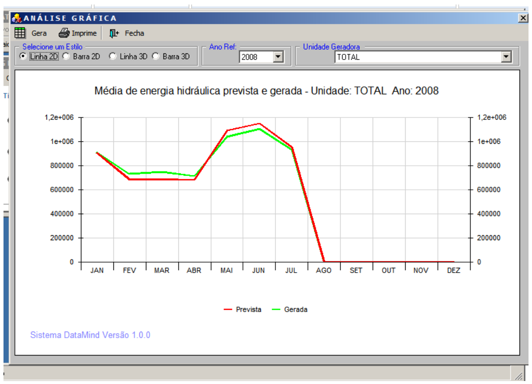
The user has the option to generate graphs selecting "average of hydropower generated and planned for months and unit of generation", "average of hydraulic and thermal energy generated by month and by area of generation" and "accumulated precipitation by month:



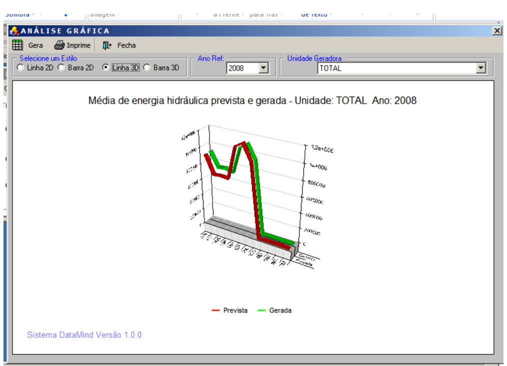
When you click create, you have different types of graphics in bars:



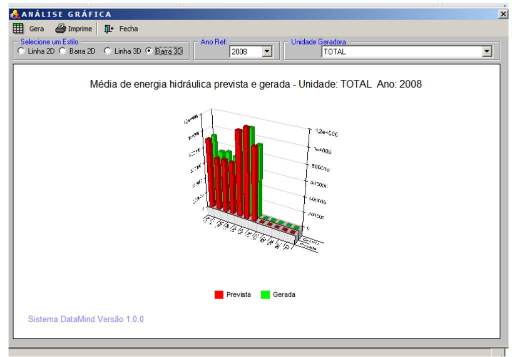
Or in lines:



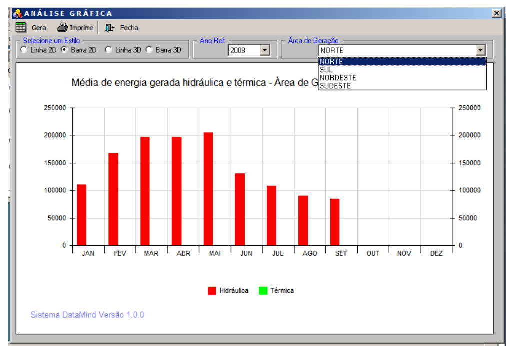
Or in lines , three dimensional:



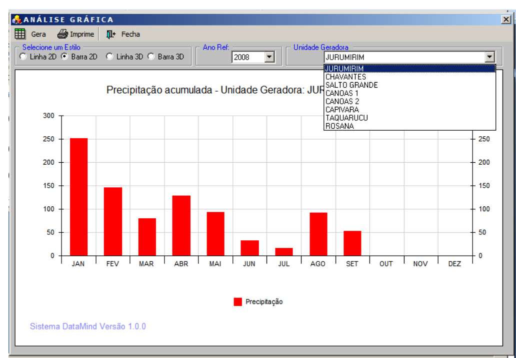
Or in bars, three dimensional:



The user can also select, by area of generation:



Or by generating unit:



4. CONCLUSIONS

This paper presented the prototype of the software for data mining that is being implemented in this project, and future paper will present the modules using neural networks and cluster analysis with the database of Duke Energy Brazil, aiming to improve the quality service to the Brazilian consumer.

5. REFERENCES

[1] Barbosa, Denise Chaves Carvalho, Gassenferth, Walter, Machado, Maria Augusta Soares. *"Data Mining as a Decision Tool for Materials Procurement"*, In: Satyendra Singh (Org.) **Handbook Of Business Practices And Growth In Emerging Markets.** Winnipeg, Canadá: World Scientific Pub Co Inc, 2009;

[2] Cios K. , Pedrycz W. , Swiniarski R. , Kurgan L. , **Data Mining: A Knowledge Discovery Approach**, Springer, ISBN: 978-0-387-33333-5, 2007.v

[3] Krzyztof Koperski, Jiawei Han. "Discovery of spatial association rules in Geographic Information Database", **Proceedings of 4th International Symp. Advances, in Spatial Database, (SSD). Vol 951**, Springer-Verlag, 47-66.

[4] Nabil Adam, Vijay Atluri, Songmei Yu and Yelena Yesha, "Efficient Storage and Management of Environmental Information", **11th Mass Storage Conference hold by IEEE and NASA,** *Maryland,* 2002.

[5] Nogueira, Edgard. "Data Mining usando o SODAS: Um Estudo de Caso" , **Dissertação de Mestrado em Administração** Ibmec-RJ, 2006;

[6] Machado, Maria Augusta Soares. **Inteligência Computacional.** São Paulo: Editora Thomson, 2007;

[7] Medeiros, Valéria Zuma, Machado, Maria Augusta Soares, CALDEIRA, André Machado, Pacheco, Giovanna Lamastra, Gassenferth, Walter. **Métodos Quantitativos com Excel.** São Paulo: Editora Cengage Learning, 2008;

[8] Raymond T.Ng, "Detecting outliers from large datasets", "**Geographic Data Mining and Knowledge Discovery**", Taylor & Francis, 2001

Affordability and Level of Assistance for Economical Weaker Section Group - A Case Study of Surat city (India)

Dr.Krupesh A Chauhan[1]
[1]Associate Professor of CED, S.V. National Institute of Technology, Surat-Gujarat, India,
Dr. N.C. Shah[2]
[2] Professor CED,SVNIT, Surat, Gujarat, India
Dr. S.M.Yadav[3]
[3]Associate Professor of CED, SVNIT, Surat-Gujarat, India,

ABSTRACT

Housing shortage in India is increasing rapidly, mainly because supply is much less than the housing demands. In urban area, the problem is more complex and complicated as the pressure for houses, and services due to both natural increase and migration. The most important resources required to purchase House is finance. Housing plays an important role in a country's economy, typically accounting for 10 to 20 per cent of total economic activity. In this present paper investigations have been made on affordable housing in India. Present paper is a classical example for affordable shelter economical weaker section for metropolitan city like Surat.
Keywords: Dwelling Unit, Housing, Finance, Affordability, Economical Weaker Section.

1. INTRODUCTION

Housing is often an individual's biggest asset. In the housing sector, the problem further compounded by the mobility of the families, both the poor and the rich, to spend enough for the basic services, which also requires finance. The availability of housing finance is, therefore, crucial for overall economic development as well as for a household's welfare and its quality of life. [1]

On the housing front, despite the substantial increase in the housing stocks in the urban areas, the housing situations continues to be daunting and even more complex by the diverse nature of its population and its geographical space. The mind-boggling forecast is that nearly 36% of India's population will be leaving in urban centre by the year 2025. [3] This forecast falls into perspective when we realize that even at present moment, urban India is 24.71 million dwelling units sort as relieve in the draft 11th plan [8] working group report on urban housing which states that the housing requirement over the 11th plan period (2007-2012) is estimated to be at 26.53 million units including the housing shortage as on 2007. Over 97% of the total housing shortages are for EWS and LIG categories.

Any attempt to solve the magnitude of housing shortage has to deal with issue of the urban households belongs to EWS & LIG. As much as 70% of urban population belongs to this two income groups (as per NCEAR study). And as mentioned above, out of total housing needs of 26.53million units in urban areas by 2012, as many as 97% will be required by the EWS & LIG people. [2].In the first National five year plan period (1951-56), housing was introduced in policy framework at the national level. Affordability was emphasized as the key issue and government support through subsidies and loans was deemed necessary and still it is continuing in 11th plan period (2007-2012).

2. STUDY AREA

Surat is one of the fastest growing cities between Mumbai and Ahmedabad corridor. The city of Surat is situated on the bank of river Tapi having coastline of Arabian Sea on its west. Surat is the main center of business and commerce in South, Surat Municipal Corporation (SMC) as shown in Fig.2.1. The area is about 334 sq. km.

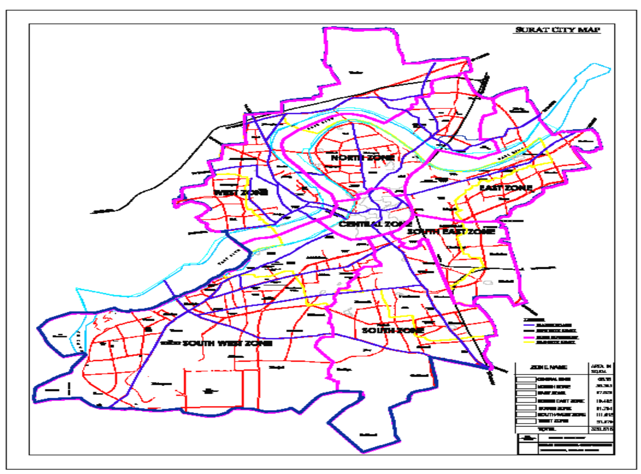


Fig. 2.1 : Surat Municipal Corporation Area

3. OBJECTIVES OF THE STUDY

- To study the affordability
- To investigate level of assistance for EWS in Surat Municipal Corporation Area.

3.1 Structure of Housing Finance Institutes

Housing is a state subject. More states have set up Housing Boards and Slum Clearance/Development Boards at the state level, Prior to the First Five Year Plan[4], housing was mainly dealt by the private sector and some budgetary allocations were made for the housing of government employees. Fig.3.1 shows a schematic representation of the formal housing finance system. [9]

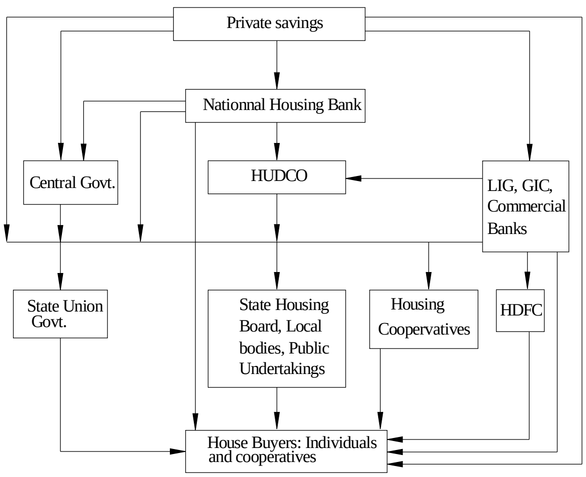


Fig- 3.1 A network of Formal housing Finance System

4. ASPECTS OF AFFORDABLE HOUSING

The government in taking new initiatives to tackle the problem of housing finance. The creation of several Housing Finance Corporation is one such step. The Finance Institution has a mandate to allocate a substantial part of their resources to meet the needs of the poor people. Affordable housing has three major aspects. a)Financial, (Government and public sector, public private partnership, financing through composite credit mechanism, self-help groups and tapping international capital); b)Technical, (cost of material, cost of construction technology and land and infrastructure development cost); c)Policy, (reorientation of central, state and local bodies, strengthening of building center, new approach to construction management). [12] households interviewed had moved into their housing area in the 90's. Keeping that in mind the corresponding norms laid down by 11th plan (2007-12) and HDFC in 2006 Table.5.1 were adopted for the affordability study. [8]

5. THE AFFORDABILITY STUDY

To understand the affordability it was analyzed using Wakely diagrams to indicate the levels of affordability at EWS.

5.1 Income Parameter

The selection criteria for housing areas, those, in which the residents had been staying for at least eight to ten years were selected, so that the residents could develop familiarity with the housing areas. The analysis of the questionnaire showed that a majority of the households interviewed had moved into their housing area in the 90's. Keeping that in mind the corresponding norms laid down by 11th plan (2007-12) and HDFC in 2006 Table.5.1 were adopted for the affordability study. [8]

In the monogram shown in Fig.6.1 housing standards in terms of area are plotted in the Y-axis of the corresponding chart, while the related capital cost is plotted in the X-axis. In the case of EWS category the capital cost as multiple of annual household income is 1.41. as shown in Fig.6.1. This means that the maximum capital cost of a house in this category could be only Rs. 55,836/- (39,600 x 1.41). At Rs.4841 per square meter cost of construction; this income category could afford a maximum of 11.53 square meters, represented in the corresponding scale by point P. The point P is then projected on the chart of capital cost as a multiple of annual household income to meet the line corresponding to 1.41 at Q point. The Q point is then projected horizontally on the curve depicting percentage distribution of household incomes to meet point R. Next, R is projected on the chart below which is showing affordability of percentage of household in the income category. [6]

Table: 5.1 Interest Rates and duration of Loan Repayment for Different Income Categories.

Category	Max. Annual Income (Rs.)	Interest Rate	Loan Duration	Percentage of Income of rent
EWS	39,600/-	5.68 %	15 yrs.	20

Source: 11th plan 2007-12, HDFC

5.2 Capital Cost, Households and Dwelling Size

The minimum cost of construction according to the 2007 prices was considered to be as follows:

- EWS housing: Rs. 4841/square meter
- The household size was taken as 4.4 (NSS report No. 505, January- Jun, 2004) [7]

In the following section, the effectiveness of the level of assistance for EWS is analyzed using Wakely charts.

6. LEVEL OF ASSISTANCE FOR EWS.

Wakely charts are primarily of two types. In the first type of Wakely charts, one can determine the capital cost for housing as a multiple of annual household income, given (i) The borrowing rate of interest, (ii) The duration of loan, and (iii) The household income devoted to rent. The second type of Wakely charts is a combination of four types of charts expressed in the form of a monogram. By using these charts in tandem one can determine the percentage of households who can afford housing of a given standard. The given parameters are: (i) The housing standard in terms area, (ii) The capital cost involved. (iii) The capital cost as a multiple of annual household income and (iv) The percentage distribution of annual household income.

For the analysis of the affordability of the EWS category, first the capital cost of housing as a multiple of annual household income was determined.

The duration of loan is assumed to be for 15 years at 5.68% effective rate of interest according to the prevailing benefit. The household income devoted to rent in assumed to be 20%. As shown in figure: no. 4 the point of intersection of an interest rate of 5.68% and duration of loan of 20 years in the point O. this point is then projected horizontally to the left to meet the corresponding to 20% of the household income devoted to rent to get the point P, and the line is projected further to the left to meet the scale of annual repayments R. The point P is then dropped perpendicular to get the point Q on the scale of capital cost as a multiple of annual household income. In this case C has a value 1. 71 Using the equation p = c x r. One can get a value of p = 1. 71 x 20. This means that the maximum cost of a house a person in this income category can afford is 1. 71 times his annual household income. [5]

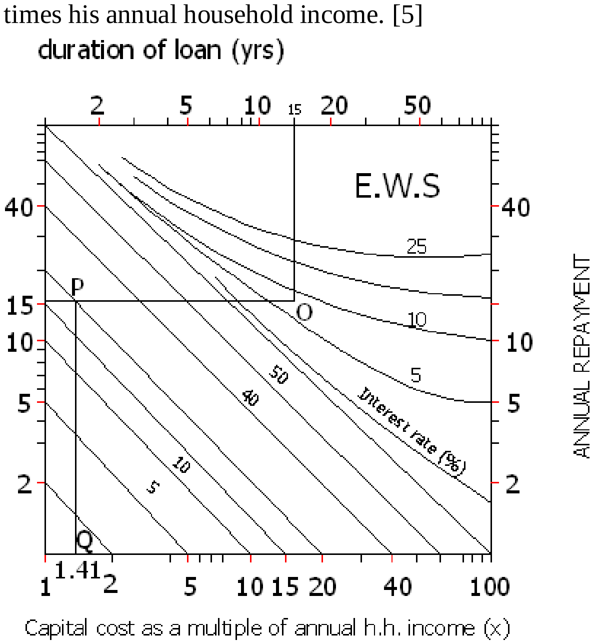


Fig:.6.1 Ability to pay for housing

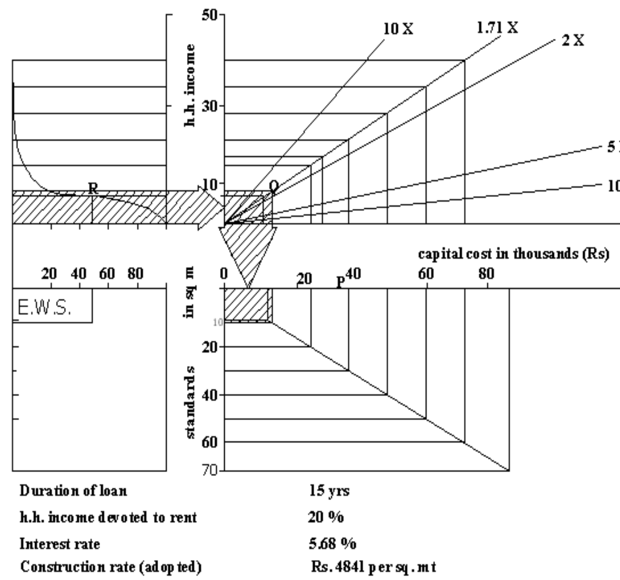


Fig:6.2 Housing finanace for EWS

The affordability bar shows that about 80% of the EWS category could afford the bare minimum of 11.53 square meters, as shown in the monogram Fig: 5. which one is better than prescribed by the National Building Code. The maximum affordable size of22.45 square meters, by the EWS, was, considered by SMC authority. [10] A household size of 4.4 numbers. This means that there would be a very medium occupancy ratio in the rooms, which would affect the satisfaction level.

7. EWS HOUSING PROJECTS

The Surat Municipal Corporation has constructed 7616 dwelling units on 24 different sites during the year 2005-06[11]. The construction work of 7424 dwelling units (DUs) has been completed and possession to the beneficiaries has been handed over. 3533 units are handed over to the beneficiaries by draw The housing design for E.W.S. Category is Ground plus Three Storied RCC framed structure. On each floor, there are four dwelling units. The total Built-up area of single D.U. is 22.45 sq.m. (i.e. 241.00 sq. ft.) The unit has single living room, kitchen, water closet, washing place and balcony. The land cost is not considered as part of total project cost. 1,49,596 sq.mt. of land has been covered for 23 sites. The average housing density is 453 D.U.s./Hector (i.e.2265 PPH).Basic infrastructure like water supply, drainage, pucca roads and street lights are also provided by Surat Municipal Corporation. The construction cost of the single dwelling unit w as Rs.58, 000/- for the project prior to 26 Jan, 2001. While for the post earthquake projects, the building design w as revised and the construction of single dwelling unit rise to Rs.68, 000/-. Rs. 5,000 per D.U subsidy is provided by the Government. the project site photograph is shown in Fig.7.1.

Fig. 7.1 Project Photograph

8. LEVEL OF ASSISTANCE BY SMC CEILING COST

The salient points like income ceiling , built up area, unit cost, govt. subsidy and other details are shown in table 8.1

Table-8.1. The salient points at this level of assistance were as follows:

Scheme	Income Ceiling In Rs per month	Built up Area per Unit (sq.mt.)	Unit cost	Contributions			Installment
				Beneficiary	Govt. Subsidy	Loan component	
EWS	Up to 2500 From 2006-07 Up to 3300	22.45	In case allotment done through draw				
			58,000 Before E.Q.	18,000	5,000	35,000	332 up to15years
			68,000 After E.Q.	28,000			
			In case hutments are shifted				
			58,000 Before E.Q.	1,000	5,000	35,000	332 up to 15 years
						17,000 SMC's without interest loan	94 up to 15 years
			68,000 After E.Q.	1,000	5,000	35,000	332 up to 15 years
						27,000 SMC's without interest loan	150 up to 15 years

The affordability bar shows that 14 % of the EWS as shown in Fig. 6 could not afford a core house with the SMC assistance. 22.45 square meters. The total built-up area of single Dwelling Unit (D U) is 22.45 Sq. m. (i.e. 241. 00 Sq ft.) The unit has single living room kitchen, W.C., washing place and balcony. This shows that SMC levels of assistance was adequate and for the majority of the EWS people a minimum housing was affordable. The actual unit cost after earth quake is 63,000/- due to Government subsidy benefit and the total monthly installment is Rs 482/- as shown in Table 8.1 which one can save easily for monthly income.

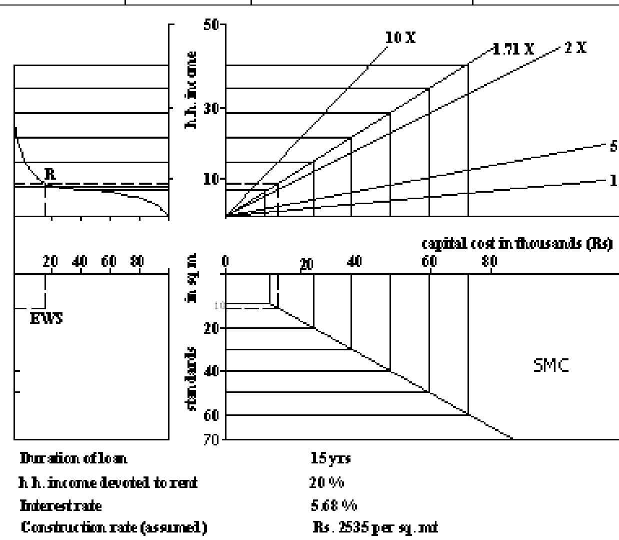


Fig: 8.1. Housing Finance for SMC ceiling costs

9. CONCLUSION AND DISCUSSION

Majority low affordability often results in a "filtration" process in which the housing does not reach the actual target groups and is often occupied by people from a higher income level. It is emphasized that given a better affordability the satisfaction level also raises, a fact hitherto over looked.

The major findings from the study revealed the following

- A minimum standard house is affordable by the majority of the EWS of Surat urban area.
- Looking to the prevailing construction rate and all infrastructural facility provision in EWS project SMC authority have been providing best level of assistance.
- More number of units are required to construct as per demand supply ratio in Surat which one observed during invitation of application form for EWS dwelling unit housing scheme by SMC authority.
- As per study, it seems that EWS are in a better position compared to the LIG. The persons at the highest level of the EWS are in a more favorable position than their immediate higher incomes categories, as they enjoy loans at lower interest rates and have the benefit of higher repayment periods.
- It found that at all levels of EWS assistance, the affordability in significant which satisfaction level of needs to consider the fact that under the present set of condition. The need for an extensive and wide spread institutional networking with accent on the housing finance system.
- Institutional building in the housing finance sector has now assumed critical importance not only in the context of 'affordability', but also for better integration of the housing finance system with the macro finance system.
- The emerging policy intervention of government of India and the changing role of Government from 'Provider' to 'Facilitator' has generated a host of Opportunities and challenges for the key stakeholders.

REFERENCES

[1] Bertrand Renaud, "Housing and Financial Institutions in Developing Countries an Overview", World Bank staff Working Papers, Number 658.

[2] Bhattacharya K.P. (1998), "Affordable Housing and Infrastructure in India"

[3] Census Data (2001), Census of India, Ministry of Home Affairs New Delhi.

[4] D. S. Sudhakar (2007), "Affordable housing and cost effective Technologies", Shelter, Vol-10, HUDCO-HSMI Publication.

[5] Krupesh A Chauhan, Dr N C Shah, Maulik P Jariwala, (2008) "A Study: An Overview On Housing Finance Mechanism and Affordability for LMIG And LIG Housing: Surat Urban Area," First International Conference On Emerging Technologies and Applications in Engineering, Technology and Sciences, Volume-2,- 1930-1935.

[6] Krupesh A Chauhan, Dr N C Shah, S M Yadav (2007) "A case study on Housing Affordability Based Area vs. Income for three Nodes of Surat Urban area" (NSSCM-07) Soft Computing Methodology Ujjain Engg. College Ujjain (M.P.), Volume-1, 32.

[7] National Housing and Habitat Policy (1998), Ministry of Urban Affairs and Employment Government of India, Nirman Bhavan, New Delhi,.

[8] Report on the 11th Five Year Plan (2007), "Working Group on Urban Housing with Focus on Slums", Government of India Ministry of Housing and Urban Poverty Alleviation, New Delhi.

[9] R. N. Sharma (1994), "Indo-Swedish Perspectives on Affordable Housing", Tata Institutes of Social Sciences, Bombay.

[10] Surat City Development Plan (2006-2012), Surat Municipal Corporation (SMC) and Surat Urban Development Authority (SUDA), Surat, 2006.

[11] "Surat vision 2020"; Surat Municipal Corporation (SMC), May 1, 2004.

[12] V. K. Dhar (2004), "Housing Finance and the Urban Poor", Urban Housing: Issues and Development Strategies (Volume-1), UHIDES, Surat.

*   *   *

A Generalized Definition of Jacobian Matrix for Mechatronical Systems

Hermes GIBERTI, Simone CINQUEMANI
Mechanical Engineering Department, Politecnico di Milano,
Campus Bovisa Sud, via La Masa 34, 20156, Milano, Italy

Giovanni LEGNANI
Dipartimento di Ingegneria Meccanica ed Industriale
Università degli Studi di Brescia, via Branze 38, 25129, Brescia, Italy

ABSTRACT

Manipulator kinetostatic performances are usually investigated considering only the geometrical structure of the robot, neglecting the effect of the drive system. In some circumstances this approach may leads to errors and mistakes.
This may happen if the actuators are not identical to each other or when the employed transmission ratio are not identical and/or not constant.
The paper introduces the so called "Generalized Jacobian Matrix" obtained identifying an appropriate matrix, generally diagonal, defined in order to:

1. properly weigh the different contributions of speed and force of each actuator.
2. describe the possible non-homogeneous behaviour of the drive system that depends on the configuration achieved by the robot.

Theoretical analysis is supported by examples highlighting some of the most common mistakes done in the evaluation of a manipulator kinetostatic properties and how they can be avoided using the generalized jacobian matrix.

1. INTRODUCTION

The behaviour of a serial or parallel manipulator can be investigated through its kinetostatic performances [1] such as repeatability, stiffness, maximum force or velocity. They all depend on the kinematic structure of the system, on its configuration in the working space and on the kind of drive system used to operate the robot.
The manipulator may have singular configurations in which the performances in some directions are extremely poor while in others are extremely good. Conversely the manipulator may have configurations where the performances are identical in all directions. This behaviour can be described through the concept of isotropy [2], [3]. Naturally, the design of an isotropic machine is desirable because it assures homogeneous performances in all the directions in terms of accuracy, repeatability, stiffness, maximum force and velocity [4].
The kinetostatic properties of a manipulator, as a function of its position in the workspace, can be analyzed through the jacobian matrix (*J*) [5] or by the ellipsoids of manipulability, strictly related to the jacobian matrix itself [1].
Generally evaluation of isotropy, however, is carried out under the assumption that the behaviour of all the actuators is independent by the pose of the robot and it is the same for all the actuators. This assumption corresponds to exclude the effects of the drive system on isotropy, thus assuming that it depends only on the robot geometry and its reached position [6]. This practice inevitably leads to an incorrect formulation of the problem and to an inaccurate assessment of the isotropy of the system [2].

The paper deeply analyzes this problem introducing the definition of Generalized Jacobian Matrix (*J**): unlike the jacobian matrix, it allows to evaluate the real isotropy of a manipulator taking into account the effects of the drive system on the performances of the robot.

2. MANIPULATOR ISOTROPY

Robot performances are usually measured referring to jacobian matrix J. The function:

$$\underline{f}(\underline{x},\underline{q}) = 0 \tag{1}$$

shows the relationships between the joint space coordinates $\underline{q}$ and the workspace ones $\underline{x}$. Differentiating eq.(1) one gets:

$$J_x \dot{\underline{x}} = J_q \dot{\underline{q}} \tag{2}$$

where:

$$J_x = \frac{\partial \underline{f}}{\partial \underline{x}}; \qquad J_q = -\frac{\partial \underline{f}}{\partial \underline{q}} \tag{3}$$

The jacobian matrix J can be expressed as:

$$J = J_q^{-1} J_x \tag{4}$$

linking velocities of joint space $\dot{\underline{q}}$ with the workspace ones $\dot{\underline{x}}$ as:

$$\dot{\underline{q}} = J\,\dot{\underline{x}} \tag{5}$$

Thanks to the so called kinetostatic duality […] the transposed jacobian matrix represents the relationship between the forces and torques acting on the end-effector F_a and forces and torques exerted by actuators F_q:

$$F_a = J^T F_q \tag{6}$$

Kinetostatic properties of a manipulator, as a function of its position in the workspace, can be analyzed through some indices related to the jacobian matrix [2,5]. Isotropy is one interesting property of a manipulator, since it defines the behaviour of the robot along each direction.

Remembering that the i^{th} singular value $\sigma_i(A)$ of a matrix A is defined as the square root of the eigenvalue λ_i of the corresponding matrix A^TA:

$$\sigma_i(A) = \sqrt{\lambda_i(A^T A)} \tag{7}$$

where is $\lambda_i \geq 0$, isotropy can be "measured" through index:

$$I = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} = \frac{\sigma_{\max}}{\sigma_{\min}} = cond(J) \tag{8}$$

which is the condition number of the jacobian matrix.
When it is verified *cond(J)=1*, the minimum and the maximum eigenvalues coincide and the manipulator is defined as isotropic. Condition on isotropy can also be expressed as [5]:

$$J^T J = kI \tag{9}$$

where k is a scalar and I is the identity matrix. That means isotropy can be achieved when jacobian matrix is proportional to an orthogonal matrix.

This definition, however, is carried out under the assumption that the behaviour of all the actuators is independent by the pose of the robot and it is the same for all the actuators.
This assumption corresponds to exclude the effects of the drive system on isotropy, thus assuming that it depends only on the robot geometry and its reached position. Moreover this classical definition does not consider that some of the gripper and joint coordinates describe rotations and other describes translations and so utilises different units (e.g. degrees and meters).
This practice inevitably leads to an incorrect formulation of the problem and to an inaccurate discussion of the properties of the system like isotropy.
To overcome these problems it is possible to introduce some "characteristic lengths" utilized to normalize the dimension of the manipulator; one length is used to correlate rotation of the TCP with its translation, while a second length is used to compare revolute and prismatic actuators [7]. However the choice of the value of these parameters is arbitrary and some criteria to select reasonable values should be developed.
The paper deeply analyzes the problem of comparing different actuators introducing the definition of *Generalized Jacobian Matrix*: unlike the jacobian matrix, it allows to evaluate the real isotropy of a manipulator, taking into account the effects of the drive system on the performances of the robot.

To better understand these concepts, a case study is presented: it is a 5R 2 dof parallel kinematic machine consisting in 4 links (5 considering the ground) connected by five revolutionary joints (R) two of which are located on ground and driven by motors (Fig.1).
It is constituted by 4 main elements:

1. the support (coloured with light grey), which is fixed and connected to the ground,
2. the driving system (coloured with green), constituted by 2 brushless motors, each actuating a joint,
3. the transmission (coloured with dark grey), which changes the torque and the speed supplied by the motor to the ones requested at joints
4. the manipulator (coloured with light blue), machine consisting in 4 connected by five revolutionary joints

The position $\underline{x} = [x_e;\ y_e]^T$ of the joint C can be expressed as function of the actuated joints coordinates $\underline{q} = [\theta_1;\ \theta_2]^T$.
Figure 2 shows a developed manipulator prototype, whose main feature is the opportunity of changing the distance between the joints connected to the ground (O_1, O_2) through two sliders. The actual configuration allows the two joints to be coincident.
This configuration allows to have the wider workspace which is a circle with its center in the origin ($O_1 \equiv O_2$) and radius $R=L_1+L_2$.

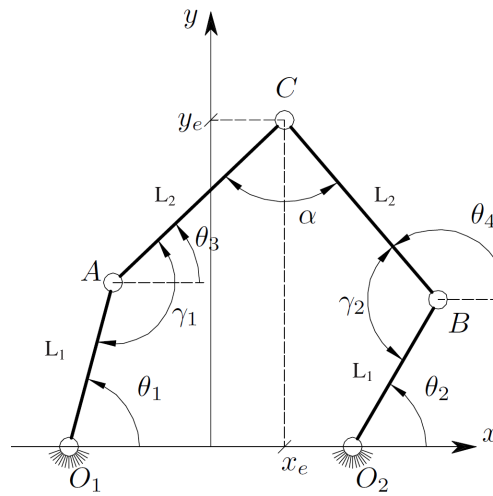


Figure 1 – 5R 2 dof PKM

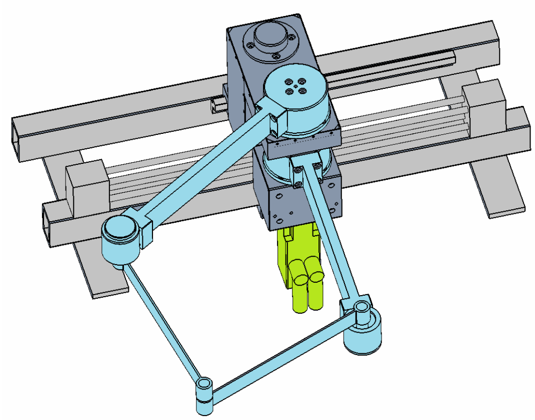


Figure 2 – 5R 2 dof PKM with coincident joints connected to the ground ($O_1 \equiv O_2$)

The graph depicted in Fig.3 highlights the trend of the inverse of the jacobian matrix conditioning number for the configuration described, inside the half workspace.
It is noted that the locus of points where the manipulator is in a isotropic configuration is a circumference. Isotropic behaviour depends only on the distance of the end effector from the origin, while it does not depends on the direction.

Figure 4 is related to a robot configuration corresponding to non coincident joints position. Workspace is reduced and the manipulator behaviour is no more radial symmetric.

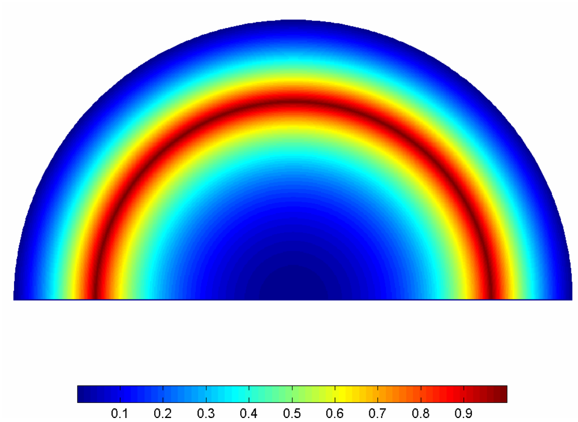


Figure 3 – Evaluation of robot isotropy inside the workspace through the inverse of the jacobian matrix conditioning number (case $O_1 \equiv O_2$).

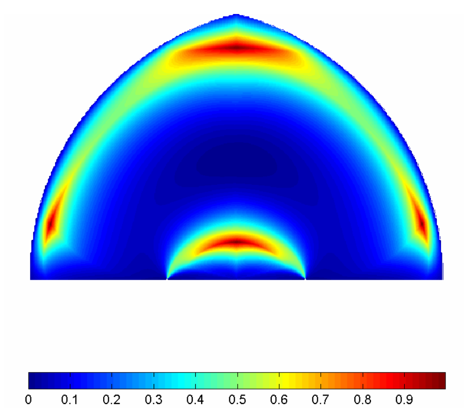


Figure 4 – Evaluation of robot isotropy inside the workspace through the inverse of the jacobian matrix conditioning number (case $O_1 \neq O_2$).

5. THE GENERALIZED JACOBIAN MATRIX

Usually in the kinematics optimization the effect introduced by the behaviour of the drive system is not considered. One of the more frequent cases is when actuators are not identical (i.e. different maximum velocities, different maximum torques, etc.). in this case, instead of analyzing the matrix J^TJ, or its inverse, it will be necessary to consider the *generalized jacobian matrix*:

$$J^* = JD \tag{10}$$

where D is a matrix (generally diagonal) to be defined in order to properly weigh the different contributions of $\dot{Q}$ or F_q [6]. Therefore it is essential for the design of the kinematics of a robot, to evaluate the performance indices previously presented to the matrix DJ rather than J.

The effect described by the D matrix can be better explained remembering some definitions about isotropy [6]:

1. *geometrical isotropy*, it's reached when the manipulator, independently by its drive system, is in an isotropic configuration. In this case:
$$cond(J) = 1 \quad \text{or} \quad J^T J = kI \tag{11}$$
2. *drive system isotropy*, it's achievable if the behaviour of the drive system is the same in all the configurations reached by the manipulator. It holds:
$$cond(D) = 1 \quad \text{or} \quad D^T D = kI \tag{12}$$
3. *effective isotropy*, it's when the robot, driven by a defined drive system, has an isotropic behaviour. In this condition, independently by the condition number of J and D, one gets:
$$cond(JD) = 1 \quad \text{or} \quad J^T D^T DJ = kI \tag{13}$$

Actuators with different performance

When the motors used to drive the manipulator are all of the same type (rotational or linear), it is often assumed that all of them have the same maximum performance both in terms of speed and torque (or force).

Conversely, a manipulator can be driven by actuators of the same type, but different from each others in terms of performances.

To visualize this fact, the generalized jacobian matrix has to be considered instead of the jacobian one. Matrix D should be defined introducing suitable weights to normalize the performances of the actuators. Such a definition is arbitrary, and there is not a "universal" choice which is suitable for all the situations. In [2] it is suggested to define two matrix, D_v related to velocities and D_f related to forces, as:

$$D_v = \begin{bmatrix} \frac{1}{\dot{q}_{1,\max}} & 0 & \cdots \\ 0 & \frac{1}{\dot{q}_{2,\max}} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \tag{14}$$

$$D_f = \begin{bmatrix} \frac{1}{f_{1,\max}} & 0 & \cdots \\ 0 & \frac{1}{f_{2,\max}} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \tag{15}$$

where $\dot{q}_{i,\max}$ and $f_{i,max}$ are respectively the maximum speed achievable by the i^{th} motor and the maximum force deliverable. This choice gives a physical and concrete value to the scaling factors since it depends on the characteristics of the actuators themselves.

Figure 5 shows the manipulator driven by two actuator with different performance ($\dot{q}_{1,\max} = 1.2 \cdot \dot{q}_{2,\max}$).

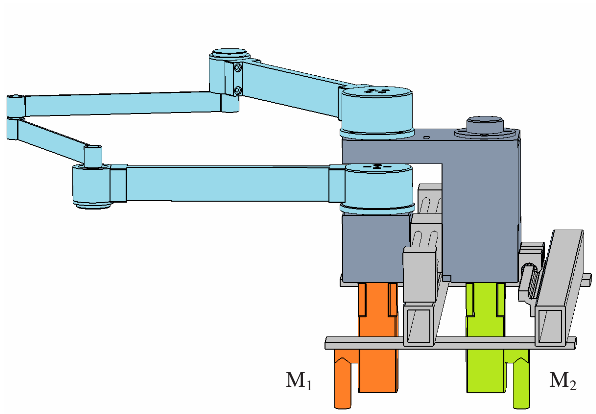


Figure 5 – 5R 2 dof PKM with coincident joints connected to the ground ($O_1 \equiv O_2$) driven by two different motors.

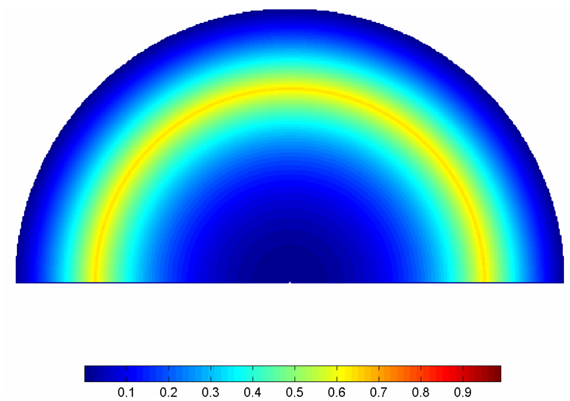


Figure 6 – Effects on robot isotropy of actuators with different performance. Evaluation through the inverse of the jacobian matrix conditioning number .
(case $O_1 \equiv O_2$).

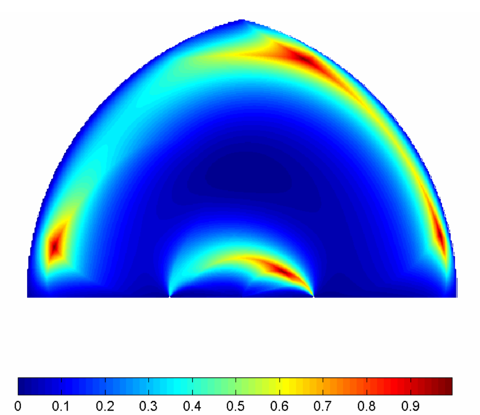


Figure 7 – Effects on robot isotropy of actuators with different performance. Evaluation through the inverse of the jacobian matrix conditioning number.
(case $O_1 \neq O_2$).

Transmissions with different transmission ratio

An effect that is usually neglected in the study of manipulator isotropy is due to the presence of transmissions interposed between the structure of the robot and the actuators. Such transmissions change the torques/forces that motors exert on the structure, as well as the speed that they impose.

This effect necessarily changes the kinetostatic properties of the robot and couldn't be observed by analyzing only the jacobian matrix of the manipulator, but a generalized one should be adopted.

While motors exert torques F_a^* and speeds Q^*, on driven joints are applied forces $F_a = D_a^{-1} F_a^*$ and speeds $\dot{Q} = D_v \dot{Q}^*$ where the matrices D_v, D_f are defined as:

$$D_v = \begin{bmatrix} \tau_1 & 0 \\ 0 & \tau_2 \end{bmatrix} \tag{16}$$

$$D_f = D_v^{-1} = \begin{bmatrix} 1/\tau_1 & 0 \\ 0 & 1/\tau_2 \end{bmatrix} \tag{17}$$

Figure 8 shows a particular of the transmission systems of the considered manipulator. While the two motors are identical, motor M_2 is connected to a belt transmission to the actuated joint. If $\tau \neq 1$ then it results:

$$\tau_1 \neq \tau_2 \tag{18}$$

Suppose $\tau_2 = 2\tau_1$, the effects on robot isotropy are depicted in Fig. 9,10. For both the cases isotropic behaviour is get dramatically worse with respect to Fig.3,4.

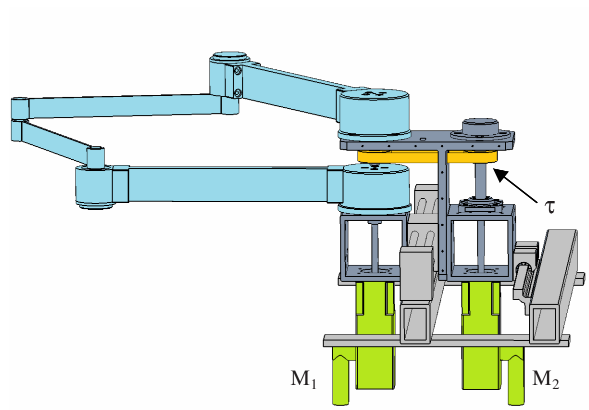


Figure 8 - – 5R 2 dof PKM with coincident joints connected to the ground ($O_1 \equiv O_2$). A particular of the transmission.

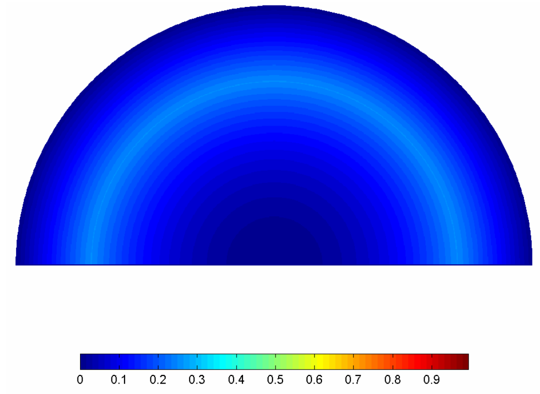


Figure 9 – Effects of different transmission ratios on robot isotropy inside the workspace. Evaluation through the inverse of the jacobian matrix conditioning number .
(case $O_1 \equiv O_2$).

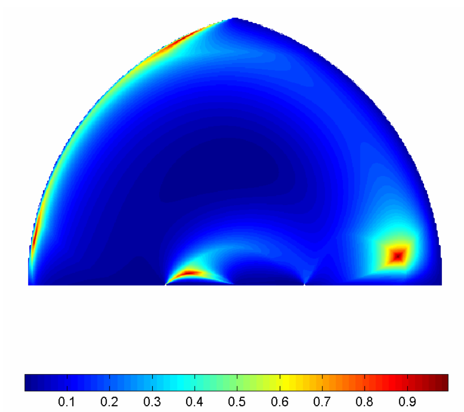


Figure 10 – Effects of different transmission ratios on robot isotropy inside the workspace. Evaluation through the inverse of the jacobian matrix conditioning number
(case $O_1 \neq O_2$).

CONCLUSION

Manipulator kinetostatic performances can be analysed through indices related to jacobian matrix expecially in terms of isotropy.

Generally this investigation is carried out considering only the geometrical structure of the robot, neglecting the effect of the drive system. This approach may leads to errors and mistakes for same manipulators.

Considering forces or velocities, the Generalized Jacobian Matrix, obtained by identifying appropriate matrices D_f and D_v, allow to properly weigh the different contributions of speed and force of each actuator. This operation is performed using some parameters related to the performance of the actuators themselves and therefore gives to the D matrix a physical meaning. Moreover it can describe the possible non-homogeneous behaviour of the drive system that depends on the configuration achieved by the robot.

REFERENCES

[1] T. Yoshikawa, **Foundations of Robotics: Analysis and Control**, 1990, MIT Press, Cambridge, MA,.
[2] Legnani, G., **Robotica Industriale**, Casa Editrice Ambrosiana, 2003.
[3] Salisbury, J.K., Craig, J. J., "Articulated hands: force control and kinematic issues", **International Journal of Robotic Research**, Vol.1, No. 1, 1982, pp.4-17.
[4] Hanrath, G., Stengele, G., **Machine Tool for Triaxial Machining of Work Pieces**, United States Patent N 6328510 B1 (Specht Xperimental 3.axis Hybrid MC), 2001.
[5] Merlet J.P., "Jacobian, Manipulability, Condition Number and Accuracy of Parallel Robots", **J. of Mechanical Design**, Vol.28, 2006, pp.199-206.
[6] Giberti H., Cinquemani S. Chatterton S. "How the Drive System Affects the Kinetostatic Properties of a Robot", **Proc. of 2nd International Multi-Conference on Engineering and Technological Innovation: IMETI 09**, Florida, USA, 2009.

Servo Motors Classification Based on the Accelerating Factor

Hermes GIBERTI, Simone CINQUEMANI
Mechanical Engineering Department, Politecnico di Milano,
Campus Bovisa Sud, via La Masa 34, 20156, Milano, Italy

ABSTRACT

This work is focused on the analysis of the so-called *accelerating factor* (α) [9,10] defined, for each motor, as the ratio between the square of the motor nominal torque and its momentum of inertia. The coefficient α is exclusively defined by parameters related to the motor and therefore it doesn't depend on the machine task: it can be calculated for each motor using the information collected in the manufacturer catalogues.
Actually there is not any theoretical study that investigates the dependence of the accelerating factor on the electro-mechanical characteristics of the motor.
One way to investigate these relationships is to collect information from catalogues of a significant number of motors produced by different manufacturers. This allows to have a statistical population on which perform appropriate analysis. For this reason a database containing more than 300 brushless motors has been created containing, for each record, information on the most important electro-mechanical characteristics.
Using the collected information, some graphs are produced showing how motors having the same size have different accelerating factors.

Keywords: Electric servo-motor; accelerating factor; continuous duty power rate

1. INTRODUCTION

The need to increase production capability, while maintaining the quality standards, requires the implementation of automatic machines performance ever higher.
In this context, it is of strategic importance in the machine design phase the correct selection of the motor-reducer unit.
Unfortunately, the choice of the electric motor required to handle a dynamic load, is closely related to the transmission choice. This operation, in fact, is bound by the limitations imposed by the motor working range and it is subjected to a great number of constraints that depend indirectly on the motor (through its inertia) and on the reducer (through its transmission ratio, its mechanical efficiency and its inertia), whose selection is the object of the design.
In literature there are many procedures for the selection of a motor-reducer unit [1-8] that, while all start from the same theoretical basis, they differ from their approach to the problem.
This work is focused on the analysis of the "continuous duty power rate" (also called "accelerating factor") [8] defined, for each motor, as the ratio between the square of the motor nominal torque and its momentum of inertia. Each manufacturer of brushless synchronous motors adopts its own technological solutions and its constructive layout, which is generally different from those of another producer. However, the designer of an automatic machine who has to choose a motor, can consider all the motors as black boxes characterized by their "accelerating factor". Actually there is not any theoretical study that investigates the dependence of the accelerating factor on the electro-mechanical characteristics of the motor, therefore the comparison of motor performances in terms of their accelerating factor is possible only in relative terms and not in absolute.
The aim of this paper is to put the groundwork for a deeper analysis of the accelerating factor, in order to give to the designer of an automatic machine a tool to critically evaluate the performance of motors, not only as compared to the other available, but also in absolute terms.

Table 1 – Nomenclature

T_M	motor torque
J_M	motor momentum of inertia
$T_{M,rms}$	motor root mean square torque
$T_{M,N}$	motor nominal torque
$T_{M,\max}^{TH}$	motor theoretical maximum torque
$T_{M,max}$	servo-motor maximum torque
ω_M	motor angular speed
$\omega_{M,N}$	motor nominal angular speed
$\dot{\omega}_M$	motor angular acceleration
P_N	motor nominal power
m	motor mass
V_N	motor nominal voltage
p	motor poles
T_L	load torque
J_L	load momentum of inertia
T_L^*	generalized load torque
$T_{L,rms}^*$	generalized load root mean square torque
$T_{L,max}$	load maximum torque
ω_L	load angular speed
$\dot{\omega}_L$	load angular acceleration
$\dot{\omega}_{L,rms}$	load root mean square acceleration
τ	transmission ratio
η	transmission mechanical efficiency
α	accelerating factor
β	load factor
$\omega_{M,max}$	maximum speed achievable by the motor
$\omega_{L,max}$	maximum speed achieved by the load
t_a	cycle time
C_{th}	motor thermal capacity
R_{th}	motor thermal resistance
τ_{th}	motor thermal constant
K_T	torque constant
i	current flowing in motor windings

2. THE MOTOR

Brushless motors (Fig.1) are the most widespread electrical actuators in automation field, which working range (Fig.2) could be approximately subdivided into a continuous working zone (called *S1*, bound by motor rated torque) and in a dynamic one (called *S6*, bound by the maximum motor torque $T_{M,max}$). Usually the motor rated torque decreases with the motor speed ω_M. To simplify the rated torque trend and to have a cautionary approach, the continuous working range is approximated to a rectangle, identifying two values $T_{M,N}$ and $\omega_{M,max}$ (Fig.3).

Fig.1 - Commercial brushless motor

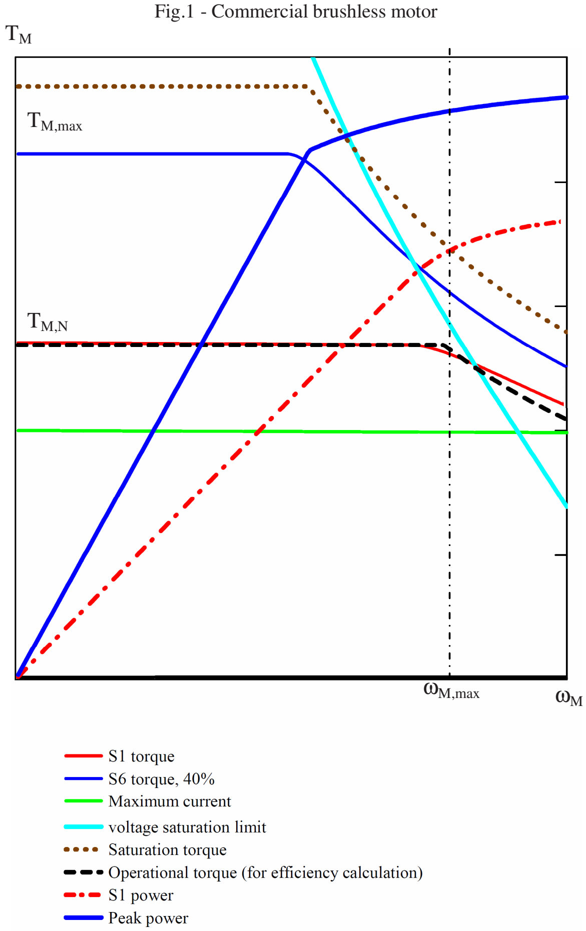


Fig.2 - Speed/torque curve of a common brushless motor

Note how the approximation used to make the *S1* field rectangular, actually has consequences on the value of $T_{M,N}$ and $\omega_{M,max}$.

Information on catalogues are often poor and, in the best case, when speed/curve torque is available, they should be managed to obtain the interesting parameters.
Note how the maximum torque achieved by the servo-motor $T_{M,max}$ strongly depends on the drive associated with it and it is generally different from the motor theoretical maximum torque $T_{M,\max}^{TH}$.

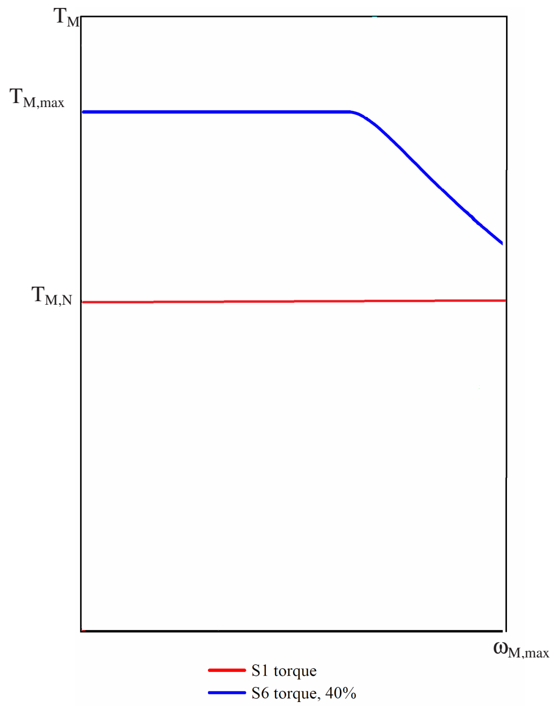


Figure 3 – Approximated speed/torque curve

At low speed, the constraint introduced by the drive systems is related to the maximum current supplied to the motor. Since torque depends on the current, this limit translates into a horizontal line on the motor working field corresponding to a maximum torque different from the theoretical one. At higher speed, this constraint is overcome by the condition on the maximum voltage endurable, which causes a reduction of the motor maximum torque with its speed.

3. THE THERMAL PROBLEM OF ELECTRICAL MOTORS

The thermal problem is of great importance in electric motors, and is generally the most binding condition in the choice of an electric motor for industrial applications.
During their operation, in fact, motors waste power W_d as heat: this is primarily because the windings are affected by the current flow (copper losses), but also for the eddy currents (iron losses) and mechanical effects.

The power lost as heat determines an increase in temperature of the motor. Heat is partially removed from the environment at least until a stationary condition is determined.
Naming θ(t) the difference of temperature at time t between motor and environment, C_{th} the motor thermal capacity and R_{th} its thermal resistance, the differential equation for the equilibrium of power is:

$$C_{th}\frac{d\theta}{dt}+\frac{\theta}{R_{th}}=W_d \tag{1}$$

that can be rewritten as:

$$\tau_{th}\frac{d\theta}{dt}+\theta=W_d R_{th} \tag{2}$$

where:

$$\tau_{th}=R_{th}C_{th} \tag{3}$$

is the motor thermal time constant (usually defined by the manufacturer and available on catalogues).

Observance of the constraints relating to thermal problem requires, when selecting a motor, that the maximum temperature reached during the operation does not exceed the maximum permissible. This requires solving Eq. (1).
However, if the task operation is cyclic, with period $t_a << \tau_{th}$ the problem can be simplified. In this case, the motor is not able to follow the fast thermal fluctuations of the power dissipation, due to high heat resistance. The temperature of the motor, then, evolves as if it were subject to constant power dissipation $\overline{W}_d$ equal to the average power dissipated in the cycle.
Assuming that the dissipation is related mainly to the Joule effect due to the resistance R, it is:

$$\overline{W}_d=\frac{R}{t_a}\int_0^{t_a} i^2 dt \tag{4}$$

$$T_M=K_T i \tag{5}$$

where K_T is the torque constant.
By substituting eq.(5) in eq.(4), it is possible to reach the value of the so called motor root mean square torque:

$$T_{M,rms}=\sqrt{\frac{1}{t_a}\int_0^{t_a} T_M^2 dt} \tag{6}$$

namely the torque, acting steadily over the cycle, which is attributable to the total energy dissipation really occurred in the cycle.
The condition on the thermal problem becomes:

$$T_{M,rms} < T_{M,N} \tag{7}$$

where motor torque $T_{M,N}$ is obtainable from catalogues given by motor manufacturers and it is defined as the torque that can be supplied by the motor for an infinite time, without overheat.

The motor torque T_M can be written as:

$$T_M=\tau T_L^*+J_M\frac{\dot{\omega}_L}{\tau} \tag{8}$$

where:

$$T_L^*=T_L+J_L\dot{\omega}_L \tag{9}$$

is the generalized resistant torque at the load shaft.
When selecting the motor-reducer unit, the transmission ratio τ and the motor inertia J_M are still unknown. In this phase, transmission is considered ideal (η=1).

Equation (8) highlights the dependence of the motor torque on this variables, while from eq.(9) it's possible to observe that all the terms related to the load are known.
The root mean square torque is obtained from:

$$T_{M,rms}^2=\int_0^{t_a}\frac{T_M^2}{t_a}dt=\int_0^{t_a}\frac{1}{t_a}\left(\tau T_L^*+J_M\frac{\dot{\omega}_L}{\tau}\right)^2 dt \tag{10}$$

Developing the term in brackets and using the properties of the sum of integrals, it's possible to reach the root mean square torque as:

$$T_{M,rms}^2=T_{L,rms}^{*2}\tau^2+\frac{J_M^2}{\tau^2}\dot{\omega}_{L,rms}^2+2J_M\left(T_L^*\dot{\omega}_L\right)_{mean} \tag{11}$$

and inequality (7) can be written as:

$$T_{M,N}^2\geq T_{L,rms}^{*2}\tau^2+\frac{J_M^2}{\tau^2}\dot{\omega}_{L,rms}^2+2J_M\left(T_L^*\dot{\omega}_L\right)_{mean} \tag{12}$$

4. THE MOTOR ACCELERATING FACTOR

Since $T_{M,N}$ is positive by definition, one can gets:

$$\frac{T_{M,N}^2}{J_M}\geq\frac{T_{L,rms}^{*2}\tau^2}{J_M}+\frac{\dot{\omega}_{L,rms}^2}{\tau^2}J_M+2\left(T_L^*\dot{\omega}_L\right)_{mean} \tag{13}$$

Let's introduce the *accelerating factor* of the motor:

$$\alpha=\frac{T_{M,N}^2}{J_M} \tag{14}$$

describing the performances of each motor, and the *load factor*:

$$\beta=2\left[\dot{\omega}_{L,rms}^2T_{L,rms}^*+\left(T_L^*\dot{\omega}_L\right)_{mean}\right] \tag{15}$$

defining the performances required by the task. The unit of measurement of both factors is *W/s*.
The coefficient α is exclusively defined by parameters related to the motor and therefore it doesn't depend on the machine task: it can be calculated for each motor using the information collected in the manufacturer catalogues. Moreover it could be reported on them, to provide a classification of the commercial motors on

the basis of this standard. Otherwise, the coefficient β depends only on the working conditions (applied load and law of motion) and it's a measure that defines the power required by the system.

Substituting α and β in inequality (13) we reach:

$$\alpha \geq \beta + \left[T^{*}_{L,rms} \frac{\tau}{\sqrt{J_M}} - \dot{\omega}_{L,rms} \frac{\sqrt{J_M}}{\tau} \right]^2 \qquad (16)$$

Since the term in brackets is always positive, or null, the load factor β represents the minimum value of the right hand side of eq.(16). It means that the motor accelerating factor α must be sufficiently greater than the load factor β, so that inequality (7) is verified.

A motor must be rejected if $\alpha<\beta$, while if $\alpha\geq\beta$ the motor can have enough rated torque if τ is chosen properly.

The preliminary motor choice is conducted comparing only the values α and β; these values are easily calculated knowing the mechanical properties of the motor and the load features.

5. SERVO MOTOR COMPARISON

The aim of this work is to put the basis of a detailed analysis of the accelerating factor (or continuous duty power rate). The starting point is the answer to the question: *"Let's assume that motors with different sizes are hard to be compared, may similar motors have accelerating factors* α *extremely different?"*

A negative answer to this question would make unnecessary any subsequent consideration, indicating that manufacturing parameters marginally influence the accelerating factor. That means the commercial brushless motors currently on the market have similar electromechanical features, presumably best suited to obtaining high values of α.

On the opposite, a positive answer would open a research field to find which are the electromechanical features of a motor that most influence the accelerating factor and which is (if it exists) the theoretical or technological value of α whose overtaking is impossible or technically not convenient.

A way to answer the question is to collect enough information from different manufacturers catalogues for a significant number of motors. The resulting database will be a useful instrument to compare different commercial devices and a suitable tool to highlight how the accelerating factor can't be the only parameter to describe the performance of a motor and how all motors features influence the design of a machine.

The database

The database collects the main information available on catalogues of about 300 motors whose power is between 15[W] and 15[kW]. Information collected relate to: brand, model, type of motor (AC or DC), torque coefficient, winding electrical resistance, number of poles, geometrical dimensions and, naturally, motor nominal torque and the rotor momentum of inertia.

The momentum of inertia J_M includes the inertia of the rotor and the one of the positioning sensor, a needed component for the machine functioning and thus a part of it. The inertia of any brake systems, or related to any additional sensors is neglected.

Data analysis

Figure 4 represents the trend of the accelerating factor (y axis) for the entire population of considered motors (x axis).

Motors are identified by a unique growing index. Notice how α can assume values really different, and how some motors have an accelerating factor extremely high compared to the considered population.

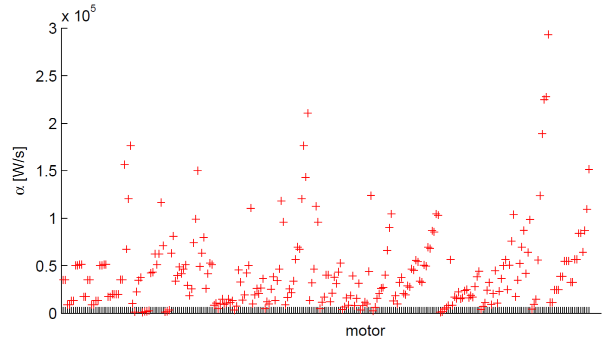


Figure 4 – Accelerating factor (α) for the motors in the database

The graph can not highlights if this high values of the accelerating factor are due to a high nominal torque, or to small rotor inertia, or to the combination of the two factors. It is also unclear whether the accelerating factor is related to the motor size or not.

For this reason, values of α, motor nominal torque $T_{M,N}$ and motor momentum of inertia J_M are reported on the same chart for all the motors in the database (Fig.5). For ease of consultation, motors are ordered with increasing momentum of inertia.

The three series of data are normalized on their respective maximum value to allow a comparison between series.

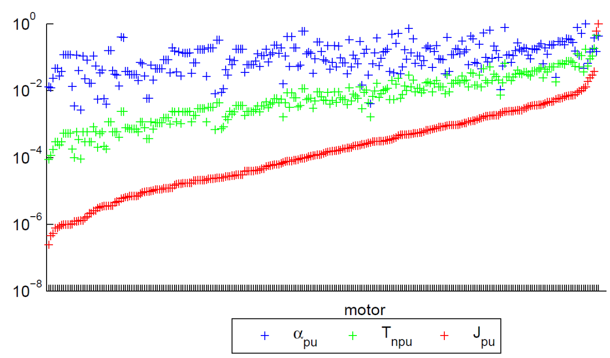


Figure 5 – Normalized accelerating factors (α), motor nominal torques ($T_{M,N}$) for the considered motors ordered with increasing momentum of inertia J_M

Figure 6 depicts the trends of motor weight (M) and nominal torques ($T_{M,N}$) for the considered motors ordered with increasing momentum of inertia J_M.

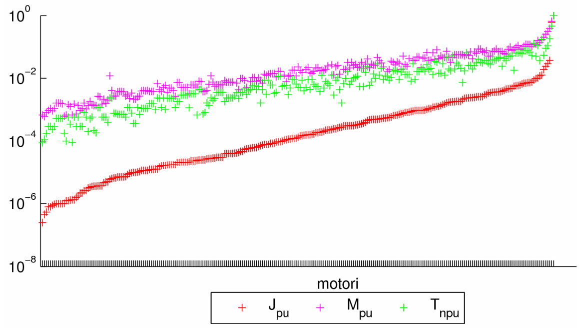


Figure 6 – Normalized masses (M), motor nominal torques ($T_{M,N}$) for the considered motors ordered with increasing momentum of inertia J_M

Looking the chart in Fig.5, 6 some interesting consideration can be done:

1. High values of accelerating factor can be obtained, even with high values of J_M, because of increased motor nominal torque;
2. Motor nominal torques and rotor inertia seem to be proportional each others;
3. Motors with same momentum of inertia can have accelerating factors extremely different;
4. Motors momentum of inertia and mass seem to be proportional

These considerations give a first answer to the question done: commercial brushless motors are built with different designs and generally have different performances.
It means that some motors are better than others.

Table 2 shows, as example, the main features of motors classified as no.44 and no.230. Despite their different characteristics and dimensions, the two motors are identical at least as regards their accelerating factors.

Table 2 – Comparison between two selected motors

N	44	230
Brand	Mavilor	TAMAGAWA
Model	MA-3	TS4813
$T_{M,N}[Nm]$	1,3	3,3
$J_M[Kgm^2]$	4,00E-05	2,60E-04
$\alpha[W/s]$	42250,00	41884,62
$x[mm]$	120	100
$z[mm]$	82	158
$P_N[kW]$	1,22	1,04
$M[kg]$	1,9	4,7
$\omega_{M,N}[rpm]$	9000	3000
$V_N[V]$	110	400

Actually, this conclusion is the starting point to investigate what are the electromechanical characteristics that allow a motor to be more performing.

Let's now consider the transmission that could be coupled with each motor, such that condition on thermal problem is verified. The range of suitable transmission ratios can be calculated by solving eq.(16).
It results:

$$\Delta\tau = \frac{\sqrt{J_M}}{T^*_{L,rms}}\sqrt{\alpha-\beta} \qquad (17)$$

where $T^*_{L,rms}$ and β do not depend on the motor.

Motors in table 2 have similar accelerating factors but rotor momentum of inertia really different. Suppose their accelerating factors were higher, for a given task, than the load factor. Then motor no.230, with a greater moment of inertia, would have a wider range of useful transmission ratios than motor no.44.

6. CONCLUSIONS

The "accelerating factor", or "continuous duty power rate", it is a parameter characterizing the performance of a motor and it is defined as the ratio between the motor rated torque and the square of its rotor momentum of inertia.
The higher is the accelerating factor α, the wider is the range of transmission ratios that can be used for coupling the motor to the load to be moved.
The designer who is choosing the motor reducer unit, however, has difficulties in understanding whether the choice done is the best or not, because there are no absolute references on the accelerating factor on which perform the selection. In other words, it is impossible, at now, to evaluate if the chosen motor is the best solution for an application, or if a smaller one with the same accelerating factor, and therefore better for weight and dimensions, is available in commerce.
The analysis reveals how motors for automation field (taking into account only synchronous brushless motors) are extremely heterogeneous in terms of performance and highlights the needing to define benchmarks for the accelerating factor to help the designer in selecting the best motor-transmission coupling.

7. REFERENCES

[1] Pasch K.A.,Seering W.P. "On the drive systems for High-performance Machines", **Transactions of ASME** Vol.106, 1984, pp.102-108
[2] Van de Straete H.J, Degezelle P., de Shutter J., Belmans R., "Servo Motor Selection Criterion for Mechatronic Application", **IEEE/ASME Transaction on mechatronics** Vol.3, 1998, pp.43-50
[3] Van de Straete H.J, de Shutter J., Belmans R., "An Efficient Procedure for Checking Performance Limits in Servo Drive Selection and Optimization", **IEEE/ASME Transaction on mechatronics** Vol.4, 1998, pp.378-386
[4] Cusimano G., "A procedure for a suitable selection of laws of motion and electric drive systems under inertial loads", **Mechanism and Machine Theory** Vol.38, 2003, pp.519-533

[5] Cusimano G., "Optimization of the choice of the system electric drive-device-transmission for mechatronic applications", **Mechanism and Machine Theory** Vol.42, 2007, pp.48-65
[6] Cusimano G., "Generalization of a method for the selection of drive systems and transimissions under dynamic loads", **Mechanism and Machine Theory** Vol.40, 2005, pp.530-558
[7] Van de Straete H.J, de Shutter J., Leuven K.U. "Optimal Variable Transmission Ratio and Trajectory for an Inertial Load With Respect to Servo Motor Size", **Transaction of the ASME** Vol.121, 1999, pp.544-551
[8] Roos F., Johansson H., Wikander J. "Optimal selection of motor and gearhead in mechatronic application", **Mechatronics** Vol.16, 2006, pp.63-72
[9] Legnani G., Tiboni M., Adamini R. **Meccanica degli azionamenti**, Esculapio, Italy, 2002
[10] Giberti H., Cinquemani S., Legnani G., "Evaluation of motor-reducer coupling in high demanding industrial applications", **Proc. of 2nd International Multi-Conference on Engineering and Technological Innovation: IMETI 09**, Florida, USA, 2009.

Impact of the Data Quality from Hydroelectric Plants in the Past Operation Analysis using a Middle Term Simulation Tool

Ieda G. Hidalgo[1], Secundino Soares F.[2], Darrell G. Fontane[3], Marcelo A. Cicogna[4] and João E. G. Lopes[5]

[1] *Ph.D. Electrical Engineering, Department of Electrical and Computer Engineering, State University of Campinas (UNICAMP) – Brazil; e-mail: iedahidalgo@gmail.com*
[2] *Professor, Department of Electrical and Computer Engineering, State University of Campinas (UNICAMP) – Brazil; e-mail: dino@cose.fee.unicamp.br*
[3] *Professor, Department of Civil and Environment Engineering, Colorado State University (CSU) – EUA; e-mail: darrell.fontane@colostate.edu*
[4] *Ph.D. Electrical Engineering, Anhanguera Educacional (AESA) – Brazil; e-mail: marcelocicogna@gmail.com*
[5] *Ph.D. Civil Engineering, Consulting Engineers – Brazil; e-mail: jelopes1@gmail.com*

ABSTRACT

This paper presents the impact of the data quality from hydroelectric plants in the analysis of their past operation. A middle term simulation tool has been applied to Brazilian hydroelectric plants which are under the coordination of the Electric System National Operator (ONS). In order to analyze the impact of the data quality this simulator is used to reproduce the past operation of the plant, once with official data and average productivity and the next with adjusted data and variable overall efficiency. The results show that the use of both consolidated data and variable overall efficiency reduces the errors between recorded and simulated variables bringing the plant's simulated and real operation closer.

Keywords: Hydroelectric Plants, Data Quality, Simulation Tool, Brazilian Model, Operation Planning.

1. INTRODUCTION

The power generated by a hydroelectric plant is a function of the water discharge by the turbines, the difference between the forebay and tailrace levels, the penstock head loss and the machines' efficiency involved in the process. The level of accuracy used to represent the hydropower generation function depends on the problem addressed. For example, in the middle term planning, it is quite usual to consider an average productivity for the plant. While, in the short-term, usually a variable efficiency is considered for each machine.

This paper focuses on the middle term and its characteristics. The main objective is to compare the past operation analysis of a plant alternating: official data and consolidated data. The tool used to reproduce the past operation is a middle term simulator of the hydroelectric plants operation.

The first section presents the Brazilian model for middle term hydroelectric operation planning. The second details the concept and calculation of variable overall efficiency. The third describes the middle term simulator used in this paper. The fourth shows a graphic and numerical analysis on the impact of the data quality in the past operation reproduction. Conclusions are given in fifth section.

2. BRAZILIAN MODEL FOR MIDDLE TERM HYDROELECTRIC OPERATION PLANNING

The most common formulations of the Brazilian system for hydroelectric operation planning include: the production function and the water balance equation [12].

The goal of the production function is to quantify the power generation of a hydroelectric plant, as Eq. (1).

$$p = k.\eta.[h_{fb}(x) - h_{tr}(u) - h_{pl}].q \quad (1)$$

where:

p Is the instantaneous power obtained in the conversion process of the hydraulic potential energy to electric energy (MW).

k Is the gravity constant, multiplied by the water specific weight and divided by 10^6. Its value is 0.00981 (MW/(m³/s)/m).

Is the forebay elevation which is function of the water storage *x* (m).

u Is the water release of the plant, that is, the sum of the water discharge by the turbines and the water spillage (m³/s).

$h_{tr}(u)$ Is the tailrace elevation which is function of the water release *u* (m).

h_{pl} Is the penstock head loss which is function of the water discharge (m).

q Is the water discharge by the turbines of the powerhouse (m³/s).

The water balance equation, Eq. (2), is used to calculate the water mass conservation balance of the reservoir.

$$x = x_0 + \left(y + \sum_{j \in \Omega} u_j - (q + s + ev + uc) \right) \cdot \frac{\Delta t}{10^6} \quad (2)$$

where:

x_0 Is the reservoir volume at the beginning of period *t* (hm³).

y Is the incremental water inflow to reservoir during the period *t* (m³/s).

Ω Is the upstream plants index set from the analyzed plant.

s Is the water spillage during the period *t* (m³/s).

ev Is the reservoir evaporation during the period *t* (m³/s).

uc Is the use of the reservoir's water without the purpose of generating energy, such as: urban water supply, irrigation and navigation during the period *t* (m³/s).

Δ*t* Is the size of the period *t* (s).

In order to aid the calculation of the parameters involved in the equations presented above are used seven physical functions: area-level polynomial, level-volume polynomial, level-release polynomial, maximum power function, maximum water discharge function, efficiency function and penstock head loss function [1].

η Is the constant or variable efficiency of the plant in the conversion process of the mechanical energy to electrical energy.

x Is the water storage in the reservoir of the plant (hm³).

$h_{fb}(x)$

3. CONCEPT AND CALCULATION OF VARIABLE OVERALL EFFICIENCY

The overall efficiency includes the losses and the efficiencies involved in the operation. The use of the overall efficiency simplifies the production function, Eq. (1), without compromising the planning and the operation record of the plant.

Fig. 1 presents the simplification of the Eq. (1) using the concept of overall efficiency, η^G.

$$p = k.\eta.[h_{fb}(x) - h_{tr}(u) - h_{pl}].q$$

$$p = k.\eta^G.[h_{fb}(x) - h_{tr}(u)].q$$

Fig. 1. Production function, Eq. (1), using the concept of overall efficiency.

The variable overall efficiency is represented as a hill curve matrix. This matrix can be a function of the gross head and the power output. For the attainment of the overall efficiency matrix using the data recorded by the plant can be used an optimization method, such as the "Solver" tool in Excel [3]. The objective function is to optimize the cells of the overall efficiency matrix in order to minimize the sum of squared error between the plant's overall efficiency for the selected record and the overall efficiency calculated, as Eq. (3).

$$Min \sum_{i=1}^{n} [\eta^G(i) - \eta^G(i)_{calc}]^2 \quad (3)$$

where:

n Is the number of operations recorded in the plant's database.

i Is the index of the operation recorded in the plant's database.

$\eta^G(i)$ Is the plant's overall efficiency for the record of index *i*.

$\eta^G(i)_{calc}$ Is the plant's overall efficiency calculated using the overall efficiency matrix for the power output and the gross head of the record of index i.

4. MIDDLE TERM SIMULATOR

The middle term simulator of the hydroelectric plants operation represents in detail the active operational restrictions on this horizon using weekly or monthly data. It can be used for planning the future operation or reproducing the past operation of a period. When it is used for reproducing of the past operation it performs the function of a tool for data analysis like others mentioned in Hidalgo 2004 [5]; Hidalgo et al. 2009-A [6]; Hidalgo et al. 2009-B [7] and Hidalgo et al. 2009-C [8]. Its simulation process is based on the production function, Eq. (1), and the water balance equation, Eq. (2).

The software project and the computational implementation of this simulator use the Object-Oriented Paradigm [3], the C++ Programming Language [10] and the Structured Query Language (SQL) [4].

For studies of this paper, the simulator's aims is reproduce the water discharge trajectory from the initial volume, the trajectories of generation, water spillage and water inflow, as Fig. 2. The advantages of this type of application are: it shows the impact of the data inconsistency in the plant's water balance and it can be used for planning the future operation.

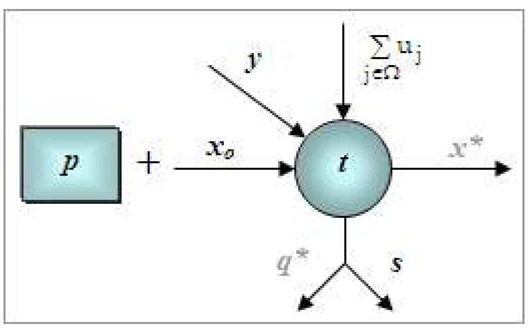


Fig. 2. In bold, input data for the simulator when its aim is to reproduce the water discharge trajectory.

5. GRAPHIC AND NUMERICAL ANALYSIS ON THE IMPACT OF THE DATA QUALITY IN THE PAST OPERATION REPRODUCTION

The middle term simulator was applied to a large Brazilian hydroelectric plant. The data recorded by the plant from 09/01/06 to 08/31/07 were compared to the resulting data from the simulation for the same period. The comparison was made in two situations. In the first the simulator worked with the official data provided by the company responsible for the operation. In the second the simulator used the consolidated data obtained according to the methodology presented in Hidalgo et al. 2009-D. Basically, the differences between the official and consolidated data are the six physical functions involved in the hydroelectric operation planning: area-level polynomial, level-volume polynomial, level-release polynomial, maximum power function, maximum water discharge function and efficiency function.

A - Simulation using official data and average productivity

In this simulation were used the official physical data of the company and the average productivity of the plant. The simulator's goal is to reproduce the water discharge trajectory recorded by the plant, as Fig. 2.

Fig. 3, 4 and 5 show the recorded and simulated trajectories. It is possible to notice that the official physical data recorded by the plant are not coherent with the reality of the operation because the simulated trajectories moved away from the recorded trajectories.

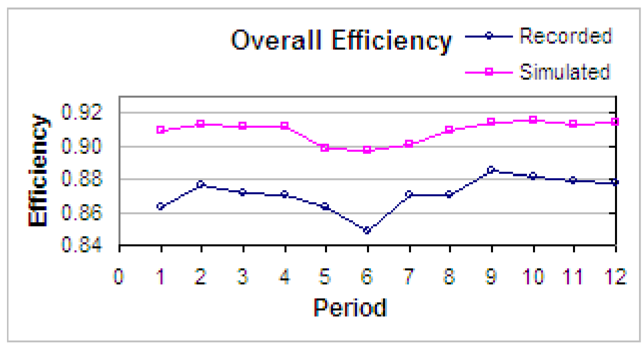


Fig. 3. Comparison between the overall efficiency trajectories.

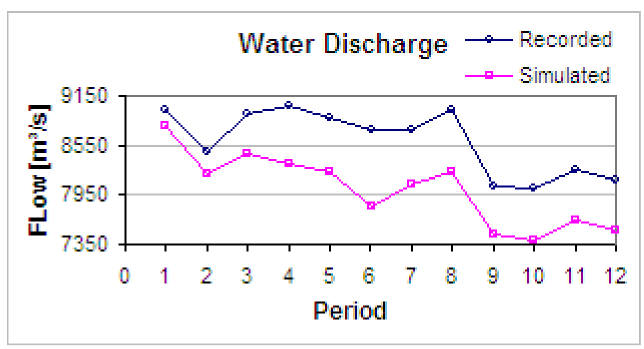


Fig. 4. Comparison between the water discharge trajectories.

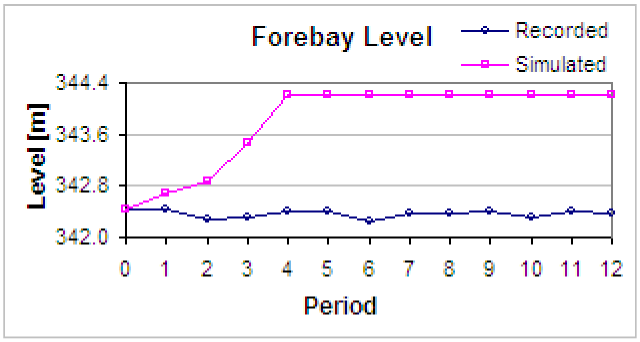


Fig. 5. Comparison between the forebay level trajectories.

The physical information that most influenced the result of this study was the plant's productivity, Fig. 3. Since it is overestimated, the simulator saves reservoir's water, Fig. 4, in order to produce the recorded generation. This justifies the increase of the forebay level, Fig. 5.

B - Simulation using consolidated data and variable overall efficiency

In this simulation were used the consolidated physical data of the plant and the variable overall efficiency calculated according to Eq. (3). Again, the simulator's goal is to reproduce the water discharge trajectory recorded by the plant, as Fig. 2.

Fig. 6, 7 and 8 show the recorded and simulated trajectories of the plant.

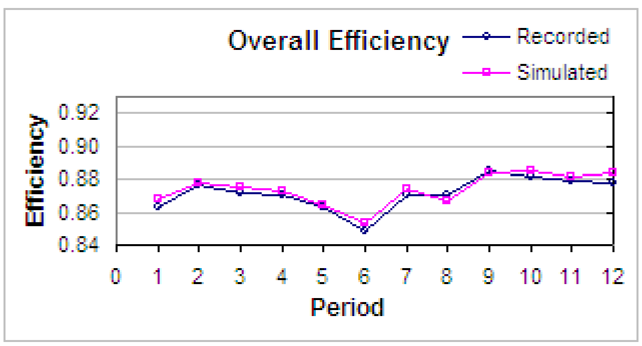


Fig. 6. Comparison between the overall efficiency trajectories.

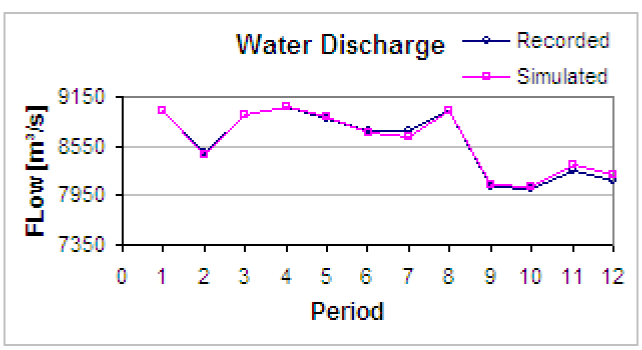


Fig. 7. Comparison between the water discharge trajectories.

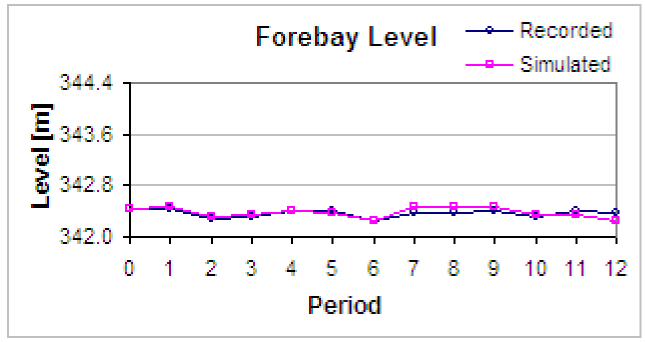


Fig. 8. Comparison between the forebay level trajectories.

These three figures show a strong coherence between the recorded data and their reproduction by simulation. This demonstrates the advantages of the use of variable overall efficiency even in the middle term.

In order to show a numerical analysis of the results presented above, Table I presents a statistical summary of the sum and mean squared error between the recorded and simulated variables when the simulator's goal was water discharge reproduction.

Table I. Statistic summary – Water discharge reproduction

	Recorded Data x Simulated Data		
Variable	**Official**	**Adjusted**	**Reduction**
	Sum squared error		
$h_{fb}(x)$	32.09	0.04	99.87 %
$h_{tr}(u)$	5.81	0.17	97.16 %
q	4,609,516.29	16,089.42	99.65 %
s	3,713,950.31	0.01	100.00%
η^G	0.02	0.00	99.13 %
	Mean squared error		
$h_{fb}(x)$	1.57	0.06	
$h_{tr}(u)$	0.70	0.12	
q	619.78	36.62	
s	556.32	0.03	
η^G	0.04	0.00	
Mean reduction of the sum squared error			99.16 %

The first column presents the analyzed variables. The second shows the sum and mean squared error between recorded data and simulated data using official information and plant's average productivity. The third shows the sum and mean squared error between recorded data and simulated data using consolidated information and plant's variable overall efficiency. The fourth column presents the error reduction between the second and third columns.

The numbers in Table I show that the data quality has great influence in the past operation analysis of a plant. Some minor differences presented in the third column of the table are explained by accuracy of 0.1 MW used by the simulator. The differences related to the tailrace level and water discharge are explained by the fact that the plant's tailrace is represented by a scattered cloud of points. It is believed that an accuracy technique for measuring of water discharge reduces the spaces among the points of the cloud further improving the results.

6. CONCLUSIONS

This paper compared the impact of the data quality in the past operation analysis of a large Brazilian hydroelectric plant. The analysis was made using a middle term simulator. The objective of this simulator was to reproduce the plant's operation from 09/01/06 to 08/31/07 using monthly data. The simulator's input data were classified in two kinds: official data using the plant's average productivity and consolidated data using the plant's variable overall efficiency.

The results were presented in the form of graphics and table. They all confirmed the importance of the data quality. For the analyzed study the mean reduction of the sum squared error between the recorded and simulated variables was of 99.16%.

The impact of the data quality in the past operation analysis indicates that others computational models used by the energy sector for optimization, simulation and streamflow forecasting may present dubious results due the quality of the data provided for them. Therefore, the search for the improvement of data is important to the choice of an economic and reliable operation policy for the hydroelectric system.

7. ACKNOWLEDGMENTS

The research reported herein was supported by the CNPq and CAPES, Brazilian Government agencies dedicated to the development of science and technology, which have funded, at different times, the Ph.D. studies of the first writer.

8. REFERENCES

[1] Arce, A., Ohishi, T., and Soares, S. (2002). "Optimal dispatch of generating units of the Itaipu hydroelectric plant." *J. IEEE Transactions on Power Delivery*, 17(1).

[2] Bloch, S. C. (2003). "Excel for engineers and scientists." 2nd Edition, *Wiley*, EUA.

[3] Cicogna, M. A. (1999). "Model for energetic operation planning of hydrothermal systems using object oriented programming." MS thesis, State University of Campinas, SP, Brazil.

[4] Elmasri, R., and Navathe, S. (2000). "Fundamentals of database systems." *Addison Wesley Professional*, MA, EUA.

[5] Hidalgo, I. G. (2004). "Search system for data analysis from hydroelectric plants." MS thesis, State University of Campinas, SP, Brazil.

[6] Hidalgo, I. G., Soares Filho, S., Fontane, D. G., and Cicogna, M. A. (2009-A). "Management and analysis of data from hydroelectric plants." In 2009 *IEEE Power Systems Conference & Exhibition (IEEE)*, Washington / Seattle – pag. 1 to 6.

[7] Hidalgo, I. G., Fontane, D. G., Soares Filho, S., and Cicogna, M. A. (2009-B). "Computer aided system for managing, controlling and analyzing data from hydroelectric plants." In 2009 *World Environmental & Water Resources Congress (ASCE)*, Missouri / Kansas – pag. 4749 to 4757.

[8] Hidalgo, I. G., Fontane, D. G., Soares Filho, S., and Cicogna, M. A. (2009-C). "A Simulator of the hydroelectric plants operation as tool for analyzing data." In 2009 *World Congress on Computer Science and Information Engineering (CSIE)* Los Angeles / Anaheim. DOI: 10.1109/CSIE 2009.1078 – pag. 14 to 18.

[9] Hidalgo, I. G., Fontane, D. G., Soares, S., Cicogna, M. A., and Lopes, J. E. G. (2009-D). "Data consolidation from hydroelectric plants." *ASCE Journal of Energy Engineering (accepted).*

[10] Hollingworth, J., Butterfield, D., Swart, B., and Allsop, J. (2001). "C++ Builder 5 developer's guide." *Sams Publishing*, EUA.

[11] ONS. "Electric System National Operator" <http://www.ons.org.br>

[12] Pereira, M. V. F., and Pinto, L. M. V. G. (1985). "Stochastic optimization of a multireservoir hydroelectric system: a decomposition approach." *J. Water Resources Research*, 21(6), 779-792.

Design of Palpation Simulator Using MagnetoRheological Fluids

Dean H. Kim
Julie A. Reyer

Department of Mechanical Engineering
Bradley University
Peoria, IL 61625

ABSTRACT

This paper describes the design and development of a prototype for a palpation simulator, i.e. to replicate the act of touching or feeling the human body to determine the condition of the underlying parts. This project work is motivated by the need for more realistic and more versatile palpation simulators. The final design utilizes magnetorheological (MR) fluids, whose properties can be changed with magnets to represent the desired abnormality to be palpated. Key design decisions include choice of MR fluid, magnetic field strength, and packaging of the MR fluid. The improved realism and increased number of scenarios for the prototype has been confirmed by local experts. This work has been performed by an undergraduate team for the two-semester capstone design course for the Mechanical Engineering Department at Bradley University. This project crosses traditional mechanical engineering boundaries because it addresses a critical training need within the medical profession. A significant and unique feature of this project is its sponsorship by the Kern Entrepreneurship Education Network (KEEN). KEEN's mission is to graduate engineers who have the necessary mindset regarding entrepreneurship and innovation. Therefore, this project work includes extensive research and market analysis regarding the designed system.

Keywords: Medical Simulation, Palpation, Design, Magnetorheological Fluids.

INTRODUCTION

The use of simulators to train medical professionals in important skills has become more common due to factors such as limited patient availability and the increasing number of scenarios for which training is needed. Medical simulation is intended to improve education and patient safety by replacing real patient experiences with guided experiences under realistic conditions [1],[2]. One such important skill is palpation, which is the act of touching or feeling the human body to determine the condition of the underlying parts. This palpation skill is needed for situations such as feeling for a breast lump, checking for a swollen thyroid, and determining the existence of internal hemorrhage after trauma. The possible variables involved during the medical professional's palpation include the number of fingers used and the applied pressure.

The presently available palpation simulators are very limited in the important areas of realism and versatility of possible scenarios. For example, one popular breast palpation simulator provides only several possible lump locations and several possible lump sizes, which severely limits the effectiveness of the training.

This paper describes the development of a palpation simulator using magnetorheological (MR) fluids to represent the abnormal anatomy which must be detected. MR fluids are a type of smart fluid whose properties such as viscosity can be changed when subjected to an applied magnetic field, due to the extremely small magnetic particles suspended in the fluid. Design decisions for this palpation simulator include choice of MR fluid, required magnetic field strength, type of magnet, and packaging of the MR fluid. The improved realism and increased number of possible scenarios for the functional prototype has been confirmed by local experts in the related fields, including practicing physicians and the director of the Clinical Skills Laboratory at the local medical school.

Other design options were considered for this palpation simulator. One rejected option was the use of mechanical, pin-sized actuators to generate the desired palpation surface. However, the achievable resolution of the palpation surface was not fine enough. Another rejected option was the use of electrorheological (ER) fluids, whose properties change when subjected to an electrical current. However, there was great concern about the safety for the medical trainees due to the continuous current flowing through the fluid.

This work has been performed by a team consisting of four undergraduate students and two faculty advisors as a two-semester capstone design project. It incorporates many aspects of a mechanical engineer's education such as mechanical design and material selection. This project also crosses traditional mechanical engineering boundaries because it addresses a critical training need within the medical profession. Hence, extensive feedback has been obtained from local experts in the related fields, including practicing physicians and the director of the Clinical Skills Laboratory at the local medical school. Previous projects have been completed and documented under the same senior project sequence. These projects each had focused on the development of a control system, for a laboratory wind tunnel [3] and a gear dynamometer [4], respectively.

Finally, a significant and unique feature of this project is its sponsorship by the Kern Entrepreneurship Education Network (KEEN). KEEN's mission is to graduate engineers who have the necessary mindset regarding entrepreneurship and innovation. Therefore, this project work includes extensive research and market analysis regarding the palpation simulator.

The remainder of the paper is organized as follows. The KEEN sponsorship of this project is described briefly, and the guidelines for the senior design capstone course for the Mechanical Engineering Department at Bradley University are provided. The market analysis for medical simulators is presented as stipulated by the KEEN sponsorship. The technical approach is described, in particular the various design concepts and options to be studied. The final design is presented. Finally, the construction of the proof-of-concept prototype is described, and the prototype validation with medical professionals is confirmed.

KERN ENTREPRENEURSHIP EDUCATION NETWORK

The Kern Entrepreneurship Education Network (KEEN), created by the Kern Family Foundation in 2005, is intended to help universities graduate engineers who are equipped with an action-oriented entrepreneurial mindset who will contribute to business success and transform the U.S. workforce. KEEN has provided the funds to sponsor this two-semester senior design project for the Department of Mechanical Engineering at Bradley University. This project incorporates many aspects of a mechanical engineer's education such as system design, material science, and design of experiments. This project also crosses traditional mechanical engineering boundaries because it addresses a critical training need within the medical profession. Hence, this project matches the KEEN mission by including extensive research and market analysis regarding the designed system.

MECHANICAL ENGINEERING SENIOR DESIGN CAPSTONE COURSE AT BRADLEY UNIVERSITY

The senior capstone design project course for the Mechanical Engineering Department at Bradley University covers two semesters, specifically the Fall and Spring semesters of the academic year. Each project has an industrial sponsor (i.e. "client") and the typical project team consists of four undergraduate students and at least one faculty supervisor. For this project, the client (i.e. KEEN) is represented by two additional Bradley faculty members, including the person directly responsible for securing the KEEN sponsorship. This design project course requires that each student team provide an initial proposal, regular progress reports, four oral presentations to classmates and faculty, and at least one presentation to the client. Each student team's initial proposal includes a schedule with milestones, a budget, and a list of deliverables.

For this project, the expected deliverables are:

- Midterm report
  - Includes research and analysis for at least 3 potential designs
  - Recommendations for design
- Functional proof-of-concept prototype
- Final report
- CD containing all project documentation

MARKET ANALYSIS

The team first has conducted extensive research on the current market for medical simulators, finding information about the quality, cost, and functionality of available products. This is a necessary step to show the team what kind of market there is for the proposed palpation simulator. There are full-body simulators available which can simulate certain human functions such as breathing (via embedded air pump) and urination. However, these simulators do not feel realistic because they use rubber to mimic human skin. These full-body simulators, shown in Figure 1, typically cost at least \$ 200,000. There also are partial simulators available which provide training for specific tasks and/or for particular sections of the body. These include breast exam simulators and generic palpation simulators, shown in Figure 2 and Figure 3, respectively. However, these simulators only have a limited number of possible abnormality shapes and locations, so their effectiveness decreases with repetition. This project work described in this paper culminates in a proof-of-concept prototype which bridges the gap between the expensive full-body simulators and the less expensive but limited capability partial simulators.

This market research also has included a visit to the Rager Clinical Skills Laboratory (RCSL) in downtown Peoria, which uses a full-body simulator in addition to some partial simulators. The team's communication with the lab director had provided a better understanding of the customer needs, as did similar communication with students from the nursing school.

TECHNICAL APPROACH

The team has made the following assumptions in order to simplify the design of the palpation simulator. The finger size of the student performing palpation would be considered to be constant. At the same time, the soft tissue and abnormality would be individually treated as homogeneous substances for the solution chosen, i.e. the density and elasticity of each would remain constant throughout the substance.

Two unique design concepts have been researched and analyzed in depth in order to determine the recommended solution. The first type requires the student to palpate a simulator, or place their finger onto the device. The simulator is designed to deliver realistic feedback by recreating the resistive force of soft tissue and an abnormality. One of the solutions for this type of simulator is using pins that would be individually controlled by servomotors. By doing this, an abnormality could be generated by controlling the height of each pin. An advantage to this design is that servomotors are easily accessible and relatively inexpensive. However, one concern is the amount of space that this device would occupy. Another solution is that of electrorheological (ER) fluids. These fluids are a special type of fluid that, when exposed to electrons, experience a change in viscosity. The fluids can be arranged into individual cells on a grid, allowing for the location of the abnormality to be chosen by sending electric current sent to any particular cell. While the space constraint was not an issue of concern for this setup, the cost could potentially have been. One liter of ER fluid, the minimum order, can cost around \$1000-\$1500. Other fluid alternatives that are less expensive were researched, which in turn led to the option of magnetorheological (MR) fluids. MR fluids are similar to ER fluids, but the change in viscosity is driven by an applied magnetic field as opposed to an electric one. Small iron filings in the fluid line up along the lines of flux, which causes the increase in viscosity.

The second design concept uses haptic feedback to provide a virtual simulation via a device placed on the student's finger. The concept behind this design is simulating the presence of an entity that does not actually exist. In this manner, the device is similar to commercially available haptic devices Ideally, the device would consist of a finger sleeve lined with sensors and connected to a computer interface, through which the user could program various cases and locations of abnormalities.

FINAL DESIGN

After extensive study of the various options and design concepts, the team has decided to pursue Design Concept 1, i.e. placing finger(s) onto the device. The team also has decided to use MR Fluids to create the abnormality to be palpated. Hence, the three categories of design decisions to be made are the MR fluid, the magnet, and the fluid packaging. The final design will be the proof-of-concept prototype of the palpation simulator capable of producing an abnormality with a modulus of elasticity comparable to that found in human tumors.

MR fluids are a type of smart fluid in which the viscosity can be changed by applying varying magnetic fields to the fluid. Generally, these fluids are used in a shear mode, where displacement is occurring perpendicular to the magnet's flux lines often found as a damper, a brake, or a clutch. However, this project uses the MR fluid in "squeeze mode", where the compressive force is in the same direction as the flux lines coming off the pole of the electromagnet. The team has chosen a water-based MR fluid because it provides a modulus of elasticity of 300 kPa, which is about the same value as a ductile carcinoma (breast tumor)[5].

The magnet is the driving force of the stiffness that the MR fluid can attain. The MR fluids become saturated when electromagnetic field strength is between 1.2 and 2.0 Tesla, at which point the MR fluids begin to lose their fluid properties and become very hard. The team has chosen to use a permanent magnet to activate the MR fluids. In particular, the prototype employs a 0.58 Tesla permanent magnet underneath MR fluid contained in LDPE packaging, separated only by a thin aluminum sheet. Magnetic modeling has been performed to model the flux lines and flux density for this design, which are shown in Figure 4 and Figure 5, respectively.

If a different stiffness is desired of the abnormality, various strength magnets could be interchanged to allow for this change. The team also has considered using electromagnets, but the achievable electromagnetic field strength is not strong enough to cause a noticeable change in MR fluid viscosity.

For the packaging of the MR fluid, the team has decided that a bladder consisting of compartments, connected by constricted passages, would allow MR fluid to flow freely in the absence of a magnetic field while providing the realistic feel of the abnormality when a magnetic field is applied to the appropriate constricted passages. The chosen packaging pattern is shown in Figure 6.

CONSTRUCTION OF PROTOTYPE

The frame for the prototype is made of 80/20 aluminum. Two layers of aluminum support the fluid packaging. The bottom layer is a ¼" thick aluminum plate providing support for the fluid packaging layer. In this sheet, a pattern matching the fluid packaging cells Is drilled to allow for magnet placement. This pattern consisted of approx. 40 holes (1 1/8" in diameter) to allow adequate clearance for the 1" diameter magnets. Figure 7 shows the machining of these holes. The top layer is a thin sheet of aluminum (0.030" thick) to provide separation of the magnet and the fluid packing. Both of these layers are screwed to a support rail on the frame with the supplied screws.

The material chosen for packaging the MR fluid is a Low-Density Polyethylene (LDPE) matrix of approximately 40 cells. The layout of the cells utilizes a bottleneck design that acted as a nozzle to restrict the flow. This enhances the overall performance of the fluid. A large syringe is used to fill these cells. The final array of MR Cells is placed on top of the aluminum supports. Figure 8 shows the filling of the packaging with MR fluid.

The tissue layer is a piece of open celled polyurethane foam to represent soft tissue. This was simply cut to size, 12"x12" to fit inside the frame and placed on top of the fluid cells. The final cover or "skin" layer is a sheet of nylon fabric that was glued to the foam piece using high strength spray adhesive. This combination was chosen due to its availability and low cost. Figure 9 shows the packaged MR fluid and the foam tissue layer.

VALIDATION OF PROTOTYPE

This project crosses traditional mechanical engineering boundaries because it addresses a critical training need within the medical profession. Medical professionals have been consulted throughout this project. The student team has given their final presentation to these two medical professionals, in addition to a representative from the Kern Entrepreneurship Education Network (KEEN). The medical professionals have tested the final proof-of-concept prototype as shown in Figure 10. The medical professionals have complimented the prototype, and they even have suggested additional potential applications for this technology. Most importantly, they have agreed that the proposed technology could be a viable and applicable method of medical simulation.

CONCLUSIONS

This paper describes the development of a palpation simulator using magnetorheological (MR) fluids to represent the abnormal anatomy which must be detected. Design decisions for this palpation simulator have included choice of MR fluid, required magnetic field strength, type of magnet, and packaging of the MR fluid. The proof-of-concept prototype has been constructed and its potential viability for palpation simulation has been validated by local medical experts. This work has been performed as a two-semester capstone design project for the Department of Mechanical Engineering at Bradley University. In addition to incorporating typical mechanical engineering areas, this project also crosses traditional boundaries because it addresses a critical training need within the medical profession. A significant and unique feature of this project is its sponsorship by the Kern Entrepreneurship Education Network (KEEN). The extensive research and market analysis required by the project has fulfilled KEEN's mission to graduate engineers with the necessary mindset regarding entrepreneurship and innovation.

ACKNOWLEDGEMENTS

This project has been fully supported by KEEN. The authors wish to thank Dr. Bob Podlasek, Dr. John Engdahl, Dr. Andy Chiou, Dr. David Buchanan, and Dr. Danuta Dynda. The authors especially wish to thank the student team members: Mr. Tim Myers, Mr. Kurt Friedrich, Mr. Carl Poettker, and Mr. Nate Adams.

REFERENCES

[1] H.R. Champion, and A.G. Gallagher, "Surgical Simulation – a Good Idea Whose Time Has Come," British Journal of Surgery, Vol. 90, Issue 7, pp. 767-768.

[2] J. Balcombe, "Medical Training Using Simulation: Towards Fewer Animals and Safer Patients," Proceedings of the Fourth World Congress of ATLA, 2004, pp. 553-560.

[3] D. Kim, M. Morris, G. Leja, T. Scarlata, and S. Wylie, "Dynamic Analysis and Control System Development for a Laboratory Wind Tunnel," Proceedings of the 2000 ASME International Mechanical Engineering Conference and Exposition, November 2000, pp. 71-76.

[4] D. Kim, M. Morris, A.M. Moellenberndt, T. Rowlands, and T. Masha, "Design of Controller and Data Acquisition System for a Gear Dynamometer," Proceedings of the 2002 ASME International Mechanical Engineering Conference and Exposition, November 2002, pp. 51-56.

[5] J.O. Ladeji-Osias, and N.A. Langrana, "Analytical Evaluation of Tumors Surrounded by Soft Tissue," Proceedings of the 22nd Annual EMBS International Conference , July 2000, pp. 2114-2117.

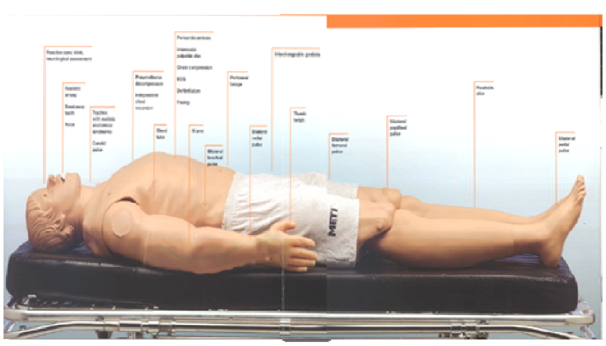


Fig. 1 Commercially Available Full Body Human Simulator

Fig. 2 Commercially Available Breast Exam Simulator

Fig. 3 Commercially Available Palpation Simulator

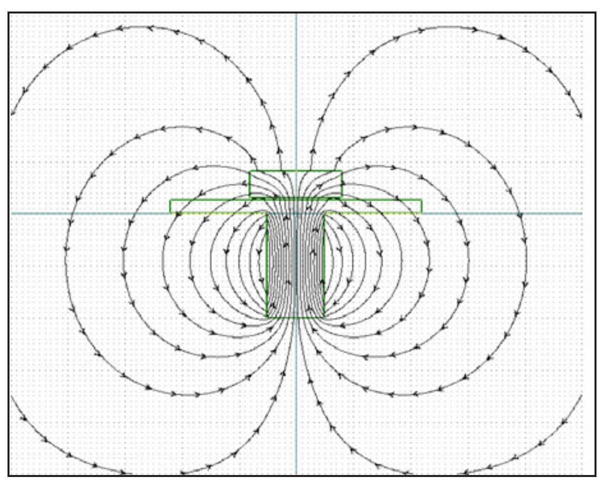


Fig. 4 Magnetic Flux Lines for Final Design

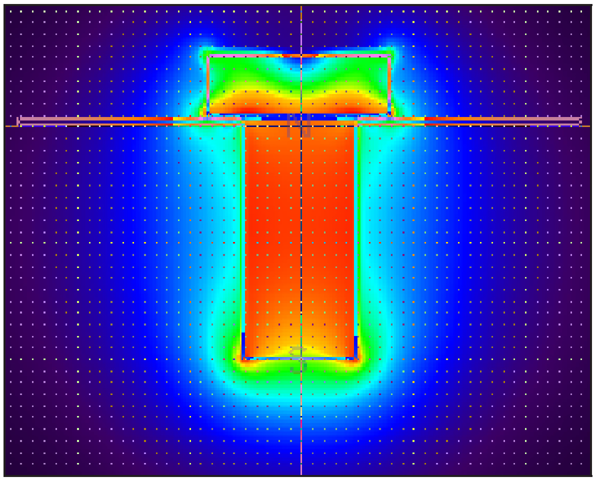


Fig. 5 Magnetic Flux Density for Final Design

Fig. 6 Packaging Pattern for MR Fluid

Fig. 7 Machining Magnet Placement Holes in Plate

Fig. 8 Filling Packaging with MR Fluid

Fig. 9 Prototype with Packaged MR Fluid and Foam Layer

Fig. 10 Testing of Prototype by Medical Professionals

"DESIGN AND IMPLEMENTATION OF A METHODOLOGY FOR THE ESTABLISHMENT OF THE MEXICAN CUSTOMER SATISFACTION INDEX FOR SOCIAL PROGRAMS: The Case of the Subsidized Milk Program"

Odette Lobato Calleros
Universidad Iberoamericana-Mexico City, Department of Engineering, Prolongación Paseo de la Reforma No. 880, Lomas de Santa Fe, Distrito Federal, C.P. 01710, México, Tel: (52-55) 59504000 ext. 4133; Fax: (52-55) 91774450. Email: odette.lobato@uia.mx

Humberto Rivera; Hugo Serrato; Frederick Golden; Crhistian León; Ma. Elena Gómez; Paola Cervantes; Adriana Acevedo
Universidad Iberoamericana-Mexico City.

Ignacio Méndez Ramírez
Universidad Nacional Autónoma de México, Mexico City.

ABSTRACT

This paper presents the methodology that was used to assess satisfaction among users of Social Programs in Mexico and the results obtained from the pilot test. The causal model, the Mexican User Satisfaction Index (IMSU), is an adaptation to Mexico of American Customer Satisfaction Index (ACSI). The methodology in development is intended to become an alternative for the evaluation and improvement of Social Programs and Social Policies in Mexico.

Keywords: Customer Satisfaction, Social Policies, Government Services, Social Programs, National Satisfaction Index and Quality.

1. INTRODUCTION

The work reported in this paper is part of the research project: Mexican Customer Satisfaction Index (IMSU for its name in Spanish), developed by the Universidad Iberoamericana Mexico City. Its objective is to design and implement a methodology for a National Index of Satisfaction of Beneficiaries of Social Programs in Mexico. The aim is to arrive at a standardized, comparable and reproducible methodology, based on that of the ACSI but adapted to Mexico´s conditions. This paper focuses on the results of the pilot test of the model and its questionnaires.

Mexico´s history is characterized by high levels of poverty and persistent inequality. According to Székely (2005), in 1950 88.4% of the population lived below the poverty line, and today 47% of the population (CONEVAL, 2007) still remains under this classification. For this reason the Government is obliged to generate a Social Policy that might strengthen social protection for the poorest. One of the mechanisms to implement this Policy is the creation of Social Programs.

In this sense the evaluation of Social Programs is very important. As mentioned in the study of Talukdar et al. (2005), the World Bank recommends that we hear what the users of these programs have to say about the goods and services they are receiving. The World Bank seeks to incorporate the "voice of the consumer" -consumers are mainly the poor- in its socioeconomic development projects. "Thus, the World Bank's goal to represent the voice of the poor - the" target market "- in its funding for the provision of public goods and services, is analogous to a customer-driven company's goal to incorporate the voice of customers in its marketing of private goods." (Talukdar, et al. 2005: 101). In the case of Public Policy, Hastak et al (2001:172) mention that "outcomes of a Policy must be evaluated in order to determine if the Policy is meeting its objectives. Evaluation must also provide feedback so that the Policy may be modified to improve its effectiveness". This is exactly what we intend to do by measuring beneficiary satisfaction with Social Programs in Mexico. This research project seeks to identify opportunities for improvement, which may translate into changes in the operation of the programs towards a higher quality of the services offered by the Government to the poorest. In addition, this project seeks to contribute to the field of Engineering in Mexico by developing a theoretical model and a methodology to assess the performance of processes and their relative importance from the user perspective; thus providing Engineers with a tool to improve processes.

2. BACKGROUND

Social Programs

The "Plan Nacional de Desarrollo 2007-2012" (National Development Plan), establishes a series of Social commitments and Social Policies for the government. The way Public Administration carries out the Social Policies is through Social Programs. Social Programs, like Gómez (2004) defines them are "technical services related to specifically identified human needs, that tend to provide care to those groups which, because of their circumstances or conditions, are in a situation of need or marginalization" (Gómez, 2004: 31). Such groups are composed mainly of people living in poverty. All programs being evaluated in this project are in charge of the Agency for Social Development (SEDESOL for its name in Spanish). Like SEDESOL says, the primary mission of its Social Programs is "to create equal opportunities for all people, regardless of their place of birth, income, family or socio-cultural conditions, so that every citizen may have access to goods and services essential to his development" (SEDESOL, 2009).

National Satisfaction Index Models

At the international level, there is a trend to establish National Satisfaction Indexes. In 1989, the first satisfaction index, called the Swedish Customer Satisfaction Barometer (SCSB) was born (Fornell, 1992). "The American Customer Satisfaction Index (ACSI) was introduced in 1994; the Norwegian Customer Satisfaction Barometer was introduced in 1996; and the most recent development is the European Customer Satisfaction Index (ECSI)" (Johnson et al., 2001). Other countries are also developing national satisfaction indices including New Zealand, Austria, Korea, Germany, Taiwan and Hong Kong.

American Customer Satisfaction Index

The American Customer Satisfaction Index (ACSI) is a national indicator that measures the level of satisfaction among Americans about the quality of goods and services consumed. The ACSI evaluates ten sectors of the economy of the United States, covers 41 industries and more than 200 companies and federal or local government services. The satisfaction index is obtained from the treatment of the American people responses to a telephone questionnaire. The ACSI model for government has been useful to describe government programs and services in the United States and has also been successfully tested in Mexico in previous studies such as the Diconsa Rural Supply Program and the Local Development (Microrregiones) Program (Lobato et al., 2006a; Lobato et al., 2006b). Therefore the ACSI was selected to be tested and adapted to the Mexican reality in order to create a National Index for Social Programs in Mexico, the IMSU. This model is the one that has been used in the seven Social Programs that have been evaluated during this project.

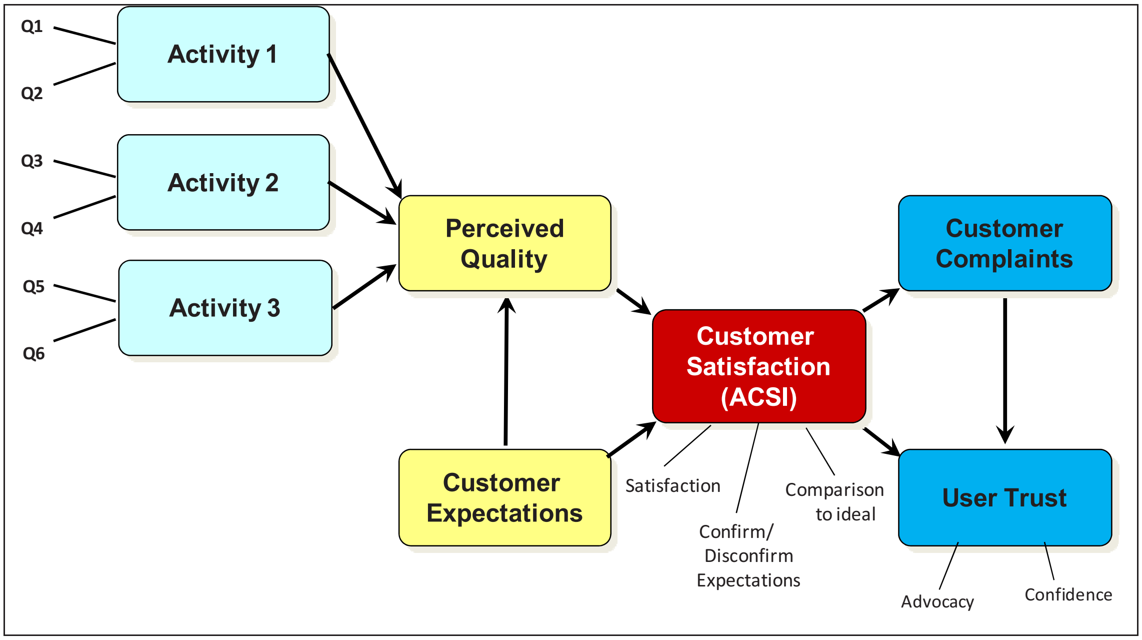


Figure 1. ACSI Model: Government Services and Nonprofit Organizations. Source: ACSI "Methodology Report" (2005)

3. METHODOLOGY

The IMSU board is an interdisciplinary group made up of Engineers in Quality, Statisticians and Social Scientists. Through a series of pilot test, the board has adapted the ACSI model to evaluate seven Social Programs in Mexico. We will briefly present some definitions and basic properties of the ACSI model.

ACSI Model

ACSI uses an econometric model that measures several indicators that conform an index of satisfaction as well as other indicators related to latent variables or constructs. The opinions of customers are collected through a survey and the data analyzed to obtain a description model by the method of partial least squares (PLS). The latter is an iterative procedure that integrates aspects of principal component analysis with multiple regression. What we measure are the manifest variables (survey questions), and through the model we find the value of the latent variables (including satisfaction), because the ACSI model is a system of cause-and-effect relationships (see Figure 1).

In the ACSI government model, Customer satisfaction has two antecedents: "Perceived Quality" and "Customer Expectations". The construct Perceived Quality has entries that must be determined for each case study and correspond to those Program Processes where the user has direct contact with the operation (left side components of the model). The consequences of satisfaction, according to the model, are: "Customer Complaints" and "User Trust". Satisfaction itself is a latent variable (central box in Figure 1 model), measured through multiple manifest variables, which are the questions that make up the satisfaction survey. The index produces results on a 0-100 scale. One of the primary objectives is to estimate the effect of ACSI on loyalty and trust of the user, a construct of universal importance in the assessment of current and future performance of the business (for further details on the ACSI model see "Methodology Report" ACSI, 2005).

One of the main advantages of this model is that it not only estimates customer satisfaction, but it also identifies the impact of each process experienced by the customer in the perceived quality. I.e. it presents an analysis of processes, where we can identify which of them need improvement and which would yield a better result, thus allowing a targeted investment.

The Federal Programs evaluated during the pilot runs included:

1. Subsidized Powdered Milk Program (Programa de Abasto Social de leche, modalidad leche en polvo). Subjects: bearers of the program's ID cards.
2. Subsidized Fluid Milk Program. Subjects: bearers of the program's ID cards.
3. Daycare Centers Program – support for parents Subjects: working mothers and single fathers receiving the benefit.
4. Daycare Centers Program – support for owners. Subjects: owners of the centers who received support for opening or remodeling the centers.
5. Senior citizens older than 70 – cash transfers. Subjects: Senior citizens older than 70.
6. Municipal infrastructure in prioritary areas. Subjects: Municipal officers in charge of the infrastructure projects.
7. Concrete Floor Program. Subjects: Residents of homes granted a concrete floor.

A causal user satisfaction model was developed for each program. In the paragraphs below the general procedure for the design and pilot testing of the seven satisfaction models is described.

- Qualitative study. Analysis of existing information about each program and of their operation rules for a preliminary identification of key processes and main users. Extensive group interviews with the programs' administrators were a major component of this step.

- Field trips. Also a part of the qualitative study, these trips consisted mainly of observation of the delivery of the benefits to the population, in-depth interviews with the beneficiaries (to know what processes are key to them and to know their lexical uses) and in-depth interviews with the programs' local administrators.

- Design of the causal model. The processes that are more likely to drive user satisfaction were identified and grouped in no more than four dimensions.

- Questionnaire design. The instruments are comprised of a set of homogeneous items for user expectations, perceived quality and user satisfaction, as well as items measuring the drivers of satisfaction, which are different according to program characteristics.

- Pilot test. Questionnaires were tested in a small convenience sample with similarities to the population of each program. The pilot run pretended to test the causal model, the logistics of fieldwork and the interviewing procedure. Possible sources of variation in the responses were sought.

- The ACSI software was used to estimate the models: satisfaction indexes and significant relations. The results up to this point are reported in the rest of this paper.

The next steps of this research project (under way while this paper is being written) are:

- Conducting interviews on a national scale for each of the seven programs.

- Estimation of the final causal models.

- Analysis and interpretation of results. Identification of improvement opportunities.

4. RESULTS

We will use the Subsidized Powder Milk Program to illustrate the results obtained from the pilot tests. The evaluation model is shown in Figure 2. Our first observation is that the main causal relations proposed by the ACSI model are confirmed for this Program: "Perceived Quality" has a significant impact (2.599) on "Customer Satisfaction", and "Customer Satisfaction" also has a significant impact (1.873) on "Trust." The program obtained a satisfaction rating of 79 on a scale of 0-100, with a margin of error of +/- 2.5, at a confidence level of 95%. The Cronbach alpha obtained for the instrument was 0.854, resulting acceptable.

The four components affecting perceived quality in this Program are: Program Access, Product, Point of Sales and Customer Service. Within each component, a group of “manifest variables” were identified. The component with more impact into perceived quality was "Customer Service" (2.691), which received a score of 93. This is where improvements could be implemented as a priority. Within this component, the variable that got the lowest rating was "Impartiality in the delivery" with a 77 rating. Here, users have the perception that some people receive more milk than their fair share, so this might be the variable where we want to start with the improvements. The second most impacting component was "Point of Sale", which received a score of 86, and within this variable the lowest score comes from "Supply" with a 71.9 rating, being another susceptible area for improvement. "Program Access” has the third most significant impact on perceived quality, and obtained a score of 72. The manifest variable with the lowest rating within this component -and within the whole model- was "Response time" (70.8). However, improvements should not start here, because the impact of this component over quality is not as strong as the impact coming from “Customer Service” (where improvements should start, as mentioned above). The “Product” component has not a significant impact over perceived quality, which might confirm the trend that the service is essential to satisfy people, even when you are providing mainly a product.

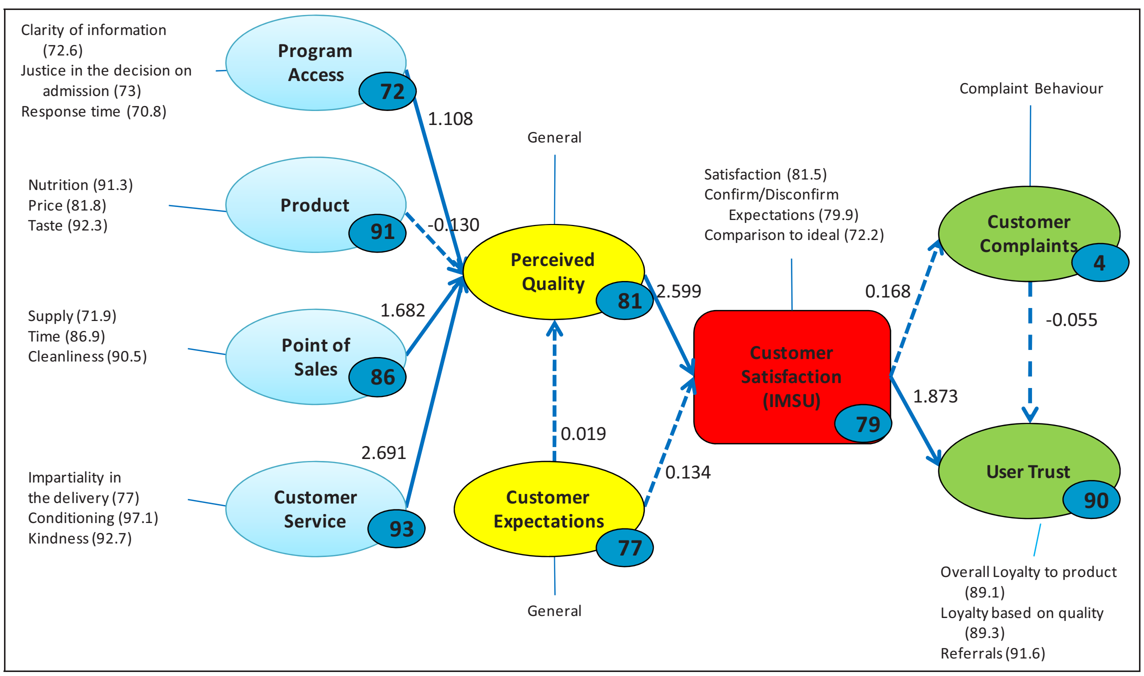


Figure 2: Evaluation Model for the Subsidized Powder Milk Program.

5. CONCLUSIONS

The evaluation model of customer satisfaction used in this paper, based on the ACSI model, is presented as a good alternative for the evaluation of the mechanisms that effect to the government's Social Policy; in this case social programs remain a way for citizens to give their opinion on whether the public policies implemented by their governments are fulfilling the objectives for which they were created and meet the needs of the population.

The evaluation model presented allows users to give feedback to the studied processes. This could lead to specific changes in the variables where improvement needs where identified, resulting in an increasing level of satisfaction among beneficiaries for these Programs.

The results presented here correspond to those obtained in the pilot test. We are currently working on the national assessment, which will render the final results of this project.

6. ACKNOWLEDGEMENTS

This research project is funded by the National Council for Science and Technology, Mexico (CONACYT).
We thank the collaboration of IMSU board members: Dr. Ignacio Méndez (UNAM), Dr. Alexander von Eye (Michigan State University), Dr. Graciela Gonzalez (CIMAT), Act. Alfredo Ramirez (CIDE), and Graciela Teruel (Universidad Iberoamericana).

7. REFERENCES

- American Customer Satisfaction Index (ACSI) (2005). “Methodology Report”. The Regents of the University of Michigan. April, USA.
- Consejo Nacional de Evaluación de la Política Social. Informe Ejecutivo de Pobreza. México, 2007. (Retrieved Mayo de 2010). From: www.coneval.gob.mx/contenido/info_public/1778.pdf
- Fornell, Claes (1992). “A National Customer Satisfaction Barometer: The Swedish Experience”. Journal of Marketing, Vol. 56, January, N°1.
- Gómez Serra, Miquel (2004). “Evaluación de los servicios sociales”. Barcelona; Gedisa.
- Hastak, Manoj; Mazis, Michael; Morris, Louis (2001). “The Role of Consumer Surveys in Public Policy Decision Making”. Journal of Public Policy & Marketing. Vol. 20, USA, 170 – 185.
- Johnson, Michael D., Gustafsson, Anders; Andreassen, Tor, Wallin; Lervik, Line; Cha, Jaesung (2001). "The Evolution and Future of National Customer Satisfaction Index Models." Journal of Economic Psychology, Nº 22.
- Lobato, Odette; Serrato, Hugo y Rivera, Humberto. (2006a) “Versión Final del Reporte de la Aplicación de la Metodología para la Obtención del Índice de Satisfacción del Beneficiario del Programa de Desarrollo Local”. México D.F.
- Lobato, Odette; Serrato, Hugo y Rivera, Humberto. (2006b) “Versión Final del Reporte de la Aplicación de la Metodología para la Obtención del Índice de Satisfacción del Beneficiario del Programa de Abasto Social Diconsa”. México D.F.
- Secretaría de Desarrollo Social. (SEDESOL). “Objetivos Estratégicos de Desarrollo Social 2007-2012”. [Retrieved Noviembre de 2009]. From: http://www.sedesol.gob.mx/archivos/1/file/Objetivos_Estrategicos_desarrollo_social.pdf
- Székely, Miguel (2005). “Desmitificación y nuevos mitos sobre la pobreza: escuchando lo que dicen los pobres". Secretaría de Desarrollo Social; CIESAS; ANUIES. México, D.F.
- Talukdar, Debabrata; Gulyani, Sumila; Salmen, Lawrence F. (2005). “Customer Orientation in the Context of Development Projects: Insights from the World Bank”. Journal of Public Policy & Marketing. Vol. 24, 100 – 111.
- Vicencio Aguilar, Anabel (2010). “El Estado del Arte de la Evaluación de la Satisfacción de Beneficiarios de Programas Sociales y su Aplicación en el Desarrollo de un Modelo de Evaluación del Programa de Abasto Social de Leche Liconsa”. Tesis para optar el grado de maestría en Ingeniería de Calidad de la Universidad Iberoamericana.

The Challenge Between Traditional And Environmental Aspects Against Modern Architectural Design, a Case Study.

Mahmoud Tarek,M.Hammad.
Al Azhar University, Faculty of Engineering, Department of Architecture.
Nasr City, Cairo, Egypt.

ABSTRACT

Modern designed buildings in the mamiluk and Fatimid part of Cairo represent an odd vision for the mere observer and the relevant visitor .This part of Cairo displays an array of historical cultures exhibiting the sequence of very outstanding and important episodes in the Islamic history of Egypt. The Children's Cancer Hospital near the River Nile at Cairo is a world- class pediatric facility dedicated to the care of children with cancer .While in the heart of Islamic Cairo it has its outlines relying more to reform than to tradition. The designer inspired the boats in the river for his structure. How this design has served the purpose and functions of the building and the compatibility to the Islamic culture of the environment, is the aim of this paper. The tools used by the author include an analysis of the plans from ecological and site environmental point of views. A comparison study was made for other new buildings in the area namely, Dar El Eftaa and Mashiakhet Al Azhar where some Islamic traditional considerations were adopted. The challenge is how these traditional outlines can best fit the function of the building and the facilities intended to be offered to the patients from one side, and how the new building has satisfied the requirements of the traditional environment.

The hospital and the two other establishments were analyzed along, data collection including plans and elevations, elements of design and architectural treatments that achieve ventilation and thermal balance. Other elements such as esthetical considerations, facades, and ornaments were also analyzed.

The results show increasing usage of glass in the hospital which does not offer sufficient heat insulation or reasonable lighting. Internal spaces were not efficiently utilized. The site and the available area are quite limited, wasting of spaces therefore, is critical. The design of this building relied on its position near the river Nile in an area of Islamic traditions which gives an odd appearance. The two other establishments although of completely different functions, their design considered and recognized the environmental requirements, and the historical background

Keywords:
Architectural design , historical landscape, Historical cultures, traditional outlines and environmental requirement

1.INTRODUCTION

Architecture history in Egypt is displaying an array of civilizations and cultures of the people in this region. It demonstrates ages of great supremacy and steep declines of dynasties that ruled through a long history beginning from ancient Egyptians to modern times. The people of Egypt adopted Islam early; only 28 years later after the descending of Islam on prophet Mohamed .Since then Islamic architecture in general and especially in Egypt, had reacted with and modified previous famous architecture styles of Greek, Byzantine, Persian and Roman styles. Islamic influence on the other hand was a major contributing factor to architecture development in Andalusia, and whole Europe during renaissance age. Islamic architecture in essence encompassed a wide range of both secular and religious styles

Cairo as the capital of Islamic Egypt, is displaying the dynasties of a long Islamic history since its establishment. Several quarters are distinguished with its dynasty style. The conservation of this heritage in such historical town is therefore of vital importance not only for Egyptians but also for humanity in general .Although this principle is violated in many ages in several districts; the survival of many structures are still a human wealth.

The foundation of Children Oncology Hospital and many other institutions is just a violation of the principle of preservation of human heritage and its environment. This study is comparative and analytical in its purpose to show degrees of contradictions to the dominant Islamic styles in the region**.**

2.ENVIRONMENTAL BACKKGROUND

The origin of old Cairo City, the core and skirts are a function of the reaction of several elements including environmental, religious, economic, social and political. Of the most important aspects that specialize and characterize the region is the Islamic culture with its social and traditional characters. Islam religion beside its nature as a relation between God and the individual is forming a way of life with its special respect of neighbours and the conservative attitudes of families. These characteristics are reflected in the dominant architecture types. The location of Egypt in general and Cairo in particular, under arid hot conditions has affected greatly the Urban and the Architectural pattern. This pattern is obvious in adopting some elements such as thick walls, Malkafs, domes, courtyards, Mashrabia, and the urban tissue including the organic system.

3.HISTORICAL BACKGROUND

Fatimid Dynasty

The city of Cairo had witnessed a sequence of historical events since its establishment by the Fatimid Commander Jawhar al Siqilli (a previous Sicilian slave) as a new quarter (969-973) onto Fostat (the capital of Egypt since Islamic conquest by Amr Ibn El Ass in 640 which was founded beside Babylon fortress).

During the Fatimid dynasty, a number of magnificent buildings were founded including Al Azhar mosque, a famous mosque and in the same time as the oldest university teaching Islamic faith in its various sects. Other surviving Fatimid structures include the Mosque of Al Aqmar (1125) as well as the monumental gates of Cairo City walls commissioned by the powerful Fatemid emir vizer Badr Al jamali (1073-1094) (CPAS 1992).Beside these elegant constructions, elaborate funerary monuments were founded. The houses were simple and closed, characterised by open courtyards and unstraight entrances, Fountains were also erected in the courtyards together with the use of Malkaf (air catching unit) and Mashrabia.

It is worth to say that since the fatimid came to power, the city expanded gradually later on exhibiting with each age a special architecture style. All of which own their imprints of Islamic art .This art in general is strongly affected by the Islamic faith and traditions and stands in harmony with climatic and environmental aspects. Islamic architecture elements facilitated codes of conduct within the multiple and historical contexts of the Islamic world.

Mamluk Dynasty

Mamluk Dynasty started at the end of Ayubid dynasty of Saladin (1169-1252) after his outrageous triumphs over the many crusaders campaigns.The reign of Mamluks (1250-1517 AD) marked a breathtaking flourishing of Islamic art which is most visible in old Cairo. Religious concepts offered them generous patterns of architecture and art with majestic domes, courtyards and soaring minarets spreading across the city. The Mamluk architecture decorative arts including enamelled and gilded glass, inlaid metalwork, woodwork and textiles flourished under their rule and had a profound impact and influence around Mediterranean both in north (Europe) and south (African north coast). Distinguished Mamiluk rulers established a patronage of public and pious foundations including madarases (schools), mausolea, minarets and pemarestans (hospitals)(CPAS 1992).

4. ELEMENTS OF ARCHITECTURE STYLE AT FATIMID AND MAMILUK DAYNASTIES

Islamic architecture style at Fatimid and Mamiluk Dynasties may be identified with the following Elements. (Fig 1)

-Minarates as Towers, and Mihrab indicating qibla.
-Sahn(courtyard).
-Central Fountains (Maida) used for ablutions.
-Iwan to intermediate between different sections
-Domes, Vaults, Moqarnas and Arches.
-The use for geometric shapes and repetitive art (Arabesque)

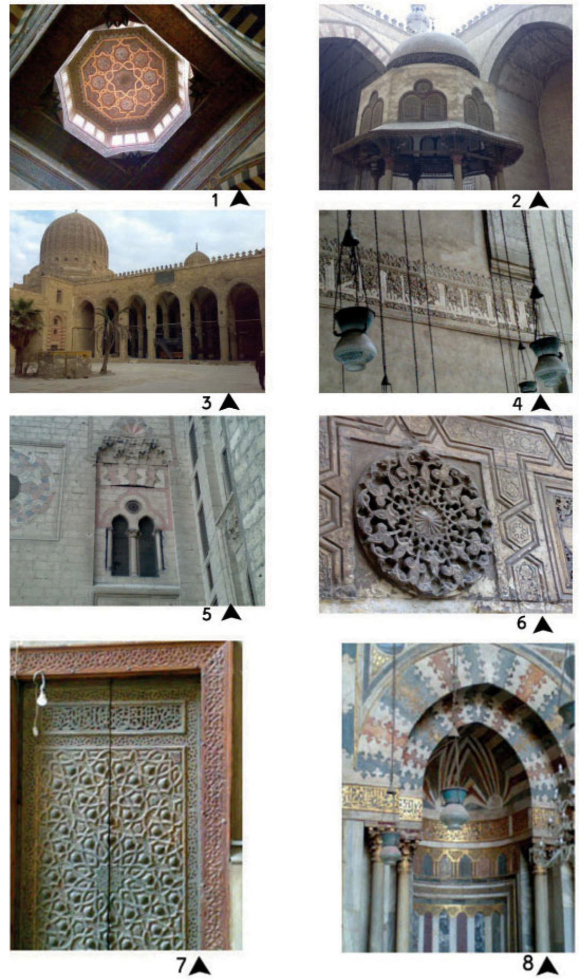


Fig.1. Photo (1) shows the use of ornaments, Photo (2) shows two vaulted Iwans, and a Fountain in the middle of the sahn, Photo (3) shows a sahn as an aesthetical and climatically element of Islamic style.,Photos (4, 6) show the use of calligraphy, Photo (5) shows a style of window with arch and moqarnasat. Photo (7) shows the use of Arabesque decoration, Photo (8) shows the bright ornaments at the Mihrab,

5. COMPARATIVE ANALATICAL STUDY OF THE THREE MODERN INSTITUTIONS.

Oncology Hospital

The location of the Hospital: The hospital is located at a short distance from Cairo Fostat west of Magra El Oyoun wall (Saladin Fortification) in the centre of a random urban district. The area of building and surrounding is about 10 thousand square meters. Three entrances are leading to the main building. The site was also surrounded mainly on the foreground by tracts of green loans and parking lots. In order to establish a better surrounding, several unplanned (randomly planned) blocks of buildings and houses were demolished. However the neighbouring areas are consisting of slaughter yards and their relevant industries forming a serious pollution source (Fig 2).

Fig.2. A satellite image showing the site plan of the hospital and the preceding shots from 1 to 4 show the surrounding streets and buildings. Source: (Google 2008)

Hospital elements and components: The hospital is a building of eight stories with a total area of $10000m^2$. Clinics, emergency and reception constitute the ground floor. These units are accessed through separate entrances from the main street. Other medical departments were distributed in the upper floors according to their functions .The hospital is provided with modern electromechanically systems including lighting, air condition, computer network and an efficient system for water treatment and waste disposal, beside the highly sophisticated medical equipments (Fig 3)

Conceptual Design: The architecture concept adopted is to build a construction with integrated and functional facilities. The core is originally a block of Falluka (boot) form with its sails inspired by its location near the River Nile The concept achieved the optimum level of service efficiency but failed to be in harmony with the prevailing Islamic style coding the region. The designer (jonathan bailley) created an architecture construction which is quite strange to the environment. The establishment in fact came to existence devoid of any Islamic element and the main building is a mere block of western pattern.

The Islamic style however, reveals a dynamic relation between blocks and spaces. To introduce Islamic elements to a modern hospital is a real challenge to the designer and could be inspired from the surrounding Islamic architecture

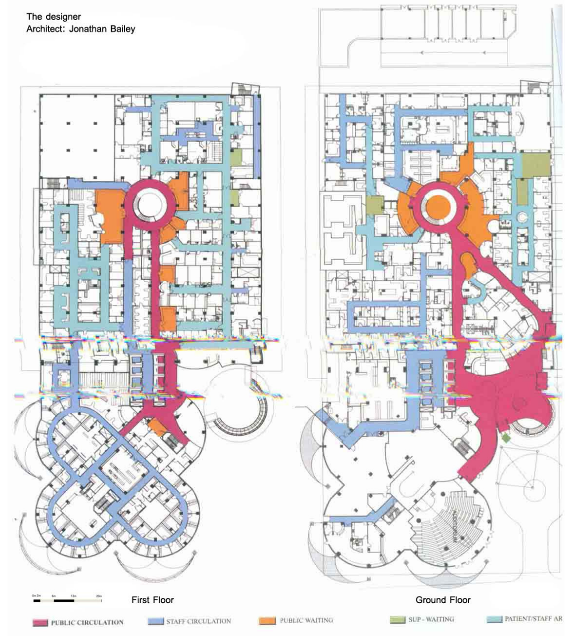


Fig.3. Design Elements of the hospital building at ground and first floors. (TCHP 2002))

The compound of Mashiakhet Al Azhar and Dar El Eftaa

This modern compound is replacing separate old institutions of both buildings. These two institutions are under the control of Al Azhar establishment, the first one (Mashiakhat Al azhar) is the head quarter of the grand Sheiks (Imams) offices and the host of specialized centres for research, publications, Faith dissemination, and international relations. The other one (Dar El Efta) is formally recognized as the only source of fatwas(interpretations of Islamic laws). The two institutions are formerly occupying old buildings, lacking in general enough space and suitable facades. These are the main reasons for having another site with enough spaces for several extended facilities.

The location of the Mashiakhet Al Azhar and Dar El Eftaa: The site was chosen for the new structures in the region of the Fatimid Cairo at a short distance from the Al Azhar Mosque (the famous Fatimid mosque in Cairo.), on a hilltop clearly elevated than the surrounding streets and grave yards at the cross road of Salah Salim and Al Azhar street. The Islamic surrounding environment has the main impact on planning and architecture design of both buildings. The compound is forming an engineering architectural model integrating all elements of Islamic architectural style and art considering in general the historical Islamic landscape (Fig 4).

Fig.4. A satellite image showing the site plan of the compound of Mashiakhet Al Azhar and Dar El Eftaa (Google 2008),
and the terrestrial shots from 1 to 4 show the surrounding streets and buildings.

Elements of Mashiakhet Al Azhar: The building is an administrational construction in general occupying an area of 6000 m^2 in a lot of 18000 m^2. It is composed of 8 stories to accommodate the various administrations according to the program of utilitarian needs and functions. These

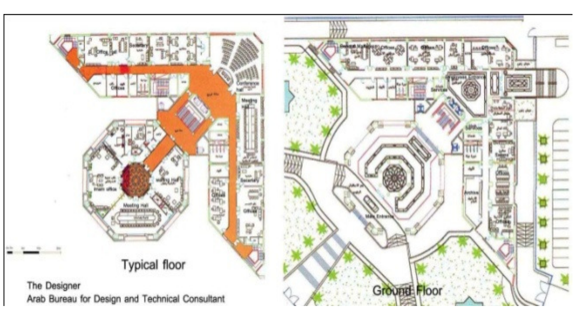


Fig.5. Design Elements of Al Mashiakhet building at ground and first floors. (AMP 19
administrational stories are connected to a central octagonal form in which the office of the grand Imam is located. The eastern principle entrance is designed to lead to the main façade. Entrances of officers and officials are at lateral facades. (Fig. 5

Elements of Dar El Eftaa: This is essentially an administrational building, where consultant services are offered to citizens and authorities. Its area is about 2000 m^2, and composed of five stories. Administration offices are accommodated according to the sequence of functions. The internal courtyard is used as a pray yard surrounded by lateral courts. The office of Mufti (Azharian sheikh) is occupying a central and special location at the main facade looking at outside by a magnificent "Mashrabia (a famous mamiluk architectural element representing modern balcony), (Fig 6).

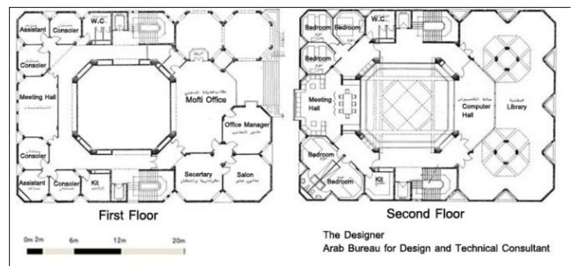


Fig.6. Design Elements of El Eftaa building at first and second floors. (DEE 1997)

Conceptual Design: The conceptual design of these two constructions is compatible with the Islamic heritage of the Mamiluk Cairo. The concept adopted by the designer(ABDTC, local bureau) depended on achieving an Islamic style in a modern and contemporary spirit, integrating in the same time with the many surrounding Islamic buildings and historical landscape, it incorporates the common elements of Fatimid and Mamiluk nearby monuments**.**

6.COMPARATIVE ANALATICAL STUDY OF THE ARCHITECTURE FORMS AND ELEMENTS IN THE THREE INSTITUTIONS.

The study revealed the following results;

External facades

All outside facades of the hospital are mainly in glass, giving an impression of transparency from a distance which is strongly contradicting with the Islamic design concept of the external facades. This, has less openings looking outside while the main and important openings are looking at internal courtyards, which are achieving the important Islamic principle of privacy. The design of both Mashiakhat Al Azhar and Dar El Eftaa came in presenting wonderfully this preceding element (Fig 7).

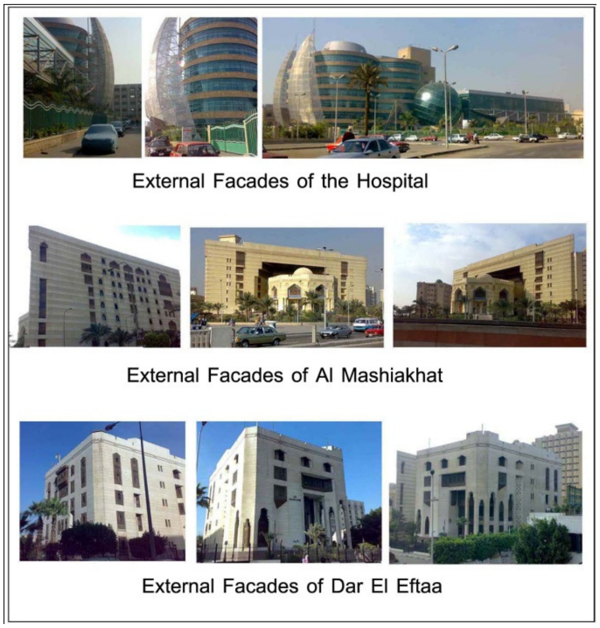


Fig.7.Terrestrial photos showing the external facades of the three institutions.

The glass facades of the hospital though of double layers did not in fact accomplish shaded and/or conditioned interiors. Curtains and centralized air conditions were extensively used. For attaining shades and light refraction, metal structures were used in many parts of the frontal facades. These structures including in some parts sails forms are quite apart from any Islamic style in the surroundings.

Fewer openings of limited lengths in the frontal facades of both Islamic constructions, and using of Islamic treatments such as Mashrabiat, ornaments, finishing coloured materials, and other elements, all came in harmony with the Islamic architectural style.

Openings

Due to the extensive surface areas of the openings at the facades of the hospital, sun rays and glare vision became a problem in day time to the wide spaces in the structure. To overcome this obstacle heavy curtains and doubled glass sheets with argon gas in between as a filter of harmful sun rays is critical in cancer treatments to the children who are put under radiation exposure and chemical treatments. The use of wide glass planes did not result in successful heat insulation. Accordingly, the hospital is using mechanical and electrical means in all times to achieve the previous objectives.

Openings at external facades of both Mashiakhat and Dar El Eftaa are few and in many instances covered with decorative wooden Arabisque to mitigate light intensity and to achieve shading and moisture conservation. The adopted design of openings in the Two Islamic institutions were modernized in a sense that is not loosing its elemental essence (Fig 8).

Fig.8. Photo 1 shows the effect of Mashrabiat at the internal façade on reducing glare and heat transmission inside the building. Photo 2 illustrates the glare resulted from extensive use of wide glass openings in patient rooms and how curtains are used to solve this problem.

Entrances

The Islamic architecture has a common characteristic style, concerning the main entrances which are generally with almost the same height of the building or the first floor, these are strong and confirmative. In the Islamic compound of Al Azhar, the two institutions are designed with this style of entrances in the fronts facing the main streets. Every entrance is obvious, strong and lintelled with a pointed arch .

In the hospital metal structures were used to define and confirm the main entrances. Other entrances are simple usually with front and intermediate glass doors.

Elements of structure form (columns, domes and arches).

Islamic Architecture in the vicinity has known to embrace important and characteristic structural elements. Columns are one of these elements and were transported in the earlier decades of Islam from Churches and Temples (Abd El Gawad 1987).

In later periods columns were modified especially during Mamiluk dynasty exhibiting elaborated forms. The designer of those Islamic institutions inspired different forms of columns using modern materials for coating, like marbles and some other manufactured materials. Dome element was successfully used to cover the bulk of the principle building. Also Aqoud (Arches) of Islamic style were used with similar scales used in the fatimid and Mamluk dynasty with some modifications to give a contemporary style to the construction (Fig 9).

In the hospital building a different concept and culture were adopted by the designer, using for instance a spherical form of glass and metal structures. He used these forms as constructional forms and as decorative patterns, which are adversely compatible with the styles of the surroundings.

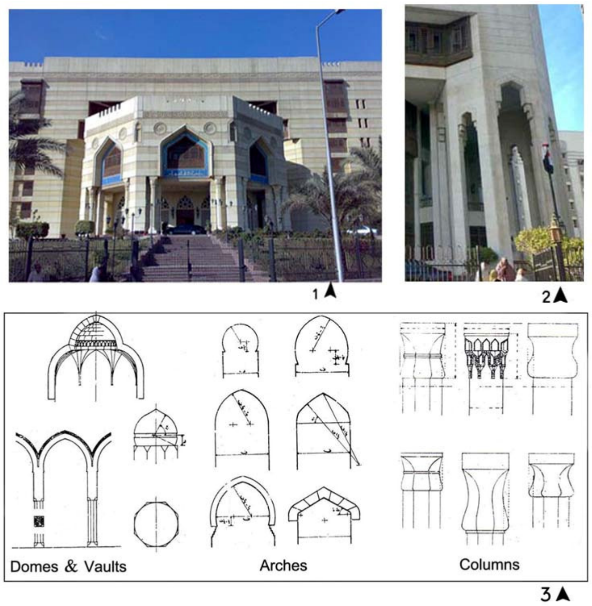


Fig.9**.** Photos 1&2 illustrate some form elements (columns, domes, arches)at Al Azhar compound inspired and developed from the Islamic prototypes as shown in drawing 3. (Abd El Gawad 1987)

Ornaments and Mouldings

Ornamental forms and mouldings in Islamic Mamiluk Architecture are quite different than those in Greece or Byzantine styles. They are reflecting islamic culture and spirit. The decorative principles are resting on the basic foundations of calligraphy, geometry, repetition and multiplication (Clevenot, D,and Degeorge, G, 2000).

Using a Variety of decorative elements, spaces are articulated. Famous ornamental patterns were used in the Islamic compound of Al Azhar. These are used to decorate frontal facades, entrances and some other surfaces (especially Mashrabiat). This is in fact not meant to serve utilitarian needs but rather to give a spirit in harmony with the surrounding buildings. In other words, to define regional identity relevant to the eternal principles of Islam.

Concerning the hospital functions, it is basically needed. The use of Islamic structures and forms would be a challenge to designers. Local designers and architects are encouraged to play this role successfully. The two Islamic buildings in Al Azhar compound are examples of having modern buildings with Islamic spirit and culture. This Islamic approach was unfortunately not adopted in designing of the hospital. All pollution protection elements or clinical regulations could be maintained with an outside of Islamic identity.

Architectural Forms

Architectural Forms established in the Mashiakhat Al Azhar and Dar El Eftaa buildings are consisting of a strong block of large scale balanced with the central axis of the internal courtyard following the important principles of Islamic architecture.

Balance and symmetry around an axis are noticed obviously in the horizontal elevations and plans of the two Islamic buildings. There is also efficient use of Islamic decorative values, resting in unity and harmony with all of the architectural elements(arches, openings, ornaments, and calligraphy)adopted in new scales that recognize the function without wasting the essence.

Hospital forms in the contrary are demonstrating blocks in symbolic shapes and using glass facades giving an inspiration of transparency and conduct with the outside. This approach is contradicting with the Islamic architectural principles. Covering the front with metal structure as a sail gave impressions in harmony with its location near the river Nile but not compatible with the nearby environment (Fig 10).

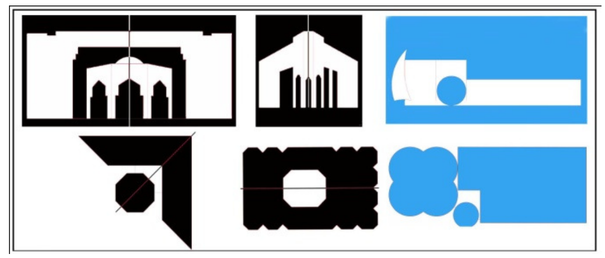


Fig.10. Black and white drawings of Al Mashikhat and El Eftaa showing the balance and the symmetry around the axis of the courtyard. The other blue drawings show the symbolic forming of the Hospital which is not in harmony with the Islamic forms.

Internal Facades

The design of the Islamic institutions depended on decorating the external facades with Islamic elements, repeated also in the internal facades using excessively calligraphic patterns. Mashrabiat and wooden arabesque are used to decorate large openings in the different facades. Marble with geometric patterns and different colours are used in floors, and walls. Ceilings are decorated with geometric units and coloured finishing materials. Carpets and furniture decorated with Islamic repetitive patterns were spread in courtyards and reception halls (Fig 11).

In the hospital interiors imported materials are used in covering walls and floors. Finishing materials used in covering walls are chemically products treated to resist bacteria and microbes. Using these materials and treatments could be used in the interiors of the Islamic constructions without contradiction and could be elaborated to be in harmony with the Islamic spirit (Fig 12).

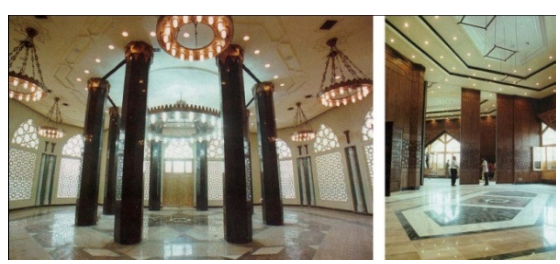


Fig.11. Photos of Al Mashikha illustrate the Islamic ornamental patterns used on internal walls and ceilings .Marbles in geometric patterns is seen in floors.

Fig.12. Photos of the Hospital illustrate the colourful chemical finishing materials and the finishing that resist bacteria and microbes, also showing the imported colourful Terrazzo floor.

7. RESULTS AND DISCUSSIONS

The Site and design of the two buildings, Mashikha and dar El Eftaa of Al Azhar compound have achieved complete harmony with the surrounding environment. The Hospital on the other hand may quite fit its utilitarian needs and functions, but lacks in general environmental and traditional balance with the whole district. Functional and traditional aspects however could be fulfilled successfully in the same time. It is the job and the art of the designers to develop certain Islamic elements to fit the purpose of construction. The reaction of local designers with the traditional and environmental requirements is noticed in many other establishments.

The location of the hospital is another unsuccessful option. It could be located at the expansion districts of Cairo. If there is a certain need to be present in this area; it should be compatible with its environment. The location of such a critical and highly sophisticated institution is greatly contradicting with environmental and traditional requirements.

8. RECOMENDATIONS

A complete compatibility with the environmental and traditional aspects is a challenge facing the designer who is also striving for the application of technological advancements in the relevant fields. The contradiction between the two cultures, traditional and contemporary would call the architect to neglect one for the other; the site in this case would dictate the decision. Environmental and traditional aspects should be carefully considered when great and important projects are to be implemented in the city. Different designs of a project are better displayed to the public opinion for discussion and evaluation. Local architects are called for their role and duty in harmonising between technological and traditional requirements. The advantage of cooperation and assistance with foreign experience is sought in applying modern technology in a frame of integrated system, satisfying both traditional and environmental needs. Architecture competition in big projects of national interest should necessarily be organised for this purpose.

9. REFERENCES

[1] Abd El Gawad.A.T, 1987, **Islamic Architecture** (in Arabic), Anglo library, Egypt.
[2] John D.Hoaa, 1977, **Islamic Architecture**, Harry n. Abrams, Inc, Publishers, New York.
[3] Centre for Planning and Architectural studies, 1992, **Principles of Architectural Design and Urban planning During Different Islamic Eras**, KSA.
[4] Clevenot. D, and. G. Degeorge, 2000, **Ornaments and Decoration in Islamic Architecture**, Thomas Hudson ltd, London, and the vendom Press, New York.
[5] Amin.M.M, Ibrahim.A.L, 1990, **Architectural Terms in Mamiluk Documents**, the American University in Cairo, Egypt.
[6] Steele James, 1994, **Architecture for Islamic Societies today**, Academy Group Ltd, Spain.
[7] Hill.D, Grabar.O, **Islamic architecture and its decoration**, 1964, Faber and Faber, London.
Rice.T.D, Islamic Art, 1975, Thomes and Hudson, USA.
[8] Mohamed, M, S, 1971, **Egypt Mosques** (in Arabic), Al Ahram Commercial, Egypt.
[9] Mostafa.l.S, 1975, **The Architectural Heritage In Egypt**, Beirut print, Beirut.
[10] The Children Hospital Project, 2002, **Tasmeem Journal Magazine, Egypt**.
[11] Al Mashiakhat Project, 1999, **Alam Al Bena Journal Magazine**, Egypt.
[12] Dar El Eftaa Project, 1997, **Alam Al Bena Journal Magazine**, Egypt.
[13] http://www.57357.com
[14] http://www.google.co.uk

Building a Reconfigurable System with Integrated Meta-Model

Ashirul MUBIN
The Graduate School, The University of Alabama
Tuscaloosa, AL 35487, USA

and

Zuqun LUO
Information Systems, Statistics and Management Science, The University of Alabama
Tuscaloosa, AL 35487, USA

ABSTRACT

In systems development activities, the usual processes of requirements analyses, design, development, implementation and testing for a target system are carried towards its production. However, soon after its deployment, a number of unforeseen issues, which were unnoticed at the time of initial requirements analyses, begin to come up as newly emerged specifications of the system. Although not impossible, yet it is very costly to incorporate such additions to a system which is already in production. Therefore, the value of this system will go down along with its user satisfaction levels; and possibly with higher attrition rates. To overcome this, we propose a parallel development strategy of a meta-model, as a counterpart of the target system under development. In this paper, we present a preliminary layout of an integrated meta-model that provides more control over the dynamic system configuration and effectively addresses newly emerged requirements towards building reconfigurable systems. This additional power to effectively manage the system is crucial for extending its lifecycle. As a result of being able to efficiently applying new changes into the system, we see significantly higher level of user satisfaction with increasing utilization of various system features.

Keywords: Meta-Model, Receptor, Probe, Knowledgebase, Meta-base, Analyzer.

1. INTRODUCTION

Even after careful and systematic development of a system, it may still need further modifications after its deployment into its operational environment. This is because, once a system is being used for services, it is encountered with different types of users with varying satisfaction levels [5] and needs at specific points in the system workflow during its operations; thus it is more likely that new unforeseen issues will eventually come up. Therefore, over a period of time, the system will either need a thorough overhaul or a full replacement of the system. But both of these options are very expensive and time consuming, since the entire system will need to be reengineered with new sets of requirements accumulated. Hence, it is difficult to achieve a continued higher level of end-user satisfaction throughout the system lifecycle. However, this highly desirable feature may be achieved to some extent with captured trends of system usage patterns, regularly identifying newly emerged requirements, and, if possible, by applying these new specifications into the system.

Our objective of this work is to build a parallel system which is simultaneously developed with the target system for tracking the shifting behavior of the users, and generating new system specifications to reconfigure the target system that will meet its ongoing or future needs. The setup of this parallel system, called a *meta-model*, helps extending system lifecycle further by making it more useful for the end users to serve over a long period of time [2] without any major overhauling.

2. SPECIFICATION OF A META MODEL

A meta-model can be viewed as a wrapper system around its target system, as shown in figure 1. Within the context of systems development perspective, a meta-model can be defined as the structural specifications of another system. These specifications can be readjusted based on the feedbacks from its operational environment, system usage pattern and the target system itself.

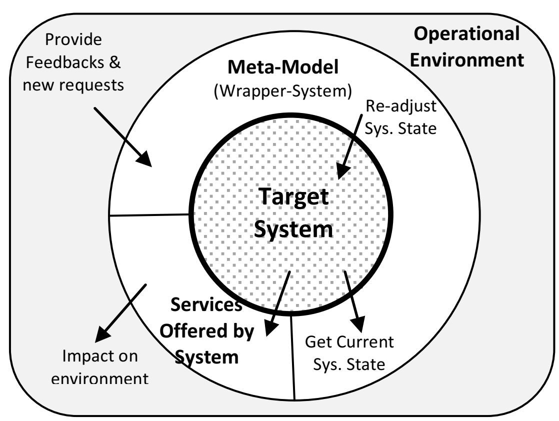


Figure 1. Role of a Meta-Model

The role of a meta-model is vital, in terms of collecting feedbacks, capturing newly emerged specifications from various sources, and utilizing system usage data. This setup helps to formulate newly adjusted system state, which is comprised of the specifications of objects and processes [1]. Such directed formulation becomes more mature over time, since analyses from previous experiences are applied in each iteration during updating system states.

Meta-model is a system by itself. Therefore, it will also need its own specific system requirements, analyses, design and implementation phases. However, since its objects and processes are in fact the blue prints of objects and processes of the target system, its development activities will continue in parallel with the target system development.

Components of a Meta-Model

A carefully designed system is comprised of objects, processes and their meaningful interconnections, that can be built through the Object-Process Methodology [1,2]. To gain programmatic control over these constituent parts of a system, we deduce parallel sets of meta-objects and meta-processes constructed from the system. The meta-objects and meta-processes contain the detailed specifications of the current system state and reflect dynamic behavior of the system for an instant of time. Table 1 describes major components of a meta-model.

Table 1. Meta-Model Components

Component	Functional Role
Meta-Object	Holds detailed specifications of its counter-part system-object. It may represent a set of possible re-adjustable parametric values.
Meta-Process	Holds detailed specifications of its counter-part system-process. It may represent a set of re-adjustable rules or conditions applied on meta-objects.
Meta-base	A repository in a relational database for recording system configuration values accumulated over time.
Knowledge-base	A collection of historical data on system usages, activity logs, newly emerged requirements, system configuration at each iteration, etc. The output from the Analyzer and the changes in Meta-base enrich the experiences in the Knowledgebase.
Analyzer	Analyzes the current system state and behavior against the configuration data recorded in Knowledgebase (KB) and outputs instructions for new specifications for meta-base entry.
System Probe	Means of collecting current system state & usage data.
System Receptor	Means of incorporating newly generated system configurations into the system.

As mentioned in table 1, we introduce two sets of instruments: *Probes* and *Receptors* to integrate the meta-model with its target system. The purpose of a *Probe* is to capture system usage patterns, rank features by popularity, collect indirect feedbacks, and infer any suggested upgrades at various points in the workflow while the system is being used. A *Receptor* provides mechanisms to dynamically apply (transport) new changes into respective system objects and processes from their counter-part meta-objects and meta-processes. Table 2 lists some of the key properties of these two components.

Table 2: Properties of a Probe and a Receptor

Properties of a Probe (P-Values)	Properties of a Receptor (R-Values)
Probe_Info	Receptor_Info
Probe_HostInfo	Receptor_HostInfo
Probe_UserInfo	Receptor_UserInfo
Probe_Location	Receptor_Location
Probe_FeatureInfo	Receptor_FeatureInfo
	Receptor_Read_MetaObjSpec
Probe_ReadSysState	Receptor_Read_MetaProcSpec
Probe_WriteKB	Receptor_Write_ObjSpec
	Receptor_Write_ProcSpec
Probe_SetActivity	Receptor_SetActivity
Probe_GetActivity	Receptor_GetActivity
Probe_Timestamp	Receptor_Timestamp
Probe_Log	Receptor_Log

Both *Probe* and *Receptor* contain similar data sets and operations, except for opposite communications between the target system and the meta-model.

3. EXTENDED SYSTEM ARCHITECTURE

With integrated meta-model, the target system architecture is, thereby, extended to address additional functionalities necessary for capturing dynamic system usage data, changing of user behaviors and applying new specifications while the system is in production. Figure 2 details the extended system architecture with system *Probes* and *Receptors* as the two interfacing components.

Organization of Architectural Components

Any system is surrounded by its operational environment, where the users and customers consume the services provided by the system. Thus a system has an impact of its resulting services on the people and its environment. This triggers external influences back into the system over a period time. By capturing these influences, we get a partial picture of the newly emerged system requirements. *Probes* continue to capture the system usage activities based on current state of the system, updating all its *P-values*. The underlying knowledgebase keeps on recording all these activities, along with the data from associated adjustments into the meta-structure through the meta-base. The *Analyzer*, in between the two databases, attempts to parse through the change in configuration data provided by KB, and formulate specifications of new configuration with new sets of *R-values* saved into meta-base. Receptors apply their respective *R-values* into the system for reconfiguration.

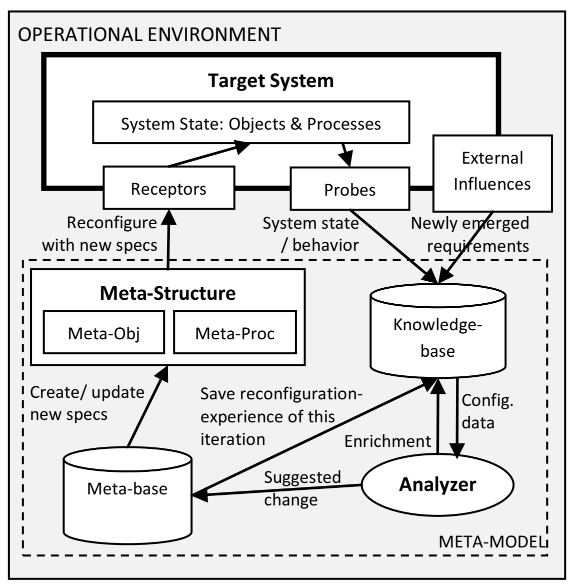


Figure 2. Proposed 3-Level Iterative Architecture

Nested Iterations

To properly configure meta-objects and meta-processes, and then, in turn, re-configure system objects and system processes, we have identified that three levels of nested iterations are necessary. The outer primary iteration is the loop initiated from the system probes into the knowledgebase, then to the meta-base through the analyzer, and then back to system receptors through meta-structure. The other two iterations are passed through the meta-base and analyzer back into knowledgebase. Together, these three iterations reconfigure the meta-structure, system objects & processes and eventually upgrade the knowledgebase by recording history of successful iterations.

Enriching the Knowledgebase

A meta-model properly integrated into the target system, a knowledgebase (KB) captures the learning experiences from system's performances, behaviors, usage patterns, shift of customer needs and light-weight feedbacks. It is important to have this knowledgebase updated during each iterative cycle [2] so that it can later direct the Analyzer to determine more correct approach to reconfigure system components and properly set the meta-structure to apply necessary changes through the *Receptors*. The experience of updating the system configuration is fed back into knowledgebase for further enrichments and more trained or educated predictions.

4. COMPARATIVE STUDY

Because of the parallel development strategies for the target system with the implementation of its meta-model, the initial development workloads are usually twice comparing to traditional system development activities. However, this initial load eventually pays off during the extended system lifecycle for long term benefits.

Traditional System vs. Reconfigurable System

A traditional system without a meta-model often cannot address newly emerged updates that may come up while the system is in production; therefore the system value will degrade to a lower level at the inception of each such update requests. During this inception phase, some end users may switch to peer service providers; but some loyal customers will continue to use the system at a lower User Satisfaction Index (USI) values [2,5,6], even with low returns out of trust-worthiness. In any case, the stakeholders will need to go through expensive re-engineering process of overhauling the system or simply replace with a new system at a much higher cost with longer downtime. Either way, it will impose negative impact on the underlying business processes. On the other hand, with the support of a meta-structure, the system becomes reconfigurable, and it can readjust itself with appropriate system meta-data over a period of time, thus continue to provide services for extended time.

Projected System Value and Maintenance Cost

Figure 3 provides simplified projected trends of system values [2] for both kinds of systems during their respective lifecycles. Once a reconfigurable system is deployed, its meta-base may require some time to establish the model from real-time system usage data. *Establishment Time* (ET) is the time to initialize the entire sets of system objects and processes, as well as corresponding meta-objects and meta-processes from its production environment. This process is required to properly setup and fine-tune system parameters; however, the target system is not affected in terms of its regular operations, since it continues to provide its intended services (the purpose it was built for) to the customers or end users.

After the initial establishment time, both the target system and its meta-structure will continue to run simultaneously with desired changes in place – at this point, users will notice increased system value due to it being able to address any desired changes into the system. On the other hand, the traditional system without a meta-model will continue to have its degraded system value [2,5].

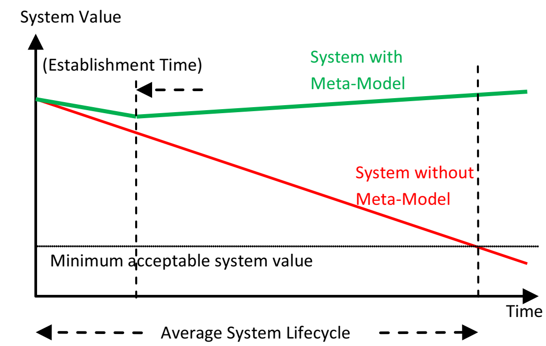


Figure 3. Projected system value (such as, user satisfaction index) over time

Figure 4 explains a projected cost of maintenance activities throughout their respective lifecycles. Initially, the cost of building a reconfigurable system may be close to twice as developing a traditional system without any associated meta-model. However, once the system is in production (after the initial establishment time), the overall cost of maintenance will go down and will continue with much lower costs toward its extended system lifecycle. On the other hand, the traditional system will experience higher maintenance costs, due to the need for alternate ways, if any, to accommodate unfulfilling requirements that will emerge throughout its lifecycle.

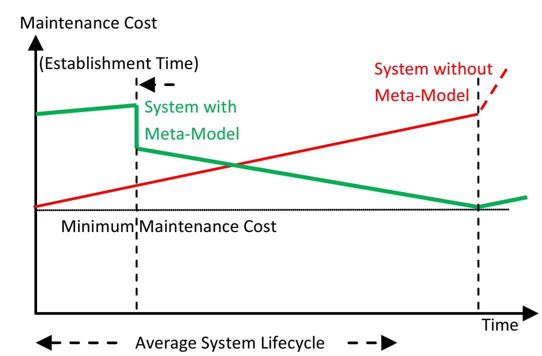


Figure 4. Projected maintenance cost (such as, efforts needed to address new changes) over time

Observations

Table 3 shows the added User Satisfaction Index (USI) values [2,5] estimated from the number of updates in system reconfigurations needed for three selected project works that are integrated with semi-automated meta-structures: (1) ADMIN-ROLES: Manages all graduate-level administrative assignments, (2) APPTRACK: a system to manage thousands of graduate applications online in a distributed fashion, and (3) ITAP: a system for managing the training data of new international teaching assistants. As we can see from the table, ADMINROLES and ITAP are becoming more mature with declining number of iterations through the system updates, with average of 3.33 and 2.00 less iterations in subsequent years, respectively.

Table 3. Effects on System Value (eg. USI changes) due to system updates through outer-level iterations

	ADMIN ROLES		APPTRACK		ITAP	
Year	#Iterations	USI	#Iterations	USI	#Iterations	USI
2006	2	n/a	-	**-**	15	n/a
2007	17	**+3**	20	**+2**	29	**+3**
2008	18	**+2**	23	**+2**	11	**+3**
2009	10	**0**	34	**+2**	7	**+1**
2010*	5	**+1**	14	**+1**	-	**-**

**activities till mid March, 2010.*

However, APPTRACK is still experiencing high dynamicity in reconfiguring of its specifications, with an average of 4.66 more iterations in subsequent years. In all cases, these systems are steadily experiencing acceptable system values throughout their continued lifecycles.

Table 4. Maintenance Cost (eg. average days to address closely related changes) due to system updates

	ADMIN-ROLES	APPTRACK	ITAP
Year	Cost	Cost	Cost
2006	88	-	21
2007	110	41	8
2008	63	9	31
2009	2	6	7
2010*	10	5	1

**Until May first week.*

In table 4, the cost of maintaining the three selected projects have been presented in terms of average days required to complete a task request, either from end users or from the system administrator; i.e. to cover any sort of maintenance activities. Here, the cost value has been approximated proportionately directly from this calculated average number of days it required for the developers to spend time and related expenses (wages, tools) for completing each task request. Comparing with the generalized behavior depicted in figure 4, we see that the cost increases during the earlier periods (establishment time) for these projects, and after sometime, each project's cost descended over gaining more maturity of the underlying knowledgebase.

5. CONCLUSION

With the wide spread adoption of internet and computer-based systems, users are more informed of similar service providers and able to quickly switch to the best service available. Therefore, in this highly competitive market, a system should be able to sense the shifting changes in its dynamic usage patterns and apply any newly emerged requirements thereof. In these circumstances, a meta-model in place can serve the purpose with minimal efforts. Our brief study presented in this paper indicates that a supporting meta-structure always tend to provide significantly more controlled administration and higher manageability towards building reconfigurable systems. Such systems are able to satisfy user needs as they come up over time, therefore, the retention rates are often high. However, not all systems need to be built in this way; the surrounding operational environment, selected groups of users and the type and complexity of the target system itself will determine whether building a meta-model will be helpful in the long run [3]. This requires a clear vision and objectives of the underlying business processes. Applying appropriate design patterns to *Receptors* and *Probes*; and adding predictive models [4] to the *Analyzer* will solidify further enhancements to the concept of meta-modeling.

6. REFERENCES

[1] Dov Dori, Object Process Methodology, Springer, 2002.

[2] A. Mubin, D. Ray, R. Rahman, Architecting an Evolvable System by Iterative Object-Process Modeling, IEEE CSIE 2009, Los Angeles, CA.

[3] M. M. Lehman, Programs, Life Cycles, and Laws of Software Evolution, Proc. IEEE, 68:1060-1076, 1980.

[4] "Using Predictive Analytics to Achieve Stellar ROI," White Paper by SPSS and Peppers & Rogers Group, 2008.

[5] Ives, B., Olson, M. H. and Baroudi, J. J., "The Measurement of User Information Satisfaction," Communications of the ACM, October 1983.

[6] Bardoudi, J. J., Olson, M. H., Ives B., "An Empirical Study of the Impact of User Involvement on System Usage and Information Satisfaction," Communications of the ACM, March 1986, Vol. 29 No. 3.

Using "System Sensing" During the Implementation of a New Mechatronics Engineering Curriculum

Billy O'Steen*
Educational Studies and Human Development, University of Canterbury
Christchurch, New Zealand

Erik Brogt
Academic Development Group, University of Canterbury

XiaoQi Chen
Department of Mechanical Engineering, University of Canterbury

J. Geoff Chase
Department of Mechanical Engineering, University of Canterbury

*Corresponding author

ABSTRACT

"System sensing" [1], or a feedback loop, has been integrated into the implementation of a new mechatronics engineering curriculum at the University of Canterbury through a sustained, three-year collaboration between engineering lecturers and academic developers. Data were collected each year from the first cohort of students and lecturers through focus groups, course evaluations, specifically designed surveys, and observations. The data were analysed by the academic developers and results and recommendations were fed back to the engineering lecturers so that they could adjust the curriculum, the teaching, and the assessments to better meet the goals they had in mind when designing the new curriculum such as: students engaged in significant design projects at every year and a strong connection with industry [2]. Positive outcomes from this approach included statements by mechatronics graduates that they had obtained core skill sets in both mechanical and electrical instead of an initial lack of identity as "neither mechanical nor electrical."

Keywords: Academic Development, Curriculum Development, Engineering Education, Mechatronics Education.

INTRODUCTION

The mechatronics engineering professional programme was started at the University of Canterbury in 2003, with a limited intake of 15 students. All students in engineering take common courses in physics, mathematics, engineering mechanics, foundations of engineering and mathematical modelling in the first (intermediate) year and specialize in the following three years (1st, 2nd, and 3rd professional years), leading to a BE (Hons) degree. Mechatronics, as a hybrid pathway between mechanical and electrical engineering, faced challenges in the development of the curriculum for these three professional years.

Originally, the mechatronics programme combined essential and existing topics from mechanical engineering, electronics, and computer engineering and was essentially a combination of relevant courses offered in the departments of Mechanical Engineering and Electrical Engineering. There was, however, a lack of coherence and a systemic approach in delivering the "synergistic integration of the three components – mechanical engineering, electronics, and computer control," which is supposed to be the cornerstone of mechatronics. As a result, several challenges soon surfaced:

- Students lacked formal prerequisites for some classes. Consequently they had limited choices for electives as their study progressed.
- A lack of laboratories and design projects led to a focus on teaching from the textbook, leaving students with insufficient exposure to practical-oriented and problem-based training.
- Students were confused about their academic identity. They felt that they were neither mechanical nor electrical engineers.

Partly as a result of these challenges, in the first graduation year of 2006, only six out of the original 15 students enrolled in the first professional (2nd) year of mechatronics completed their degrees.

These challenges called for a curricular overhaul of the programme in order to continue offering the degree pathway. This process began in late 2006. The new 2nd year curriculum was rolled out in 2007, the 3rd year curriculum in 2008, and finally the 4th year curriculum in 2009.

The curriculum development process deliberately sought collaboration from colleagues outside the College of Engineering. In particular, these included academic developers from the University Centre for Teaching and Learning. The academic developers' role was to do "system identification"; obtaining input from students, academics, and industry, and "system sensing"; acting as a feedback loop where information from the output of the (curricular) system is monitored, evaluated and fed back in order to better accomplish a goal.

This curriculum development model allowed for monitoring the learning outcomes against a set of parameters in a timely manner, continually refining the course components and assessments, and optimising the delivery of the degree programme. Particular attention was paid to getting feedback on and adjusting the three new courses that were developed for the professional years. These course are: ENMT201: Introduction to Mechatronics in the second year, ENMT301: Mechatronics System Design in the third year, and ENMT401: Mechatronics Research Project in the fourth (final) year. They are taken by students who enter the Mechatronics Engineering degree program, BE (Hon), after completing the common engineering curriculum in their Intermediate Year (first year).

CONSIDERATIONS FROM THE LITERATURE

The following considerations from the fields of academic development and engineering education guided the collaborative efforts in the mechatronics programme.

Using Student Feedback

While collecting feedback from students has been used for several decades as a means of measuring perceptions of teaching quality, its usefulness in improving teaching and curriculum development "is dependent on the extent to which staff respond to and apply the information obtained in this way" [3]. Thus, to create a more responsive system of delivering a curriculum would suggest determining how to incorporate students' data into ongoing program design.

In Situ Academic Development

Prebble, et al. found in their synthesis of research on academic staff development that "the academic work group is generally an effective setting for developing the complex knowledge, attitudes and skills involved in teaching" [4]. Therefore, the combination of engineering content experts and academic developers, each bringing a different skill set, could be fruitful in the development of a quality mechatronics engineering curriculum.

Redesigning Engineering Education

According to an article by Basken in *The Chronicle of Higher Education* [5], a new report from the Carnegie Foundation for the Advancement of Teaching, *Educating Engineers: Designing for the Future of the Field*, is a reiteration of warnings from the National Science Foundation and the National Academy of Engineering "that American engineering education is too theoretical and not hands-on enough." While Basken indicates that colleges of engineering have known for quite some time that both students and employers desired a more relevant curriculum, both faculty members and accreditation practices are often more wedded to the traditional approach. Hence, the envisioned emphasis on practical and design work in the mechatronics curriculum was in accordance with international directions for engineering education.

These considerations regarding using student feedback, in situ academic development, and redesigning engineering education, indicate that a responsive and effective approach to curriculum design would include:

- collecting student and lecturer feedback in ways that go beyond standard teaching and course evaluations.
- using that feedback in situ and in a collaboration between academic developers and discipline-based lecturers.
- placing that feedback within the context of calls for redesigning engineering education in a more hands-on manner.

METHODS OF DATA COLLECTION AND ANALYSIS

Starting from an inquiry-based learning [6] approach where the engineering lecturers' questions guided the collaboration, data were collected by the academic developers in 2007, 2008, and 2009 from the first cohort of students and lecturers as they experienced the new curricula. Focus groups, course evaluations, specifically designed surveys, and observations served as the primary collection instruments.

The data were analysed by the academic developers and results and recommendations were fed back to the engineering lecturers so that they could adjust the curriculum, the teaching, and the assessments to better meet the goals they had in mind when designing the new curriculum such as: students engaged in significant design projects at every year and a strong connection with industry [2]. In addition, final reports were generated and shared with the Board of Studies that oversees the Mechatronics Program and consists of academics from the Departments of Electrical and Mechanical Engineering.

FINDINGS

The data collected from the same cohort of students at the conclusion of each new course for three years provided feedback specific to both the individual courses and the whole programme. A summary of the findings per course will be followed by the conclusions and implications for the overall mechatronics curriculum.

ENMT201: Introduction to Mechatronics, 2007

This second year course is the first full mechatronics design course that students in the program take. It is both an introduction to the discipline of mechatronics and a combination of mechanical and electrical engineering knowledge. Its content includes introduction to mechatronics, sensors and actuators, basics of instrumentation, circuit analysis, computer-aided design, and introduction to control.

Alongside coursework, this design course consists of a series of laboratories in the first semester. Each lab project is a self-contained project exercise addressing a specific application. Students working in pairs have to implement control interface, design and write control logic. These lab projects are:

- Introduction to ladder logic
- Control inputs, outputs, and sensors
- Car washing process automation
- Water tank level control
- Stepper motor control
- DC motor velocity control
- AC motor control

In the second semester of the course, students are tasked to develop a fully functional control system using Programmable Logic Controller to control a 5-story elevator driven by DC motors. Figure 1 shows a Programmable Logic Controller (PLC) rig that is built in house, and the 10:1 scaled down elevator modeled after the actual 5-sotry elevator in the Mechanical / Civil Engineering Building at Canterbury.

Figure 1. Elevator control project using PLC (left) and tested on the model elevator (right).

The data collected from students in the ENMT201 course in 2007 suggested that students: enjoyed the class, found the content appropriately challenging, and developed a sense of programme community or camaraderie through their experiences. In the qualitative data, the areas that students thought could be improved were primarily logistical with: more equipment for particular labs, coordination of assessments with other courses, same location for lectures, and more explicit coherence or explanation for sequencing of topics.

These findings from the ENMT201 course were fed back to the lecturers, programme coordinator, and the Board of Studies that oversaw the curriculum development. The curriculum for the following year was developed and implemented while considering these findings. One of the improvements was to streamline the lab projects with the aim of maximising the learning outcomes within desired contact hours. Also, the continuous assessments have been distributed more evenly throughout the year, and avoided bottlenecks.

The elevator design project exposes students to controller design using the PID control theory covered in the course work. There was a caution whether such design skill in the junior year was too hard for students. The evaluation of the course proves that students are capable of mastering that skill set. The number of model elevators has doubled from 2 to 4, which allows each team to have more on-machine time for debugging and testing.

ENMT301: Mechatronics System Design, 2008

This course provides students with an intensive opportunity to apply their knowledge from lectures to the creation of a robotic search and rescue vehicle in the Canterbury RoboCup Competition. The project is an integral part of the whole year design course. Students, in teams of three, work in a dedicated Mechatronics Design Laboratory supervised by two instructors and one senior mechatronics technician. The design project requires students to design and build a mobile robot capable of quickly locating and gathering three objects within the field of play. No human intervention is allowed once the robot begins operation. Figure 2 shows the truck base fitted with a Qwerk controller, which forms the standard development platform.

The robotic system must have the following capabilities:

- System hardware is to be attached to the provided truck base and the interfacing managed via a Qwerk microcontroller, operated remotely from a networked computer.
- Targets must be collected unharmed and stored securely on the vehicle.
- The robot must be able to collect cups from any possible locations including corners or alongside walls.

The students are expected to achieving the following learning goals:

- ability to identify the problem requirements;
- ability to generate and evaluate design concepts;
- ability to design and fabricate a manipulator for handling the targets;
- ability to design and fabricate appropriate sensing mechanisms;
- ability to design robotic control software to accomplish the prescribed tasks;
- ability to integrate, test and debug the system; and
- ability to communicate, document, demonstrate and present the design and results.

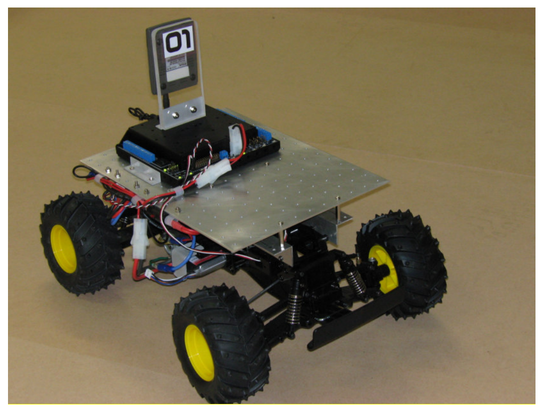


Figure 2. Basic mobile robot platform fitted with an embedded controller.

The data collected from students in the ENMT301 course in 2008 indicated that they:

- appeared to immerse and to enjoy themselves in designing and building a search and rescue robot. Five different students used the word “fun” in individual surveys and 100% of the respondents believed that they had accomplished something significant in the course and 100% would recommend this course to others.
- did not find the project too daunting. There was a discrepancy as to what level of guidance students thought they needed, either the same amount as this year or an increased amount.
- saw the identity of themselves in the programme to be inherent in the nature of the course (a designated lab space, the team approach, a “cool” project).
- saw the lasting lessons of the course to be what they learned about: the design process, project management, and working in teams.

These findings from the ENMT301 course were fed back to the lecturers, programme coordinator, and the Board of Studies that oversaw the curriculum development. The curriculum for the following year was developed and implemented while considering these findings. One of the adjustments to the course was to introduce computer vision in the classroom. Hence, students are now able to design and implement a vision system for searching the targets.

ENMT401: Mechatronics Research Project, 2009

This final year capstone research project consists of a year-long mechatronics design exercise. Students can work either in teams or individually. Most projects are sponsored by industry and students are responsible for all aspects, including organization, management (both time and budget), project proposal, design and prototyping, and final reporting. Each project has an academic supervisor and an industrial mentor, addressing a real industrial problem that does not have an off-shelf solution. As such, it requires substantial research and innovative design. Figure 3 illustrates the delivery of a wall climbing robot for welding a stainless steel tank.

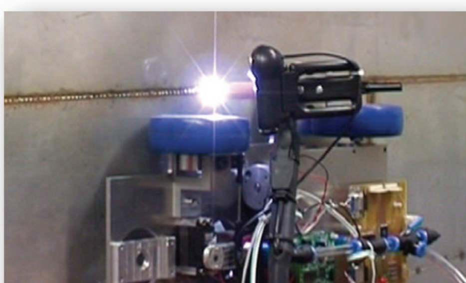


Figure 3. Wall climbing robot for the automatic welding of a stainless steel tank.

The data collected from students in the ENMT401 course in 2009 indicated that they:

- thought they learned considerable skills in the project, with an emphasis on non-technical, managerial, skills.
- saw areas of improvement could include increasing the timing of the lectures, clarity of project briefs, clarity about assessment, more specific mechatronics projects and, to a lesser extent, support and logistics.

These findings from the ENMT401 course were fed back to the lecturers, programme coordinator, and the Board of Studies that oversaw the curriculum development. The curriculum for the following year was developed and implemented while considering these findings. The improvements implemented included scheduling the separate management course in synchronisation with the research projects. Also, the assessment schedule was structured with more clarity. There have now been more mechatronics research projects sponsored by industry, resulting in plenty of projects for the capacity of the class.

Programme review, 2009

The graduating class from 2009 (the first to have gone through the re-designed curriculum) was asked to reflect upon the programme as a whole, and to identify strengths and weaknesses. Students indicated that they:

- were very pleased with the mechatronics programme overall and that the department has succeeded in creating an “academic home” for the students.
- considered the programme to be very time-intensive and demanding and noted that several topics are covered multiple times in different mechatronics papers.
- had a desire for more, and structured exposure to industry throughout the programme.

Collectively, this review along with the data from the courses will be considered by the lecturers, programme coordinators, and the Board of Studies as the mechatronics curriculum continues to be developed.

CONCLUSIONS AND IMPLICATIONS

The combination of engineering and educational expertise in developing the new mechatronics curriculum has proven to be a successful endeavor. The system sensing and feedback facilitated by the academic developers brought in an objective perspective and new impetus. The non-engineering academics complemented engineering academics by bringing valuable insights in terms of setting and achieving learning goals, managing students’ expectations, and advising on collecting feedback.

Arguably, students, staff, and the departments were more open to collaboration, feedback, and data collection as the academic developers were outside of the traditional line management structure, and were thus seen as neutral. This experience at the University of Canterbury has led to the implementation of several effective approaches for mechatronics education, which included integrating labs and design projects into and across courses and cooperative learning. In addition to the curricular adjustments, other positive outcomes involved the students with mechatronics graduates stating that they felt "both mechanical and electrical" in the core skill sets instead of their initial lack of identity as "neither mechanical nor electrical."

After 3 to 4 years' concerted effort, the Mechatronics Engineering Programme at the University of Canterbury has developed into a premier engineering programme that attracts top students nationwide and overseas. It has grown to an intake of 30 students per year, with room for expansion. The graduates are sought after by industry. Further work is needed to monitor the graduates' profiles and industrial acceptance, which will serve as another feedback in our work toward excellence in mechatronics engineering education.

This merger of mechatronics engineering content and expertise with the field of academic development has provided all involved with a unique opportunity to experience a best-practices model of inter-disciplinary collaboration with the subsequent students of the mechatronics program being the ultimate beneficiaries. It is anticipated that further beneficiaries of this transferable process may be other departments who develop their curricula by collaborating with academic developers.

REFERENCES

[1] D. Shetty, J. Kondo, C. Campana & R. Kolk, "Real Time Mechatronic Design Process for Research and Education", **Proceedings of the 2002 American Society for Engineering Education Annual Conference and Exposition**, 2002.

[2] X. Chen, P. Gaynor, R. King, G. Chase, P. Bones, P. Gough, & R. Duke, "A Project-based Mechatronics Program to Reinforce Mechatronic Thinking – A Restructuring Experience from the University of Canterbury", **Proceedings of the 17th World Congress the International Federation of Automatic Control**, Seoul, Korea, July 6-11, 2008.

[3] R. Ballantyne, J. Borthwick, & J. Packer, "Beyond Student Evaluation of Teaching: Identifying and Addressing Academic Staff Development Needs", **Assessment & Evaluation in Higher Education**, Vol. 25, No. 3, 2000, pp. 221-236.

[4] T. Prebble, H. Hargraves, L. Leach, K. Naidoo, G. Suddaby, & N. Zepke, **Impact of Student Support Services and Academic Development Programmes on Student Outcomes in Undergraduate Tertiary Study: A Synthesis of the Research**, Wellington, New Zealand: Ministry of Education.

[5] P. Basken, "Why engineering schools are slow to change", **The Chronicle of Higher Education**, retrieved on January 23, 2009 from: http://chronicle.com/daily/2009/10003n.htm

[6] V. Lee (Ed.), **Teaching & Learning Through Inquiry: A Guidebook for Institutions and Instructors**. Sterling, VA: Stylus Publishing, 2004.

Internet of the Future Non-Engineering Challenges

Gilson Schwartz *, Edison Spina †, José Roberto de Almeida Amazonas ‡
* School of Communication and Arts
University of São Paulo
Email: gilson.schwartz@gmail.com
† Department of Digital Systems and Computing
Escola Politécnica of the University of São Paulo
Email: Edison.spina@poli.usp.br
‡ Department of Telecommunications and Control Engineering
Escola Politécnica of the University of São Paulo
Email: jra@lcs.poli.usp.br

***Abstract*—The Internet of the Future is a fascinating theme from the point of view of both engineering and education as well. Terms like pervasive, ubiquitous have become known for a growing number of digital *netizens* due to the presence of Internet in our daily lives. Quality of Service (QoS) and Quality of Experience (QoE) have become the buzzwords in the network engineering community. However, we dare to say that the engineering challenges faced by the Internet of the Future are *easy*. They are easy by the single reason that we know what is at stake. So, in this paper we address the challenges of the Internet of the Future from the perspective of the Systems Engineering analyzing it as a socio-technical complex and from the perspective of the Iconomics considering the beings, things and icons vertices.**

I. Introduction

The Internet of the Future is a fascinating theme from the point of view of both engineering and education as well. Terms like pervasive, ubiquitous have become known for a growing number of digital *netizens* due to the presence of Internet in our daily lives. Quality of Service (QoS) and Quality of Experience (QoE) have become the buzzwords in the network engineering community. However, we dare to say that the engineering challenges faced by the Internet of the Future are *easy*. They are easy by the single reason that we know what is at stake.

The Internet of the Future engineering challenges may be summarized in obtaining the best transmission quality between any set of end-users. This problem may be *easily* solved by laying a *perfect* wire between any pair of communicating users. *Perfect* is to be understood as a transmission mean where:

i) the delay would be zero;
ii) the signal attenuation would be zero;
iii) the noise influence would be zero;
iv) the signal distortion would be zero.

Despite of knowing that there is no solution for such problem, first of all because the speed of light, thanks to Einstein, is limited to $c = 300,000\ \frac{km}{s}$, second, it does not exist such *perfect* wire and, last but not least, it is not neither economically nor environmentally feasible to lay a wire between any pair of communicating users, on the other hand, assuming such hypothetical solution we know what is the best quality that could be achieved and this is the direction that the engineering efforts should follow, trying to get closer to this ideal level of performance. In short, *engineering knows where to go*.

However, does anyone know, in what is Internet related:

- what is its economic value?
- what should be the rules governing access, exploitation, intellectual property rights, content distribution?
- how do people represent and communicate values and expectations associated to Internet related actions, projects and technologies?

Conceptually, the challenges of network design and implementation (“social weaving”, so to speak, as in assemblages and re-assemblages of actor-networks) are compounded by the simultaneous interaction of space, time and symbol - the playful evolution of this human e-infrastructure corresponds with values, projects and icons for the audiovisual e-superstructural grids.

The contribution of this paper consists on analyzing the challenges of the Internet of the Future from the perspective of the Systems Engineering looking at it as a socio-technical complex and from the perspective of *Iconomics*.

Iconomics relies on the triad: icons, things and beings, meaning that the actors will benefit from this new era if they are able to make appropriations in all three dimensions.

Iconomic appropriation depends on the symbolization process that, in the present context, has to be performed in the realm of groups and networks.

After this brief Introduction, in Section II we present a review of recently published conclusions about the non-engineering challenges of the Future Internet. In Section III we discuss the Internet of the Future from the perspective of the System Engineering and in Section IV from the perspective of *Iconomics*. Section V summarizes our conclusions and indicates some future work.

II. A review of the recognized non-engineering challenges of the Future Internet

This section presents a brief review of recognized non-engineering challenges of the Future Internet based, mainly, on three recent publications: [1], [2] and [3].

A. Future Internet Socio-Economics - Challenges and Perspectives

According to [1], socio-economics aims to understand the interplay between the society, economy, markets, institutions, self-interest, and moral commitments. It is a multi-disciplinary field using methods from economics, psychology, sociology, history, and even anthropology. Socio-economics of networks have been studied for over 30 years, but mostly in the context of social networks instead of the underlying communication networks.

Over the past decades, the Internet has grown and evolved to unprecedented size. However, its architecture is still based on the original design principles for an academic network in a “friendly” environment. In addition to the academic usage, the Internet is now used as a business platform and has become a central part of social life.

The overall socio-economic context is an important one, as it can significantly boost or hamper the success of an innovation - issues include the “degree of mobility” in the life-style, the balance of “privacy vs. sharing”, the need for security, the importance ascribed to health, and the distribution of wealth. Important socio-economic aspects include markets of Internet Service Providers (ISPs) and Telecommunication Providers, ISP peering agreements and/or transit contracts, as well as customer usage behaviors and selections of content. A study of all these aspects has to include investigations of regulations for the e-services market and security regulations, as well as the physical environment of e-services in terms of availability - world-wide vs. highly focused (cities) - and dependability for commercial services. This approach will enable to determine (if possible) the economic growth, providers’ revenue maximization, and customers’ benefits.

Socio-economics challenges can be identified in all domains of the Future Internet including the areas of networks, services, and content. As far as the economic challenge faced by all three areas it is worth mentioning that the rules applied for sharing are extremely vital for the healthy operation of the Internet ecosystem and directly affect the value of the network to its users. Such challenges can only be addressed by merging the disciplines of computer science and economics. The key question is: what is wrong with today’s Internet sharing technologies? Are these consistent with economics? More specifically, since TCP is the dominant sharing technology, is TCP sensible from an economic point of view? Is deep packet inspection (DPI) technology good or bad for the Internet community? Which network sharing technologies justify the end-to-end (E2E) paradigm from an economics perspective? What is required to make peer-to-peer (P2P) a blessing instead of a curse? Are there bad applications or just inefficient combinations of sharing technologies and pricing schemes? [1]

Besides the economic dimension, the Internet faces an important social challenge. The current Internet penetration has reached 20% worldwide and should reach 30% by 2015 and 50% by 2020. The Future Internet shall be able to support daily life in developed countries as well as within developing countries. Telecommunication infrastructures must be conceived to guarantee access to the Future Internet also where currently it is poor. As mobile, wireless, optical, and broadband communication infrastructures become even bigger and more interdependent, the number of Web services is expected to grow exponentially in the years to come. These trends lead to a future Internet of billions of services in a network of equals - large enterprises, small and medium enterprises (SMEs), and citizens - in which these services will be indistinctly produced or consumed by “prosumers”.

In this new context *trust* will become a major issue and Web 2.0 technologies are already starting to support trust and reputation within and between computers and humans.

A critical issue in the Future Internet research is the current proliferation of separate efforts due to the various initiatives world-wide. This may on one hand be good for innovation, as it can produce more ideas. However, if initiatives remain separate throughout the development of the Future Internet, many technologically incompatible Internets could emerge. In contrast to the current global Internet, these separate Internets could cause market fragmentation and even social seclusion. To avoid such adverse possibilities, design and implementation of the global Future Internet should proceed with a growing degree of cooperation between initiatives. The mere separation of Future Internet initiatives, if left unchecked, could become a schism leading to many incompatible Future Internets.

B. Challenges of Internet Evolution: Attitude and Technology

Another dimension can be added to the challenges of Internet evolution. This new dimension is the *attitude* towards new technologies. In [2], the author states that according to economic theory, in a competitive environment with limited resources, rational species’ behaviors change so that individuals can maximize a utility function that depends on these resources and on the satisfaction of their needs. The behaviors of rational individuals are constrained by their attitudes. *Attitude* is understood as the disposition, tendency or orientation of mind with regard to a thing, particularly with regard to technology. These attitudes determine the range of behaviors that are valid according to mental disposition. To let a wider range of behaviors be assessed by a rational individual, a change in attitude towards the elements of reality involved in the satisfaction of their needs is needed. There is a limit at which better behaviors cannot be enabled by any new technology, because they cannot be conceived according to existing attitudes. At this limit, the only way to achieve better performance in a competitive environment is a change in attitudes. When the attitudes involved in the satisfaction of needs evolve, new behaviors are considered and consequently new technologies that enable these new behaviors can be developed.

The foundations of this simple co-evolution model are that attitudes demand new technologies to improve the satisfaction of needs and, reciprocally, the new technologies enable the improvements allowed by new attitudes.

Internet is a tool involving technologies that is used by users to satisfy their needs according to their attitudes. This tool has another two characteristics that determine the need to complete this model with two more requirements. These characteristics are: (1) Internet locates at multiple locations both geographically and within society and (2) Internet has no single ownership. According to the two Internet characteristics, two requirements [4] exist for changes on attitudes and technologies to be adopted by Internet. (1) Universality: new technology can be used and new attitudes can be adopted by any user anywhere; and (2) Independency: new technology can be used and new attitudes can be adopted by some users even if others do not use or adopt them, i.e., there is no need to orchestrate change.

The Internet evolution model is completed by two fundamental characteristics: creativity and economic feasibility. The first one must be present whenever a change occurs intentionally. Unless it is a hazardous change, creative minds conceive changes out of their knowledge and experience. The second one is necessary so that evolution is possible, according to economy, in a context of competition for scarce resources, i.e., so that the underlying agents of evolution have the economic capabilities to invest on new technologies and adopt new attitudes obtaining a benefit out of them. Therefore, creativity, economic feasibility, universality and independency are necessary requirements for the co-evolution explained before to be realized. They ensure that changes will occur, that they will be financed, that they will not be reduced to a specific community and that they have the potential to be virally spread through the Internet.

In terms of goals, there are two interests using the Internet: interests of individuals and interests of organizations. Both individuals and organizations use the Internet to satisfy their needs. But in recent times, and particularly during the evolution of the web towards the web 2.0 [5], individuals prove to obtain far more benefits than organizations do. Blogs, wikis, social networks, file sharing, podcasts, online applications and other Internet progresses help individuals to innovate in satisfying their needs more efficiently; accordingly, the success of these new technologies and behaviors coalesces into their increasing number of users [6]. Furthermore, it is widely admitted that these new technologies have co-evolved together with attitudes [5]. Among others, the main components of this shift in the attitude of individuals are: participation, collaboration, confidence and sharing.

This success does not seem to spread likewise through enterprises. Although companies are introduced to web 2.0, not all of them manage to find its benefits. Sometimes, they even abandon these incipient tools, because they do not always turn out to be representing a greater value for their businesses. It is recognized that certain organizations are undeniably profitable thanks to new Internet technologies, but this benefit tends to focus on the tertiary technological sector, rather than pervasively spread through every sector of economy including the primary and secondary sectors. The problem is that societal evolution of new economic sectors fundamentally relies on higher productivity rates in the primary and secondary sectors. As the performance and competitiveness of these organizations are not strongly favored by current Internet trends, thus, the economic capabilities required to ensure the economic feasibility of investments in new Internet technologies and attitudes are limited.

It seems reasonable that a solution consists of promoting innovative paradigm shifts on the attitudes of organizations towards Internet, i. e., to incorporate collaboration, confidence, participation and sharing. However, this paradigm shift does not consist of applying, on an "as is" basis, patterns existing in communities of individuals to enterprises. Competition in business is much fiercer. Collaboration, confidence, participation and sharing must be reconciled with competitiveness. New behaviors have to assess how to create value out of confidence, sharing, participation and collaboration.

If a paradigm shift pervades organizations, performance of every sector of economy can be rewarded by new Internet technologies and can pull Internet technology evolution. Other businesses may appear out of new attitudes and new value can be created for economy with these new attitudes. Unexplored, unexploited and unmerged information might be the clue to follow so that companies can capitalize on new information technologies and gain competitive advantage.

C. Roadmap for Real World Internet Applications - Socio-economic scenarios and design recommendations

In [3], the authors' vision is to realize ambient intelligence in a future network and service environment, and to integrate Wireless Sensor and Actuator Networks (WSAN) efficiently into the Future Internet. Three scenarios are analysed to roadmap some Real World Internet applications. The three phases involve different levels of societal changes, business innovation and technical feasibility. They are not discreet but show a continuous timeline which depends on the context of the actual end use.

1) The first phase - Now - is *evolutionary* from a societal point of view and incremental from technological angle since it is the least integrated: the infrastructure of a mall is used for applications dedicated to the stakeholders of this place.
2) The second phase - New - is more *futuristic* from the socio-economics point of view and innovative from the technology side since it implies the deployment of connections between different and separate areas in the city and it starts to integrate different entities in extension to the shopping mall, e.g. private residential WSAN infrastructures.
3) The third phase - Next - is the most *revolutionary* one from the society point of view because it involves holistic applications of RWI. It proposes a fully horizontal vision of RWI applications with integration of all types of WSAN infrastructures in the city for the provision of an unlimited scope of applications. This is a disruptive vision compared to the existing Internet technology.

A RWI system has to challenge to provide benefits to the user and society through FI applications in key domains such as environment, mobility, safety, professional and industrial activities, citizenship, and ethics.

Today's Internet will change from the distinct network, providing specific services accessible through dedicated terminals, to an Internet dissolved in the artifacts of the physical world accessible via heterogeneous networks enabling users to browse the world as they browse the Internet. The RWI framework should support *horizontal* use and reuse of common WSAN infrastructures to develop a variety of applications. It should not therefore require as many WSANs as applications. RWI system architecture should be *scalable* to enable its functions to evolve in order to meet the future requirements of technology changes and growth. The RWI system must ensure the *continuity* of the services that the user needs with an adequate quality despite the user?s mobility. The RWI framework should *reduce complexity* to enable an easy access of user applications to the sensing and actuation services that are available everywhere. The RWI framework should provide mobile users with a good level of *security and privacy* protection. RWI applications will improve the users? safety in various activities, in particular in the transport, built environment, crisis management and healthcare domains. RWI applications are expected to increase the sense of the community by making perceptible the side effects of individuals? behavior. RWI will support professional and industrial activities. These benefits will be perceptible at short term as shown in the *Now* scenario. RWI system must support new business opportunities and new industrial partnerships by optimizing the integration of sensed and controlled physical phenomena to the Internet. With the integration of a real world dimension to the Internet, privacy and related ethical issues will increase. Even if RWI technology integrate the appropriate mechanisms, privacy and ethics can persist as critical issues and mistrust may slow the adoption despite it enabling an open and secure market space for context-awareness and real world interaction.

D. Concluding Remarks

The three reviewed papers do provide an in-depth view of the non-engineering challenges faced by the Future Internet. However, this discussion is far from being concluded. It seems that a more integrated assessment of such challenges are required and, at the same time, emphasizing the human aspects. Additionally, new dimensions remain to be included in the analysis. The first issue will be addressed in Section III that employs the concepts of the System Engineering, while the second issue will be addressed in Section IV that introduces the *Iconomics* vertices of beings, things and icons.

III. The Future Internet: a system engineering perspective

One of the main difficulties to understand the challenges of the Future Internet is its complexity. The System Engineering perspective provides a means to harmonize the different dimensions that compound the Internet. In this section we show how the System Engineering perspective helps to have a whole view of a such complex problem providing the human being with a protagonist role.

A. Engineering Problems Solving

Descartes's dictum that every problem should be broken down into as many separate simple parts as possible - reductive analysis - is the most successful technique that has ever been used in science. Engineering, as a constructive problem solving science, uses this principle to reduce the problems into as small as possible parts to get to the disciplines assigned to each smaller problem and, based on fundamental phenomena and materials, to cope with the necessary or possible solution. This is the main method to engineering solutions to systems construction.

B. System Engineering

The word system has a subjective nature. It is used to refer to organization forms that are associated to the way that men recognize them; the constructivist view of reality determines that a system does not exist in real world regardless human mind [7].

Systems engineering, differently from others traditional engineering disciplines, does not follow a fundamental phenomenon?s set based on physical properties and relations. Instead, it deals with the necessary knowledge to manage these phenomena, dealing with the system emergent properties, looking for a way to get control about the system entropy [7], [8].

The reductive analysis and the relatively simple construction from the parts became more difficult to deal with as the systems became bigger and bigger.

C. Complexity

The new main problem in the big systems engineering came to be the complexity. Complexity arises when there is a set of characteristics of the system that are not present in any of its parts alone. They are characteristics of the whole or of the co-existence of the parts working together. They are called emerging properties: perhaps the most simple example could be the human body. Life is an emerging property that does not exist in any part aside of the body.

Internet is complex; it is based on a huge amount of subsystems working together that is usually collectively called e-infrastructure. Many layers, not only technical ones, are interacting to fulfill the tasks assigned to them.

D. Socio-technical systems

Information systems are made considering stakeholders (man and social institutions) and technology. It is a socio-technical system, a system in which there is a social infrastructure (man and social institutions) and a technology infrastructure. The consideration of these two infrastructures is crucial in order to identify the correct factors for the quality of services and to identify which are the stakeholders? expectations, to give them the experience that they expect, surprising them whenever it is possible [8], [9], [10].

Considering an e-infrastructure alone, a computer grid, for example, it is only a technological artifact. It has a purpose, a meaning, only when one or more people use it to accomplish some task, as information search or data process to solve problems.

The technological, human and social components of an e-infrastructure system cannot be seen only as the sum of its components. There is a complex interaction among them, with emergent properties.

Another issue that contributes to e-infrastructure system complexity is that many of the systems used today were not developed in an integrated way. They were put together in a gradual way, resulting in a kind of patchwork, with new and old technologies, people and social institutions. New designs must respect this scenario, considering the new, and the old, technologies and several actors (as user, consumers and social institutions). These actors want to optimize their decisions, thinking about their own subsystems, proposes and interests [11].

Big systems engineering had very good answers before and during the WWII but, as the war ended, another new set of problems arose. These new problems came from the dawn of a completely new player in the game: the consumer. What does the consumer wants? In other words: for what is he willing to pay for? What are the requirements?

E. Requirements Engineering

Requirements engineering is engineering discipline alone, crucial in the development of any product or service. This engineering has a life cycle that leads the systems engineer in the process of requirements elicitation, negotiation, documentation and validation of the systems to be developed. The system engineer makes use of this process to execute a task that Kossiakoff and Sweet [12] call concept definition phase, and INCOSE [13] calls concept stage. Both refer to the initial phase of various life cycle models placed by the engineering statements to system information development.

In the requirement process, the elicitation phase concerns itself with people. This requirement gathering process needs to draw upon the knowledge and experience of the organization directors, managers, employees, etc., that are demanding the system. The system engineer needs to talk with people that are demanding the new systems and to the people that will be affected, positively or not, by the system. Usually all these stakeholders are organized in groups, formals or not, with different purposes; such that the whole has no clear purpose and the groups pull in different and often conflicting directions. The elicitation phase is essentially a human activity system that can bring some degree of order to the situation of multiple demands, purposes, issues and problems.

Using appropriate methods to progressively increase order to the requirements gathering process, and achieve a point where specific designs and solutions can be manifested, the system engineering has an approach to achieve the three requirements types that Kano [14], [15], [16] states that must be present on a product or service. This requirements allows the engineering to understand how meeting or exceeding the stakeholders expectation affects satisfaction in the relationship with the system. These requirements types are:

- Normal requirements: these are the requirements that are explicitly required.
- Expected requirements: these requirements are so basic that sometimes the stakeholders may fail to mention them, because they think that it was unnecessary request them explicitly. A system without these requirements is very dissatisfying, but meeting these requirements often goes unnoticed by most stakeholders.
- Exciting requirements: these requirements are the ones that if not present in the system, their absence will not be perceived, will not dissatisfy the stakeholder. As these requirements are not formalized by the stakeholders, i.e., the stakeholders are not apt to voice them, it is the engineer responsibility to explore the problem and opportunities to uncover such unspoken items. For example, as the engineer increases his knowledge about stakeholders needs, he can use his experience to propose features that were not requested but that can improve the system efficiency and effectiveness.

F. Final considerations

Man has personality, hopes, fears, dreams, values and intentions. Do not consider these human dimensions to build systems ultimately dehumanize human-system interaction, and is costly!

IV. The Future Internet: an iconomics perspective

"Iconomics", from a very broad perspective, results from a critical review of the political economy and the macroeconomics of technology transfers and market design aligned with the Center-Periphery system. The evolving actor-network develops and unfolds in the twenty-first century, generating new tools for the creation, management and critique of the information economy as a relatively open and simultaneously global and local network.

Existing technological and economic gaps are compounded by social and cultural differences which are immaterial or intangible, and which are related more closely to the realm of icons than to the requirements of things (hardware, software) and beings (evolving social networks). The iconicity of this evolutionary development is also an index of new metrics for consumption and audiovisual knowledge creation chains. The intangible assets thus produced (real, digital or virtual) are differentially appropriated by individuals, groups and property rights owners (in all classes of assets). The recovery of the world economy depends on this new accountability as much as on the survival of this bank or that company.

However, a precarious regulatory ecology coexists with the global Internet where no one is totally sure with respect to

- what is its economic value?
- what should be the rules governing access, exploitation, intellectual property rights, content distribution?
- how do people represent and communicate values and expectations associated to Internet related actions, projects and technologies?

Conceptually, the challenges of network design and implementation ("social weaving", so to speak, as in assemblages and re-assemblages of actor-networks) are compounded by the simultaneous interaction of space, time and symbol - the playful evolution of this human e-infrastructure corresponds with values, projects and icons for the audiovisual e-superstructural grids.

The emergence of mobile and immersive (audiovisual, virtual and real) applications and infrastructures will expand significantly the uses of the available grids, on the other hand the skills and knowledge required for the adequate creation, production, management, funding and distribution of information-rich devices will be lagging behind at a more than proportionate rate.

The management of audiovisual tools for human and local development engages the storyteller as well as the surveillance manager in the same local neighborhood. Privacy, intimacy, governance and intellectual property issues are at stake. On the other hand, access and use must be weighed by the skills to sustain a balance between supply and demand for information in the long run. However, the regulation of information asymmetries is not only an economic issue as such, it involves control of strategic energetic and telecom infrastructures as well as interference with content production and consumption streams, environmental effects and national identity (iconic) issues.

A. The brazilian case

The Brazilian iconomy has evolved through three stages of digital inclusion frameworks as designed by federal and state level agencies: access, open source and audiovisual, with a growing number of public funding mechanisms as well as articulation with other public policies in areas such as education, science, technology and innovation, culture and telecommunications. But without an overall ICT for development policy, which may be among the explanations for the decline in relative position of Brazil in the ICT Development Index.

A second, more political and institutional issue comes to the fore, given the emphasis on public funding of local content production and recent attempts at reconstruction of State-led broadcasting, social control of communication and regulation of digital TV in Brazil. Scenarios for future audiovisual policies and their impact on local development strategies must be thoroughly discussed, taking into account the limited impact of current policies on income generation and distribution as well as on the creation of sustainable markets for local audiovisual production.

B. Final considerations

Maybe the ideal scenario is that of an emerging "mediapolis", as in Livingstone, where it is "the mediated space where we can communicate, learn about others and take responsibility for one another". As a space where multiple mediated voices talk about the media and its centrality in everyday life. A space where the media and its work in culture, politics, economics, and ethics is critically discussed. A space where scholars, students, producers, and consumers speak the unspeakable and engage with the challenges of a multiply mediated society. A space where the presence of multiple voices in a single discourse is acknowledged and respected. A space where criticism is practiced with the spirit of plurality and hospitality [17].

Silverstone draws on Hannah Arendt and "her deliberations on the notion of republican democracy in the face of totalitarianism, imperialism and of course the threat of mass society. Often unjustly dismissed as a conservative critic, Silverstone seeks to rediscover through Arendt the public art of being with others. In particular Arendt stresses the role of public judgement, responsibility and perhaps above all the human capacity to think as the best shield against political catastrophe. A new global political culture then is not brought about through a McLuhanite technological transformation, but depends upon our shared moral and intellectual capacities. In particular the media?s ability to be able to stretch relations of time and space poses questions related to our civic imagination" [18].

This mediated space or mediapolis is a public sphere open to language patterns such as digital emancipation and other creative expressions of civic intelligence [19].

V. Conclusions and future work

In this paper we focus on the non-engineering challenges of the Future Internet as the most difficult issues to be addressed. After reviewing some recent publications concerning the socio-economic dimensions to be considered in the development of the Future Internet we postulate that the discussion is not over. We extend such discussion by presenting the systems engineering perspective that provides a more integrated approach and puts the human dimension in the center of process. Moreover, we present the *iconomics* perspective that shows that digital inclusion and appropriation of the digital technology by "prosumers" depend on considering a complete new set of values.

As future work we intend to perform an exhaustive study of the Future Internet, to develop a model that includes technological, societal and economical dimensions to produce integrated roadmaps for different technologies, applications, services and businesses.

References

[1] D. Hausheer and et. al, "Future internet socio-economics - challenges and perspectives," in *Towards the Future Internet*. IOS Press, 2009, pp. 1–11.

[2] J. M. Rubina, "Challenges of internet evolution: attitude and technology," in *Towards the Future Internet*. IOS Press, 2009, pp. 12–23.

[3] F. forest, O. Lavoisy, M. Eurich, J. V. Gurp, and D. Wilson, "roadmap for real world internet applications - socioeconomic scenarios and design recommendations," in *Towards the Future Internet*. IOS Press, 2009, pp. 325–334.

[4] S. Ratnasamy, S. Shenker, and S. McCanne, "Towards an evolvable internet architecture," in *Proceedings of the 2005 Conference on Applications, technologies, architectures, and protocols for computer communications*. ACM, 2005, pp. 313–324.

[5] T. O'Reilly, "What is web 2.0," Available in http://oreilly.com/web2/archive/what-is-web-20.html, 2005, retrieved in may, 2010.

[6] "Top 15 most popular web 2.0 websites," Available in http://www.ebizmba.com/articles/web-2.0-websites, 2010, retrieved in may, 2010.

[7] L. Skyttner, *General systems theory: Problems, Perspectives, Practice*. World Scientific Publishing Company, 2nd edition, New Jersey, 2005.

[8] D. K. Hitchins, *Systems Engineering: A 21st Century Systems Methodology*. John Wiley and Sons, Chichester, 2008.

[9] M. Ottens, M. Franssen, P. Kroes, and I. V. D. Poel, "Modeling infrastructures as socio-technical systems," *International Journal of Critical Infrastructures*, vol. 2, no. 2-3, 2006.

[10] V. Bryl, P. Giorgini, and J. Mylopoulos, "Designing socio-technical systems: from stakeholder goals to social networks," *Requirements Engineering*, vol. 14, no. 1, 2009.

[11] M. Houwing, P. W. Heijnen, and I. Bouwmans, "Socio-technical complexity in energy infrastructures - conceptual framework to study the impact of domestic level energy generation, storage and exchange," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*. IEEE, 2006.

[12] A. Kossiakoff and W. N. Sweet, *Systems Engineering Principles and Practice*. John Wiley and Sons, New Jersey, 2002.

[13] "Systems engineering handbook, version 3," INCOSE-TP-2003-002-03 (2006).

[14] G. H. Mazur and A. Jurassic, "Jurassic QFD," in *Transactions of the 11th Symposium on Quality Function Deployment*. QFD Institute, Michigan, 1999.

[15] G. H. Watson, T. Conti, and Y. Kondo, *Quality into the 21st Century: Perspectives on Quality and Competitiveness for Sustained Performance*. ASQ Quality Press, Milwaukee, 2003.

[16] "The kano model," Available in http://kanomodel.com/.

[17] "Mediapolis by levingstone," Available in http://community.livejournal.com/mediapolis.

[18] N. Stevenson, "Roger Silverstone: an intellectual appreciation," *European Journal of Cultural Studies*, vol. 10, no. 4, 2007.

[19] G. Schwartz, "Digital emancipation," *Public Sphere Project*, 2008.

The Engineering Virtual Enterprise: A Framework for Soft and Entrepreneurial Skills Education

Edgar E. Troudt, Christoph Winkler, A. Babette Audant, and Stuart Schulman
Institute for Virtual Enterprise, CUNY Kingsborough Community College
Brooklyn, NY 11235, USA

ABSTRACT

This work seeks to tackle common workplace deficiencies present in associates-level, Pre-Engineering students, as well as cultivate entrepreneurial spirit and ability. These deficiencies cause attrition in the number of Engineering graduates, and have been identified by employers as being important to the modern engineering workforce [9]. Our proposed platform, the STEM-based Virtual Enterprise (VE), is a business simulation program in which students start and run a virtual firm within their classroom. The program is currently being successfully used in other STEM areas, notably BioTechnology and Information Technology[1].

The STEM-VE program, a combination of in-class pedagogy, software tools, and an international network of participants, is designed to develop entrepreneurial competencies, interpersonal skills (e.g., teamwork, effective communication in working with those in different roles and positions), critical and analytic thinking and problem solving. Each student assumes a position within the firm and carries out the responsibilities of his or her department.

This paper specifically lays out the framework for two Engineering Virtual Enterprise (***ve****eng*) engagements that campuses could adopt to provide an encapsulating experience for Pre-Engineering majors.

[1] The previous work described is based on funding from the National Science Foundation (under DUE-0501711 [14] and DUE-0802365 [13]). Any opinions, findings and conclusions or recomendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).

1. BACKGROUND

The American Recovery and Reinvestment Act of 2009 supports innovative solutions to pressing national problems in science, health, the environment, education, public safety, and other critical areas, as a way of stimulating much-needed economic recovery and jobs creation. The Act, which provides tax breaks and increases funding for research and science, is based on the belief that entrepreneurs can innovate America's way back to prosperity [1]. Last December, the President proposed another round of economic stimulus measures that target job creation and support small businesses by providing loans, tax credits, among other measures [3].

In the engineering fields, the importance of entrepreneurial innovation is recognized: "Engineering is a key component of innovation and our technological society. Changes on a global scale are rapidly occurring for engineering, and Federal leadership is needed to respond quickly and informatively" [11]. "Engineering lies at the interface between science on the one hand and society on the other" [6]. Engineers apply scientific and mathematical principles towards designing practical solutions. While engineering education has traditionally favored STEM competencies, "the constraints on engineering problem-solving today are increasingly not technical, but rather lie on the societal and human side of engineering practices" [6]. In a Special Issue of *Journal of Engineering Education*, Sheppard, Pellegrino, and Olds [15] write, "[i]t would be naive to treat technical and non-technical challenges and opportunities as separable". The boundaries are "increasingly blurred" and engineers are being called on to design solutions for our increasingly complex world; preparing them to meet these challenges and opportunities will require designing and delivering re-tooled engineering courses and programs. Active learning strategies like Virtual Enterprise teach and encourage students, to be creative, experimental, and entrepreneurial preparing them to overcome these constraints.

There is wide recognition of the need to challenge "traditional" ways of teaching engineering and related STEM fields. Researchers have found that the creation and perpetuation of "impersonal" and hierarchical classroom environments discouraged collaborative learning. Using a strategy/intervention that "alters traditional patterns of interaction," encourages communication and teamwork "without compromising the quality of learning" [2]. VE would fall within this category of interventions. Going further, Turner [17] concludes that non-technical, or soft skills are not developed in isolation from technical skills. Rather, the development of hard and soft skills interacts in significant ways, with each process strengthened and informed by the other.

The rationale for the Engineering Virtual Enterprise program draws upon studies of project-based learning and curricula that infuse entrepreneurship into engineering and technician education courses, supporting and strengthening the development of soft skills, as well as improving student recruitment and retention: Tubaishat [16] demonstrated that including topics such as project management and product development in a Computer Science course helped students deal with non-technical aspects of project development, strengthening their communication and other soft skills. Problem-based learning, much like that of VE, was successfully integrated into an engineering program where students worked in teams to design a product that met pre-determined criteria. Students developed familiarity with the design process, and with the importance of working within a strict time frame as they would in a real-world setting [5]. Integrating entrepreneurship and related project management skills into engineering curricula enhances student profiles as integral participants in the knowledge-based economy [12]. The University of Edinburgh developed a series of interdisciplinary courses to improve students' ability to work across disciplines and "improve their preparation for industry [7]. By focusing efforts on a single project, student teams were able to see the connections between "engineering design and economic viability". Students' interest in pursuing careers related to this engineering project was also enhanced. Dabbagh and Menascé [4] examined students' overall perception of the engineering profession in a first year course that introduced entrepreneurial concepts using a project-based model. In the course, student teams formed IT companies that competed to develop business support software in a simulated market environment. Compared to students in a "traditional" course, students in the projects-based course had "significantly improved" perceptions of engineering professions and of entrepreneurial opportunities in engineering. Researchers suggest that infusing engineering curricula with entrepreneurship and projects-based learning can increase student recruitment and retention [4].

2. PLATFORM: VIRTUAL ENTERPRISE

Virtual Enterprise (VE) contextualizes disciplinary content by having students develop realistic, commercially viable projects in the classroom. The lines between disciplines are blurred as students take a holistic view of their firm and its tasks; soft skills are integral. The authors, principals of the Institute for Virtual Enterprise (IVE) at the City University of New York (CUNY) have nine years of experience in the design and administration of such programs – both academic and non-credit – that help students develop the soft and entrepreneurial skills essential to success in the workplace. CUNY-IVE operates programs that reach CUNY's twenty-three campuses, which serve about 450,000 students.

VE has teams of students act entrepreneurially as they design and operate simulated firms in the classroom. The simulation is a combination of **in-class, active-learning pedagogy**, a virtual economy (the **IVE MarketMaker**) with additional software tools, and an international community of simulated student-run firms (the **IVE Partner Network**). Through VE, students acquire soft and entrepreneurial skills while making concrete use of the content of their academic majors. Soft skills are the interpersonal skills (such as teamwork, effective communication in working with those in different roles and positions), critical and analytic thinking skills, and problem solving skills that are part of the League for Innovation's 21st Century Learning Outcomes [18]. Each student takes on a position within the firm and carries out the responsibilities of her or his department (such as, Research and

Development: establishing timelines and ethics protocol; Purchasing: setting up inventory procedures). Students have a blended learning experience in which they participate in face-to-face activities both within the simulation (such as, taking part in virtual firm staff meetings, writing up business plans, developing surveys for potential customers) and outside the simulation (such as, discussing articles about industry issues, exploring various organizational models).

The STEM-VE experience, which customizes the simulation for students in STEM majors, has proven to be a powerful method for instilling a passion for the underlying STEM discipline and a sense of entrepreneurial efficacy. Results from NSF funding by the Advanced Technological Education Program [13,14] where the STEM-VE curriculum was infused with IT content, produced a high level of student interest and engagement. Especially interesting is one such study performed with students from populations typically underrepresented in IT. This study found these entry-level students reporting a great shift in attitude toward the discipline, and aspiring to study and work in the field [8].

3. THE VIRTUAL ENTERPRISES FOR ENGINEERING *(VE^{ENG} CAREERS* AND *VE^{ENG} PROJECTS)*

This framework proposes two engagements as "bookends" on an associates-level, pre-engineering major as follows:

***ve^{eng}* Careers**

A non-credit, pre-semester institute designed to help students entering the pre-engineering major understand the breadth of engineering careers, the many fields and types of organizations in which engineers work, and the variety of technically and socially important problems they help to solve. This approach mirrors recommendations made by the National Academy of Engineering: "students who are introduced to engineering design, engineering problem solving, and the concept of engineering as a servant of society early in their undergraduate education are more likely to pursue their engineering programs to completion. The same approach is also more appealing to women and underrepresented minority students" [9,11]. Students explore the match between engineering careers and their own strengths, interests, and values. The engagement will have small groups of students staff and simulate several types of engineering firms, identifying and providing rudimentary solutions to relevant industry problems. The aim is to strengthen students' interest in engineering and motivate them to persist in engineering degree programs and ultimately in engineering careers.

***ve^{eng}* Projects**

A three-credit course in which pre-engineering majors near the end of their associates degree operate a simulated, student-run engineering firm. This course gives students opportunities to apply and practice their engineering and applied math knowledge while participating in activities that promote the development of entrepreneurship and workplace competencies. Student participants will act as entrepreneurs taking the firm from its business plan to engineering project conceptualization through planning for deployment. This experience is expected to enhance participants' (1) understanding of the types of problems that engineers help to solve, (2) understanding of the process of starting an engineering venture, and (3) entrepreneurial and soft skills.

Focal Projects

The National Academy of Engineering has organized the state of engineering problems into fourteen categories as part of EngineeringChallenges.org [10]. Of these categories, seven have both sufficient ongoing examples within the Engineering community, sufficient literature, and are a level that is accessible for community college pre-engineering students. These categories will form the topical foci of the two ***ve^{eng}*** engagements *(**ve^{eng} Careers*** and ***ve^{eng} Projects)***:

1. Provide access to clean water.
2. Restore and improve urban infrastructure.
3. Advance health informatics.
4. Engineer better medicines.
5. Prevent nuclear terror.
6. Secure cyberspace.
7. Advance personalized learning.

Both the ***ve^{eng} Careers*** and the ***ve^{eng} Projects*** will draw their choices of engineering sub-disciplines and projects from this list. The engineering instructors responsible for teaching the courses will elaborate on facts about the careers involved in these areas by giving students access to a series of online resources (e.g. links, career websites, and transfer institutions).

Expected Results
The authors hypothesize that the net result of these engagements will be:

1. student participants in the ***ve**^eng^* ***Careers*** Institute will be better aware of the breadth of engineering projects and careers, and will plan choices of future study and careers accordingly.
2. the ***ve^{eng} Careers*** Institute will improve retention and recruitment in the pre-engineering major, as demonstrated by a significant change in attitude towards the engineering disciplines.
3. project-based learning experiences, as offered in ***ve^{eng} Projects***, will instill demonstrable improvement in problem solving skills, soft skills and entrepreneurship self-efficacy.

4. VE^{ENG} CAREERS

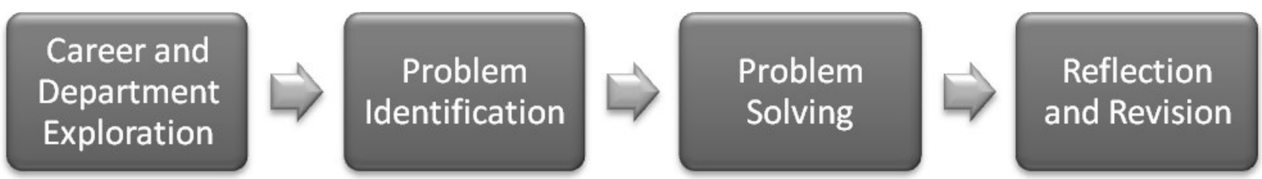


Figure 1: The flow of the VE-eng Careers engagement.

The ***ve^{eng} Careers*** (*non-credit, pre-semester institute*) will help students entering Pre-Engineering majors understand the breadth of engineering careers, the many fields and types of organizations in which engineers work, and the variety of technically and socially important problems they help to solve. The engagement will initially follow this general plan (also see Figure 1):

Phase I: Career and Department Exploration
During this phase the organizational structure for an engineering firm is established. Students research the staff positions. The instructor can facilitate this process by using a sample organizational chart from a partner or fictitious firm and by directing students to career research websites. This process allows the students to explore the breadth of careers in the industry, identify particular career tracks, and make informed decisions about their choices within the field.

Phase II: Problem Identification
During Phase 2, the firm's clients will identify potential problems/issues/needs and supporting information (such as existing solutions and users' ancedotal demands) will be gathered for the problem solving process in Phase 3. The demands of management and the needs and capabilities of the eventual users should be uncovered during this phase. As part of this process, students will familiarize themselves with the breadth of products and services that are already provided by other firms. Students will conduct a needs assessment for the company and agree on a priority project in conjunction with the company's management.

Phase III: Problem Solving
The firm devises a possible solution to the problem identified in the previous phase and presents it to the management. This solution will likely be simplistic and may involve bringing together several existing solutions from other firms. Members of the management may be simulated by faculty colleagues, corporate partners, or other members of the international VE network. Students will prepare and deliver technical whitepapers on their solutions plus presentations geared towards a non-technical audience.

Phase IV: Reflection and Revision
Based on feedback provided in the Problem Solving phase, students will revise their product descriptions. Rudimentary SWOT (strengths, weaknesses, opportunities, and threats) reports will be prepared about the firm and its proposed solution(s).

5. VE^{ENG} PROJECTS

The ***ve^{eng} Projects*** (*3-credit, 45 contact hours*) course will have students act as entrepreneurs in the classroom in a simulated engineering firm from business plan development, to engineering project conceptualization, and through planning for deployment. Students will leave the course understanding how entrepreneurs operate and how "***intra***preneurs" start novel ventures within a larger enterprise. Much of the development work will involve the project team developing numerous projects within the focal areas. The operation of the course will follow this general pattern:

1. Review a sampling of the breadth of engineering companies and products.
2. Select a niche within engineering; create a corporate identity.
3. Research and design the firm's management structure including departments and officers. (Typical departments for the simulation include: marketing, finance, human resources, and R&D.)
4. Staff positions within the hierarchy (organizational chart).
5. Design a flagship product or service; produce a demonstration product.
6. Conduct market research (on VE students or corporate partners); or in the case of responding to an outside firm's need, present the limited version of the product for feedback.
7. Solicit funding by developing the enterprise's business plan, including identifying target consumers and marketing strategy.
8. Implement product, starting with implementation plans and Gantt charts, and ending with proper technical documentation.
9. Develop promotional materials, including analysis specifying quantities such as TCO and ROI, if applicable; develop website and presentation for non-technical audiences.
10. Presentation of product(s) to prospective consumers and/or integrators.

6. CONCLUSIONS

The framework described here extends the highly effective STEM-based VE program, which evaluation research [8,13,14] has shown to be extremely successful in being adapted to the Information Technology (IT) and BioTechnology disciplines.

This engagement adds to the possible modalities for engineering entrepreneurship education at American institutions. This would meet a national priority for redeveloping the workforce both with 21st Century communication skills and entrepreneurial spirit, as outlined in the American Recovery and Reinvestment Act (ARRA) and other educational initiatives of President Barack Obama.

While the described scope of this work is focused on offerings at Community Colleges, the potential for broader impact on the national discussion about engineering curriculum reform is great. If offered both at the Community and Senior College levels, it could serve as an articulation bridge between these institutions. Further, the ***ve^{eng}*-*Careers*** construct can be used to help students at Community Colleges explore the career potentials opened by transferring to baccalaureate programs at local senior colleges.

7. REFERENCES

[1] ARRA. (2009). American Recovery and Reinvestment Act. Retrieved from: http://www.recovery.gov.

[2] Barker, L., & Garvin-Doxas, K. (2004). Making Visible the Behaviors that Influence Learning Environment: A Qualitative Exploration of Computer Science Classrooms. Computer Science Education, 14(2), 119-145.

[3] Calmes, J. (2009, December 8). Obama Offers Plan to Small Businesses. The New York Times. Retrieved from: http://www.nytimes.com.

[4] Dabbagh, N., & Menascé, D. (2006). Student Perceptions of Engineering Entrepreneurship: An Exploratory Study. Journal of Engineering Education, 95 (2), 153-164.

[5] Denayer, I., Thaels, K., Sloten, J., & Gobin, R. (2003). Teaching a structured approach to the design process for undergraduate engineering students by problem-based education. Journal of Engineering Education, 28(2), 203-215.

[6] Grimson, J. (2002). Re-engineering the curriculum for the 21st century. *European Journal of Engineering Education*, *27*(1), 31-37.

[7] Harrison, G., Macpherson, E., & Williams, D. (2007). Promoting interdisciplinarity in engineering teaching. European Journal of Engineering Education. 32(3), 285-293.

[8] Morgulas, S. (2007). Virtual Technology Information Enterprises (VEIT): An Integrated Vehicle for Technology Education Reform: Final Evaluation Report. Center for Advanced Study in Education.

[9] National Academy of Engineering. (2005). Educating the Engineer of 2020: Adapting Engineering Education to the New Century. Retrieved from: http://books.nap.edu/catalog.php?record_id=11338#toc.

[10] National Academy of Engineering. (2008). Grand Challenges of Engineering. Retrieved from: http://www.engineeringchallenges.org/cms/challenges.aspx.

[11] National Science Board. (2007). Moving Forward to Improving Engineering Education. Retrieved from: http://www.nsf.gov/pubs/2007/nsb07122/nsb07122_2.pdf.

[12] Papayannakis, L., Kastelli, I., Damigos, D., & Mavotas, G. (2008). Fostering entrepreneurship education in engineering curricula in Greece. Experience and Challenges for a Technical University. European Journal of Engineering Education. 33(2), 199-210.

[13] Schulman, S., & Troudt, E. (2008). Enhancing Soft and Entrepreneurial Skills Training for Two-Year College Technicians Using a Contextualized Business Simulation Program. (NSF ATE grant: DUE-0802365, $749,217).

[14] Schulman, S., & Deutsch, J. (2005). Virtual Information Technology Enterprises (VEIT): An Integrated Vehicle for Technology Reform. (NSF ATE grant: DUE-0501711, $149,990).

[15] Sheppard, S., Pellegrino, J., & Olds, B. (2008). On Becoming a 21st Century Engineer. *Journal of Engineering Education, Special Issue: Educating Future Engineers: Who, What, and How.* 97(3) 231-234.

[16] Tubaishat, A. (2009). IT Systems Development: And IS Curricula Course that Combines Best Practices of Project Management and Software Engineering. Issues in Informing Science & Information Technology. 6, 257-268.

[17] Turner, R. (2004). Towards a Structural Model Connecting Hard Skills, Soft Skills and Job Conditions and the IS Professional: The Student Perspective. Issues in Informing Science & Information Technology. 1, 977-991.

[18] Wilson, C. D., Miles, C. L., Baker, R. L., & Schoenberger, R. L. (2000). 21st Century Learning Outcomes: New Competencies & Tools for Community Colleges. League for Innovation in the Community Colleges. The Pew Charitable Trust. (ERIC Document Reproduction Service No. ED439751).

Refocussing Engineering Design for a Sustainable Living Environment

Larry Arno VINT
Design Department, Griffith University
South Bank, Queensland 4101, Australia

ABSTRACT

Designers are considered as an integral part of the creative industry; however, in reality they represent the consequence business. Designing involves problem solving and improving people's lives; thereby, what designers create with every decision has an environmental price. The dilemma many designers are facing is seeking an internalised balance between readiness to make informative decisions to incorporate sustainable practices and constraints of being engaged in a commercial project, run by profit. Thus, ethical design with its degree of sustainability forces designers to make a choice.

Keywords: Design, Engineering, Sustainable Design, Climate Change, Global Warming and Education

1. INTRODUCTION

Design engineers, whether disciplined in electrical, mechanical, civil or architectural are considered as an integral part of the creative industry, however, in reality, they represent the consequence business. Whether designing a particular new product, system, or inner workings of a design; from the preliminary development stage to designing the most critical parts; the synthesis of combining different ideas, influences, and/or objects in the design stages are paramount. Engineering design involves problem solving and improving people's lives; thereby what engineers create with every decision, can be multiplied by thousands and often millions through mass production. Conversely, every produced item has an environmental price. Chochinov [2] in his manifesto for sustainability in design wrote "we are suffocating, drowning, and poisoning ourselves with the stuff we produce, abrading, out-gassing, and seeping into the air, water, land, food … and designers are feeding this cycle, helping to turn everyone and everything into either a consumer or a consumable". Engineers need to understand the role and impact manufactured products and built environments have on the world. The engineering process is part of a set of tools for solving problems and improving life, not only in the short term, but for generations that lay ahead. Moreover, the design is a means, not an end.

In order to develop sustainable solutions an engineer has to transform concepts, considering the impact of the produced design. The impact is much greater than just interaction with the intended design's consumer; as the consequence of the design production and usage has an impact on human kind globally, with ramifications to people, the environment and the economy. Moreover, the design's consequence not only has influence in the present sense, but also has implications for generations to come. The materials used to make the design, the resources required to manufacture, package and sell it, the quantity, quality, and longevity of the product, and whether it should have been designed at all in the first place, have a major bearing on sustainable design.

Design engineers are often implicated in today's environmental crisis due to their active involvement in promoting a culture of oversupplying the market with unnecessary products, over-engineering and encouraging mass material consumption [7]. Detailed in Papanek's [21] book, Design for the Real World, the designer has "become one of the most harmful of professions". Yang & Giard [34] state that the design profession is both the problem and the solution, for which design engineer practitioners and students must understand the ecological impact of their profession. Findeli [5] writes that without a responsible designer, you will not produce responsible design. It has become imperative that sustainability is fostered within the profession and taught in the engineering/design curriculum at both school and university level.

2. CLIMATE CHANGE

Climate change is one of the greatest challenges facing our generation [23]. The design engineers and politicians of our time will be judged by future generations on their ability to rise to this challenge. The latest research shows that climate change will damage all economic, social and environmental aspects of life [28]. The Prime Minister of Tuvalu, Telemi [27] stated "Never, in the history of humankind have we faced such a global challenge. We, leaders, must do this *(tackle climate change)* for our children and our children's children". Fry [11] wrote, "I would actually say that we're at a watershed, the future of humanity as we understand it is really before a choice which says do we change direction, or do we try to maintain what we already have? … The challenge that we have now is to deal with the world that we've created, and sustainment in that sense is both a kind of a process and a project that is about that exercise, of dealing with the world that we've got, creating a future. … it's about creating another kind of direction, other kinds of ways of living, other kinds of economies, recognising that we're in a very dangerous situation, and that to be sustainable, we have to be able to eliminate conflict as well as damage to the environment."

Organisations including the United Nations [31][32], UNESCO [29][30], International Association of Universities [15], UK Government [13] and internationally recognised climate change theorists Clark [3], Fry [8] Laszlo [17], Rebelo [24] indicate that climate change is a design problem, and through sustainable education humans can address one of the key priorities of the 21st Century. Ecological architect Van der Ryn, an acknowledged researcher, theorist, educator and a leader in sustainable architecture, wrote:
"In many ways, the environmental crisis is a design crisis. It is a consequence of how things are made, buildings are constructed, and landscapes are used. Design manifests culture, and culture rests firmly on the foundation of what we believe to be true about the world. Our present forms of agriculture, architecture,

engineering, and industry are derived from design epistemologies incompatible with nature's own" [33].

A statement in September 2009 from the Chairman of the Intergovernmental Panel on Climate Change (IPCC) Dr Pachauri [20], on the UN Summit on Climate Change, Copenhagen, quoted findings from IPCC's Fourth Assessment Report (AR4); collective research from four thousand specialists over a 5 year period; that if no action is undertaken to stabilize the concentration of greenhouse gases in the atmosphere, then the average temperature by the end of this century would increase anywhere from 1.1 degrees to 6.4 degrees C. Figure 1 illustrates the IPCC average temperature forecast taken in 2007.

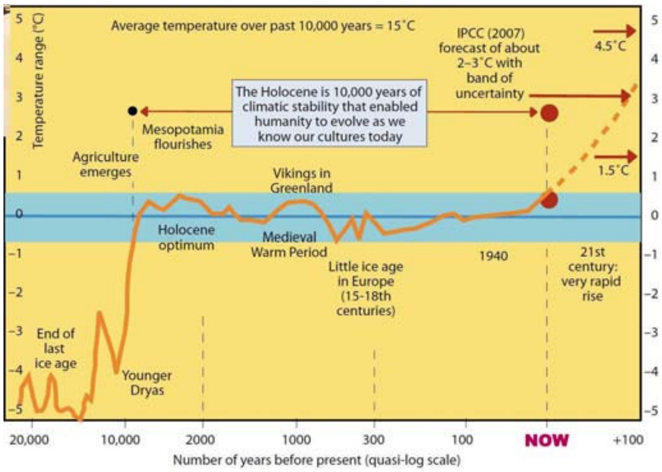


Figure 1. The average temperature by the end of this century. Source: Robert Corell, Heinz Centre. (After: Safe Climate Australia Prospectus, July 2009, Page 11)

The depletion of natural resources is faster than what nature or humans can replenish; sea levels are rising as a result from the melting ice caps and warming oceans; extremes in temperatures are causing increased precipitation, tropical storms and cyclones; pollutants are found widely in waterways, oceans, soil and the air due to non-sustainable manufacturing and agricultural processes, greenhouse gas emissions and overproduction [16]. In the absence of not taking sustainable action, the likely hood will lead to:

- Economies will falter. The success of many national economies are closely linked to their natural resources [26];
- Possible disappearance of sea ice by the latter part of the 21st century;
- Experts estimate that climate change will force millions of people to abandon their homes over the next fifty years, due to increased floods, fire, drought, and deadly heat waves;
- If sea levels were to rise by 1 metre, it is believed the majority of land would be underwater in Bangladesh (population 162 million), Sri Lanka (20.2 million), Tuvalu, Nauru, Antarctic Peninsula, Maldives, Singapore, Carabian States, Papua New Guinea Islands, Micronesia, Kiribati, Indonesia, Samoa and Egypt. Many other countries will have the problem of having their fresh water supplies contaminated with salty water. According to the United States Environmental Protection Agency (EPA) [4] sea levels will continue to rise for several centuries, even if global temperatures were to stop rising by 2020;
- Increase in frequency of hot extremes, heat waves and heavy precipitation;
- Increase in tropical cyclone intensity;
- Decrease in water resources due to climate change in many semi-arid areas, such as the Mediterranean Basin, western United States, southern Africa and north-eastern Brazil;
- Possible elimination of the Greenland ice sheet and a resulting contribution to sea level rise of about 7 metres. Without mitigation future temperatures in Greenland would compare with levels estimated for 125,000 years ago when palaeoclimate information suggests 4 to 6 m of sea level rise;
- Approximately 20 to 30% of species assessed so far are likely to be at increased risk of extinction if increases in global average warming exceed 1.5 to 2.5 degrees Celsius.

Science leaves designers no space for inaction now. Designers need to urgently act and make sustainable reforms within their own design fields.

Research conducted by the Emission Database for Global Atmospheric Research 2000 project provides a snapshot of global annual greenhouse gas emissions. Industrial processes 16.8%, residential, commercial, and other sources 10.3%, transportation fuels 14%, waste disposal and treatment 3.4%; see Figure 2. These values provide a snapshot of global annual greenhouse gas emissions in the year 2000.

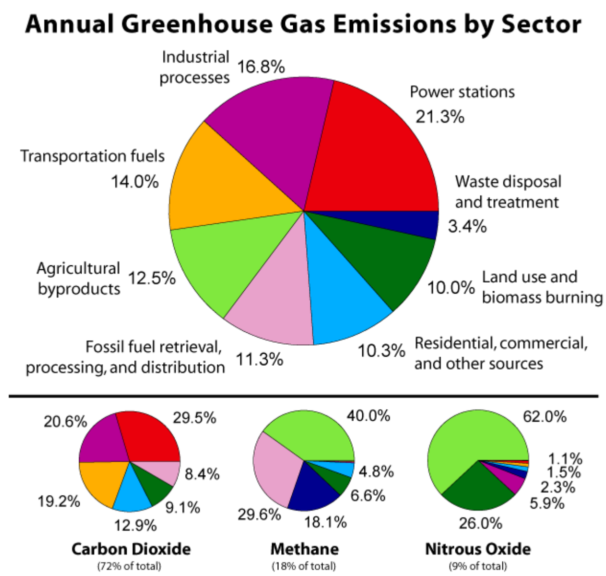


Figure 2. Greenhouse Gas by Sector. Source: "Image created by Robert A. Rohde / Global Warming Art"

3. DESIGN SUSTAINABILITY

Design sustainability is putting into place a process of transformative change to an agreed sense of direction to respond to the circumstances design engineers find themselves in. These directions are coming from the engineering design sector on a global scale [15]. It is through bad design decisions and their consequences that people's environmental, cultural, economic and social futures are being severely changed or taken away. Through design, people are becoming unsustainable. Design engineers can no longer be ignorant to the ramifications of their own design decisions. Most designed products directly or indirectly are leading people to unsustainable lifestyles. It is crucial for design engineers to learn how to design for a sustainable future.

3.1 Design domain

Design responds to the world in which we live addressing three particular kinds of ecologies (biophysical, social and ecology of mind) [12] that have all been damaged through climate change and all have a relation to design both in terms of how these ecologies became distorted and what needs to be addressed to achieve a viable future [6]. There are two important points to this: first of all, climate change is just one of the problems, and it is a problem which generates other problems, so the problem that humans face is more serious than just the implications of climate change. To put this into context, the United Nations November 2009 monitoring on emissions reduction indicated that the earths warming is progressing and greater than anticipated [32]. Using the 1990 levels as a benchmark, the temperature has increased by 41 percent. The global warming spectrum is arranged from 1.8 to 7 degrees. At the moment the earth is looking at a 7 degrees temperature increase. Therefore, the way governments have been talking about working towards an emissions reduction to 5% is too low to what is required. To add to this, the speed of government action to reduce climate change has been incredibly slow in terms of the pace of the problem. The speed and acceleration of the problem on one hand and the very slow response on the other is another way of characterising the problem. Therefore there are two issues, the inherent or intrinsic problem created and the problem of responding to that situation in an inferior or inadequate way.

3.2 Biophysical ecology

The first ecology is a familiar one, what people understand to be biophysical ecology [6][8], but what has happened to a significant part by design is that humans have actually made the distinction between the natural and artificial impossible to distinguish. For example, not even the water we consume is natural; with additives such as fluoride, chlorine and antibacterial chemicals. People are dependant on the artificial as it has become indivisible from the natural. This has become one of the reasons why design has to be on an increasingly significant steep trajectory of importance in relation to the situation that people are in. But it is more than just physical understanding of the complexity of the one or the other, it is also a perceptual indivisibility between the natural and artificial; in other words people do not see the world naturally. People only see the world from what people have learnt artificially. The relation between what it assists empirically and what it assists perceptually has become a significant part of the problem.

3.3 Social ecology

The second ecology that people live in and exist by is the social ecology. The fact that people only exist by virtue of each other and have no ability to exist as simply independent identities, so that the social and the notion of community, has a direct correlation to ones ability to flourish and survive. There is a relation between one and the other, so that as the biophysical ecology becomes critical there is a very strong chance that the social ecology also becomes critical. Again, going back to climate change to illustrate the fact that by the end of the century it is quite possible that about 10% of the world's population will be displaced under the term climate refugees. To put this in perspective, tens of millions of people will become refugees. These people will not be completing forms to migrate to different countries. People are going to move, and travel where they can in the circumstances they find themselves in. The notion of border protection and immigration procedures will be completely gone, but replaced by an enormous amount of social disruption and dysfunction. The Australian Defence White Paper published May 5, 2009 [1] indicated that there is an expectation that this problem will arrive in this country from the North. The government's short term solution is to increase the size of the Australian Navy and to start to deploy more troops in the North of the country as prevention. With the large number of water craft expected to arrive, the government's solution does not seem probable. From this example, one can see the way the problem is unfolding and is at odds with the ability of people to consider the problem and address it.

3.4 Ecology of mind

The last ecology, ecology of mind relates to design education. People exist in a particular kind of way of thinking and that particular way of thinking has a direct relation to their way of seeing. "… people see with their eyes (physically) as well as with their minds (interpretively) the existing reality (perceptively), and some see requirements and (with foresight) possible solutions for the future …" (author unknown). Eyes are simply instruments that facilitate the capability of sight, but people see the result of what it is they know. If what people know is to how to act in the world in a destructive way, then that way of thinking determines what it is they do. So a great deal of the problems designers have is that people wrongly continue to think the world as a place of infinite resources; when in fact the resources are limited. People think in terms of being, enduring and continuing as a species when in fact they are a finite species, and the more that they mistreat the conditions that they depend upon, the shorter the amount of time that they have. Thereby, the biophysical, the social and the ecology of mind are all inseparable. If you pull them apart they are an explanation of the design domain, and it is not the way you would normally understand design.

4. CONSEQUENCES

It is necessary for designers to be able to confront and address design issues with an awareness of time. To assist in the understanding and implications of timeframes, the life of CO_2 in the atmosphere by general estimates lasts for 200 years, so no matter what humans do they will have this predicament for the next 200 years. CO_2 has always been in the atmosphere, it is the only way that plant life on land can obtain its carbon to allow it to grow. The industrial revolution had really not yet ended. But since the start of the industrial revolution, the proportion of CO_2 in Earth's atmosphere has increased dramatically, and continues to increase whilst the absorbing plant-life is depleted, and that is the important part of this real problem.

The way in which the temperature of the world is regulated is through a sub thermostatic process. The thermostat being the earth's deep oceans, and terms of temperature adjustment, the deep ocean takes approximately 200 - 250 years to change a degree or two. Scientists have acknowledged that the sea levels rises are projected to keep occurring from between 300 to 400 years. Therefore, the problems will not be solved in the near future. The situation people are in at the moment is that between

1.8 degrees and 2 degrees of warming is going to occur no matter what people do. This has already been determined by the damage that has already been done. If people carry on doing what they are doing now, then there is a possibility of having a rise of 7 degrees. If the earth's temperature was to rise by 3 degrees, Australia would loose its Great Barrier Reef, species of flora and fauna including some eucalyptus species and have many coastal properties flood. The world would change dramatically. By the end of this century, much of this planet as people know it would be unrecognisable. Some areas on the earth will actually be more liveable than what they are at the moment, but many more places will be dramatically worse.

Designers have a very simple choice, they can keep on designing as it contributes to the problems, many of which have arrived by design or they can try to design by another way. In simplified terms designers can either be part of the solution or part of the problem; the decision comes down to choice. One can be paralysed by the choice, or be stimulated and motivated; and even to a degree become excited by the challenge. Designers need to change direction. Included also are the people who have an impact on the design outcome, these are clients, managers, project partners, vendors, project leaders and lead designers, who can influence the final outcome of the project design.

5. DESIGN ENGINEERING SOLUTION

Designer engineers need to deal with the world they have already created. The design solution needs to balance with its problem and its effects on the environment and ultimately everyone. For example, people do not need a battery powered scooper to pick up dog droppings, and they do not need cars that achieve 17 miles or less per gallon (17 mpg is the average gas mileage of the average car in the United States, EPA 2009) [4]. Sustainable architecture is positive for the environment, but many companies do not deal with buildings and/or cities that already exist. It is imperative to be able to deal with what already exists, which does not mean you never design anything new, but it means that quantitatively you design in relation to what is the greater problem. The world's biggest retailer Wal Mart changed to sustainable packaging of its products through re-design. In 2007 the company identified $10 billion in savings from packaging efficiencies through making new sustainability decisions within the first two years of making the change [25]. The sustainability changes involved reducing the packaging waste by 5%. There were many factors associated with making the packaging more sustainable, including greenhouse gas reductions during manufacture of the package, substrate material choices and chemical composition, eliminating PVC from their private-brand packaging, and integrating recycled materials into new products. The packaging reduction on the entire packaging supply chain was designed through a "cradle-to-gate" approach.

In Figure 3 designer Lotersztain illustrates his belief that sustainable design has no boundaries. Refurbished second hand yacht fenders, primarily used to protect boats during mooring have been transformed into a sofa. Its design is functional but informed by a sense of environmental responsibility. It recycles the energy and resources already expended in the production of large nautical boats, into a new form. These marine fenders, intended for the most radical of weather conditions, become the seat and back support of the sofa. They are supported by a recyclable stainless steel frame to support the fenders which becomes the seat and back supports of the Crusoe Sofa. The fenders and frame can be deflated and packed for transport.

Figure 3. The Crusoe Sofa (refurbished second hand yacht fenders) designed by Alexander Lotersztain, Studio Derlot

Figure 4. The Hinkler Bench (Moso bamboo) designed by Kent Gration, Wambamboo

Figure 4 illustrates designer Gration's Hinkler Bench. Gration uses Moso bamboo in his designs to achieve the environmental benefits due to its versatility, biodegradability and rapidly renewable growth cycle. This species of bamboo has been used for centuries in buildings, as a food source and as an ornamental plant. Due to growing global demand for environmentally preferred materials, it now features in textiles, cross-laminated boards, veneers, plant-based polymers and a diverse range of carpentry products.

The design engineering solutions, elimination design and platforming are two different kinds of re-directive practices to design for sustainment [9].

5.1 Elimination design

Elimination design is learning how to design things away as well as designing things into existence. It is an interesting and often difficult exercise. Identify something that is a problem, something that is actually doing harm, and then find out a way of actually getting rid of it. In terms of creative challenge this exercise is the most complicated and time and effort consuming area within a design process; however, despite the complexity, namely the developed capability to undertake elimination design is a much demanded sphere.

Over-engineering is an example where elimination design can be used to design things away to reduce environmental impact. The outcome of over-engineering has been a consequence of high-end and specialised market acceptance resulting in product designs that are more complicated and resource hungry in terms of productivity and materials than necessary. In many design engineering cases, less is more. Keeping a design simple is defined by trained architect and French aviator Antoine de Saint-Exupéry who stated “In anything at all, perfection is finally attained not when there is no longer anything to add, but when there is no longer anything to take away.” [22].

5.2 Platforming

Platforming is recognising that as designers, companies are in a situation where people cannot simply stop designing the way they previously had, and suddenly change direction [18]. A company cannot have things crash around them in the transition of where they are now and where they want to be. By building a whole range of different kinds of platforms, a company produces items which it needs not to produce, but it also needs to survive. So a platform is a way that a corporation starts to build its future within itself, and then puts its platform as it were, into competition with itself. Hence, the company should not stop serving the design market that someone constituted but should recognise things have to change.

Honda was one of the first engine manufacturers who used Platform Design as a basis of sustainable responsibility in their design development process to become the industry leader in the Power Equipment range. Honda realised in the early 1960’s that traditional carburetted two-stroke engines were responsible for an alarming rate of emissions and contributed significantly to pollution [14]; therefore, the company adopted the policy to only manufacture four-stroke engines. Honda started to design and manufacture 4-stroke engines while continuing production of 2-stroke engine range. Today, their product assortment of lawn mowers, line trimmers, brushcutters, generators, outboard motors, snowblowers and other power equipment use ultra-low emission 4-stroke engines. Nowadays, Honda’s management strongly supports government’s legislation to ban 2-stroke engines. In addition, despite being known as the world's largest manufacturer of internal combustion engines, Honda has never built a V8 for passenger vehicles, considering petrol consumption and efficiency.

6. HISTORICAL CONDITION

In the 1930’s design engineering was introduced into the United States to speed up consumption to try to increase spending to cover the economic crisis that existed at that time [19]. This became the foundation of modern consumerism and is inseparable from the trajectory of globalisation. People can link a great deal of consumer society back to that moment. It represents separated destruction rather than simply a proliferation of consumption. Individuals have managed to find incredibly seductive and efficient ways of taking the future away by design. What is required to do is to learn to do the reverse. People need to bring the future into existence as something that has viability, recognising that the future is not an empty void in front of them which people are travelling towards. On the contrary, it is rather this thing that has thrown so much from the past into. The way that people travel into the future is by negotiating their way through all that already exists within it. One can only do that by design; people cannot get to the future by accident. As far as putting into a platform of educating sustainable design, one has to see today’s events as an opportunity. This is why leadership and opportunity is there to be sought, for if you really embrace design, practicing designers, and design educators have the position and responsibility to cause directional change that will make a significant difference. Seizing that opportunity is difficult, but has become a necessity.

7. THE CHALLENGE

It is far more difficult to change a design engineer’s thoughts, than to educate an engineer/designer in their earlier years of education. Good design contributes to the possibility of a viable future; bad design is what takes it away. A lot of things in the past have been classified as good design but did not have a good outcome in terms of what they delivered, either environmentally, socially or economically.

8. CONCLUSION

To become an ethical design engineer one must be accountable for the objects one creates, whether industrial, architectural, mechanical, civil or electrical. In addition, to be a design engineer is not about being trendy; nor is it about being seen as creative or being a problem solver; it is about being accountable to what one brings into existence. Conversely, even though in most instances the client designates what is to be designed, the engineer has the significant share in accountability for the overall effect of the designed product or project on the environment, its consumers and the world at large.

In simplistic terms, designing ethically means taking responsibility for giving form and function that minimise the use of natural resources and prevents or minimises pollution and environmental damage. An ethical design can be evaluated according to its degree of sustainability; which also involves the elimination of products which are unsustainable. Rather than create more ‘green’ things that simply add to the consumer choice, products can be eliminated by design [10]. Designers have the skills, resources and reasons to deal with sustainability; therefore, they need the will to act.

REFERENCES

[1] Australian Government, **Defence White Paper 2009: Defending Australia in the Asia Pacific Century: Force 2030**, Department of Defence, Commonwealth of Australia, 2009.

[2] A. Chochinov, "1000 words: A Manifesto for Sustainability in Design", **World Changing Team**, NY, 2007.

[3] Rt. Hon C. Clark, MP., **Sustainable Development Action Plan for Education and Skills**, UK, 2005.

[4] EPA., **United States Environmental Protection Agency**, 2009.

[5] A. Findeli, "Rethinking Design Education for the 21st Century: Theoretical, Methodological, and Ethical Discussion". **Design Issues**, Vol. 19, No. 1, 2001.

[6] J. Fischer, A. D. Manning, W. Steffen, D. B. Rose, K. Daniell, et.al, "Mind the Sustainability Gap", **Trends in Ecology & Evolution**, Vol. 22, Iss. 12, 2007, pp. 621-624.

[7] K. T. Fletcher, & E. L. Dewberry, Demi, "A Case Study in Design for Sustainability", **International Journal of Sustainability in Higher Education**, Vol. 3, No. 1, 2002.

[8] T. Fry, "Dead Institution Walking: The University, Crisis, Design & Remaking", **Design Philosophy Papers,** Issue 5, Australia, 2003.

[9] T. Fry, **Design Futuring: Sustainability, Ethics and New Practices.** UNSW Press Book, Sydney, 2009b.

[10] T. Fry, "Elimination by Design", **Design Philosophy Papers**, Issue 2, Australia, 2005.

[11] T. Fry, "Innovation Cities", **Future Tense, ABC Radio National**, 20 August, Brisbane, 2009a, Transcript.

[12] T. Fry, **Remakings: Ecology, Design, Philosophy**, Envirobook, Sydney, 1994.

[13] H.M. Government, **Securing the Future: Delivering UK Sustainable Development Strategy**, 2005, Executive Summary.

[14] Honda Motor Company Limited, **Honda Ecology: Honda Environmental Conservation Activities**, Japan, 2003.

[15] International Association of Universities (IAU), IAU Priority Themes: **Sustainable Development**, 2010. Retrieved 2010 February 02 from http://www.iau-aiu.net/sd/index.html

[16] IPCC, **Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change** [J.T.,Y. Houghton, D.J. Ding, M. Griggs, P.J. Noguer, X. van der Linden, K. Dai, Maskell, and C.A. Johnson (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2001, pp. 881.

[17] C. Laszlo, **Sustainable Value: How the World's Leading Companies are Doing Well by Doing Good**, Stanford Business Books, 2008.

[18] Z. Liu·,Y. S. Wong·& K. S. Lee, "A Manufacturing-Oriented Approach for Multi-Platforming Product Family Design with Modified Genetic Algorithm", **Journal of Intelligent Manufacture**, Published online 11 December, 2009.

[19] D. F. Noble, **America by Design: Science, Technology, and the Rise of Corporate Capitalism**, Oxford University Press, 1979.

[20] R. K. Pachauri, "Summit on Climate Change. National statement from the Chairman of the Intergovernmental Panel on Climate Change (IPCC)", a pre-recorded video for the **Conference of the Parties (COP) 15, United Nations Climate Change Conference** hosted by Denmark in Copenhagen, 2009. Video taken 22 September.

[21] V. J. Papanek, **Design for the Real World: Human Ecology and Social Change**, Thames & Hudson, London, 1971.

[22] J. Porter, **Five Principles to Design By: Guidance for Creating Designs that Work.** Techsoup, 2007 August 08. Retrieved 2010 February 02 from http://www.techssoup.org/learningcenter/webbuilding/page7396.cfm.

[23] L. L. Rasmussen, "Summit on Climate Change. National statement from Prime Minister of Denmark, a pre-recorded video" for the **Conference of the Parties (COP) 15, United Nations Climate Change Conference** hosted by Denmark in Copenhagen, 2009, Video taken 9 September.

[24] D. Rebelo, "What is The Role for Higher Education Institutions in the UN Decade of Education for Sustainable Development?" **International Conference on Education for a Sustainable Future: Shaping the Practical Role of Higher Education for a Sustainable Development**, International Association of Universities (IAU), Charles University (IAU), Karolinum, Prague, Czech Republic, 10–11 September, 2003.

[25] L. Scott, "Better Sustainment", **Live Better Sustainment Summit,** Bentonville, Arkansas, USA, 10 October, 2007.

[26] A. Smith, "What the Finniston Report Should Have Said", **Engineering Designer -** Journal of the Institution of Engineering Designers*,* United Kingdom, Vol. 7 No. 3, 1981, pp. 25-28.

[27] A. Telemi, "Summit on Climate Change. National Statement from the Prime Minister of Tuvalu", a pre-recorded video for the **Conference of the Parties (COP) 15, United Nations Climate Change Conference** hosted by Denmark in Copenhagen. Video taken 22 September, 2009.

[28] N. M. Triet, "Summit on Climate Change. National Statement from the President of the Socialist Republic of Viet Nam", a pre-recorded video for the **Conference of the Parties (COP) 15, United Nations Climate Change Conference** hosted by Denmark in Copenhagen. Video taken 22 September, 2009.

[29] UNESCO, "UN Decade of Education for Sustainable Development 2005-2014", **The DESD at a Glance**, France, 2005.

[30] UNESCO-UNEP, "Environmentally Educated Tachers: The Priority of Priorities". **Connect**, Vol. XV, No. 1, 1990, pp. 1-3.

[31] United Nations, "Process of Preparation of the Environmental Perspective to the Year 2000 and Beyond". **General Assembly Resolution,** Vol. 38, No. 161, 19 December, 1983.

[32] United Nations, "World Population to 2300", **United Nations,** New York, 2004.

[33] S. Van der Ryn & S. Cowan, **Ecological Design**, Washington, DC, Island Press, 1996.

[34] Y. Yang & J. Giard, "Industrial Design Education for Sustainability: Structural Elements and Pedagogical Solutions". Paper presented at the **Industrial Designers Society of America (IDSA) National Education Conference**, Massachusetts College of Art, Boston, 2001.

Data-driven and Integrated Engineering for Virtual Prototypes

Stephan Vornholt
Otto-von-Gurericke University Magdeburg
Magdeburg, Germany
vornholt@iti.cs.uni-magdeburg.de

and

Veit Köppen
Otto-von-Gurericke University Magdeburg
Magdeburg, Germany
koeppen@iti.cs.uni-magdeburg.de

ABSTRACT

The increasing usage of components in automotive systems is accompanied by complex dependencies of products as well as engineering and economic domains. Thus specialized analysis approaches and simulations are used to identify influences of domains and components on each other. Current research deals with database systems and data management approaches to integrate different engineering domains into a component oriented model. On the other side integrative management solutions made huge approaches in the field of computer controlled economic analysis and optimization. ERP systems and Data Warehouse approaches enable a wide range of data storage and analysis solutions. A challenging task is the interaction as well as the influence of different engineering and management disciplines on each other. In this paper a comprehensive solution for data driven and integrated engineering of prototypes in the automotive domain is presented. This approach enables early updates of concepts and structures as well as re-combination of virtual products including the planning process. Based on the description of challenges the field of domain expert integration, an integrative architecture is presented.

1. INTRODUCTION

The increasing usage of software based approaches in automotive systems is enabled by the support in the fields of computer applications and further research on control systems for complex dependencies. These allow tests and analyses on virtual prototypes before implementing and manufacturing real prototypes. A challenging task is the interaction of different components as well as the influence of different engineering and management disciplines on each other. On the engineering side the CAx techniques and concurrent virtual/digital engineering approaches enable complex product engineering, taking i.e., the phases design & construction and analysis & simulation into account. On the other side the data based management systems made huge approaches in the field of computer controlled part lists, cost control, and delivery optimization. The interaction of both research areas promises a faster and cheaper design phase as well as continuous open decisions. To support this goal an integration of both domains, management information systems and concurrent engineering, is necessary.

The early adaption of ideas and concepts is enhanced by virtual engineering (VE) in the field of mechanical engineering solutions, which lead to a huge amount of variants. Many influences have to be considered and tested including the market. Business questions have influences on the product re-design as well as the design process itself can influence business constraints. Thus, an active cross linking of tools and production partners is required. Models and strategies can change and make a re-design of a prototype necessary. This approach enables early updates of concepts and structures as well as re-combination of virtual products including the planning process.

EXAMPLE

The simplified development of an automotive system, denoted as CAR, is presented as an exemplary process following the product lifecycle, see Figure 1. Following on the basic goal definition a shared concept is developed, where constraints are defined and the conceptual design is more or less specified. Hereupon, in the design & construction phase engineers develop iteratively simulation models and virtual prototypes which are analyzed and, if the result fits the requirements, forwarded to the process and production planning as well as market and economic analyses. The white arrows represent desired steps of enhancing the model up to production, while the black arrows illustrate possible re-engineering decisions.

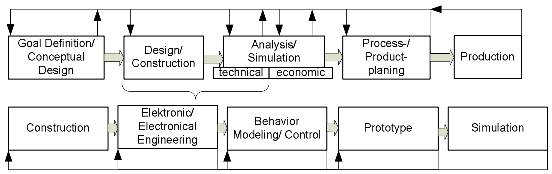


Figure 1: Life Cycle

The design itself concentrates on the geometry and connections as well as on mechanic systems, electrical drives, and controlling components. A designer of the elementary geometry could assume the structure: one body, two axis constructions, which have themselves an axis, two wheels, and two connection elements, see Figure 2a. The connection between different elements is also defined. Many mechanical definitions like weight, volume, or material properties can be used as well as predefined connection types between elements, see Figure 2b.

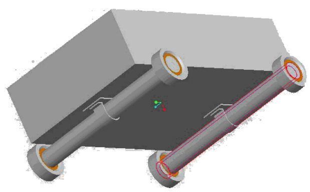


Figure 2a: Example CAR (3D CAD)

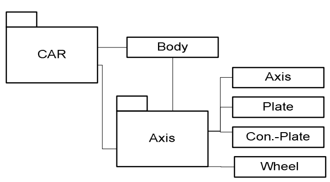


Figure 2b: Common concept CAR

According to our example, the following domains cooperate in this process:

- Electric design: the engineer designs the electric drives of the system and the electric power management. Information from conceptual design, geometry, and product libraries, e.g., engine databases, are combined.
- Simulation & Analysis: the effects are simulated with models. The results of simulation are used to identify weak spots or deficiencies and give information for further improvements. The backbones of product analysis are typically the following domains:
- Mechanical analysis, e.g., SimMechanics: mechanical simulation models are used for analyzing the mechanical behavior of prototypes, e.g., kinematic behavior, testing of possible collisions, and allowed movements.
- Mechatronic analysis, e.g., Modelica described in [4]: the mechatronic model adds electric drive and control components to the mechanism model.
- Finite element (FE) analysis (ANSYS & NASTRAN): FE models are used to analyze the elastic behavior of certain system parts. Typical tasks of FE analysis are the computation of resonant frequency or bending stress [6].

After these technical analyses several economic management tests arise, i.e., marketing, production planning, cost, and inventory. The product can be chosen to build, re-designed with new or adapted constraints, again including economic and technical decisions. Depending on the suggested changes the process of engineering is restarted.

2. RELATED WORK

The combination of approaches in Virtual Engineering (VE) and Enterprise Resource Planning (ERP) systems and an integrative architecture, combining the positive effects of new database approaches for VE and ERP systems require a deeper view on the systems.

Virtual Engineering

The product development process involves several specialized domain experts with own vocabularies, knowledge, and tasks. Thus, every expert uses domain tools and enclosed information management systems, but all work on the same conceptual design and exchange information with other groups. The interaction and heterogeneity of used data models and information systems lead to data interoperability or integration aspects and management problems within product development processes. VE targets the computer-supported parallelization of design & construction as well as simulation & analysis in order to reduce product development time and cost. A virtual prototype (VP) represents a computer based prototype of a real world artifact. It can be tested from different points of view. The VP as well as tests can be visualized in Virtual Reality (VR). Therefore, VPs comprise and combine all information, e.g., CAD geometry designs, product data, behavior models based on FE analysis, or mechanism models, and special geometry models for tests in VR. Synthesis steps, e.g., parameterization, and analysis steps are alternating during product development. In order to shorten and improve the development process, existing and verified designs, simulation models, and VR scenarios have to be re-used, modified and re-combined for new developments. Several solutions including the integration into one common data structure, e.g., presented in [1, 2], transfer of information, e.g., in [3], integration into one common tool and a storing system [5] are in the focus of research. All approaches lack in either exchanging information or specialized view or re-design.

ERP

ERP is a company-wide computer software system used to manage and coordinate all resources, information, and functions of a business from shared data stores. ERP software is a recent addition to manufacturing and information systems that have been designed to capture and organize the flow of data for the whole product life cycle. ERP software attempts to link all internal company processes into a common set of applications that share a common database. It is the common database that allows an ERP system to serve as a source for a robust data warehouse (DW) that can support sophisticated decision support and analysis.

ERP software usually has a central database as its hub, allowing applications to share and re-use data more efficiently than previously permitted by separate applications. The database of an ERP system is functional- or process-oriented organized [14]. In an ERP system the data cannot be efficiently analyzed directly, because ERP uses OLTP (Online Transaction Processing) to handle the data. OLTP is a class of program that facilitates and manages transaction-oriented applications, typically for data entry and retrieval tractions.

Data can be exchanged among disparate systems especially disparate CAD systems by Step or related solutions like MechaStep or IGES. For storing the design data it should be first archived within the organization that produced the part. Today many big corporations have archived data using a CAD/CAM format that is no longer supported by any vendor. Existing database approaches such as constraint databases are being investigated to archive design data in neutral exchange formats. To avoid geometry predominance, integrated design could be organized around an integrated product model managing design information. Current approaches usually assume that integrated design needs a unique integrated model. Typical solutions are described in [7, 8].

Product Data Management (PDM) systems are used to be integrated into a common environment within the product lifecycle. PDM is a tool that helps engineers to manage engineering data and product development processes. As PDM systems are widely used to reduce the product development time, they need to exchange product data with CAD systems. It is necessary to integrate CAD and PDM systems, because CAD systems generate the product data and PDM systems manage these data. The management of product structure data is the main function of a PDM system. PDMs should enable engineers and other users to search design models and re-use knowledge and company best practices by combining artificial intelligence techniques such as neural networks and expert systems with CAD and object-oriented databases.

Comprehensive solutions, integrating different disciplines are just researched and implemented for individual companies. These solutions are not useable for engineering clusters with many participating companies. Although the list of solutions in the disciplines themselves is long, i.e., product development, economic solutions, data exchange, an integrated approach is missing. An open architecture integrating heterogeneous systems in the same database or a common database schema as a solution for concurrent control and exchange of information is necessary and will be presented in the next section.

3. VE & ERP INTEGRATION

In this section the integration of both views of the engineering process is described. This solves on the one hand the addressed challenges and on the other hand enhances further the development process, due to use of all available information in the complete product life cycle.

Challenges

Although there are several data based solution systems in the engineering and management disciplines, where in each discipline itself the integration of different domains is in the focus of current research, a data management of both is not yet one of these solutions. The result is that both disciplines work on the same product with high concurrency and data redundancy, but the interaction between both design lines is limited to an input and output. The other discipline is seen as a black box.

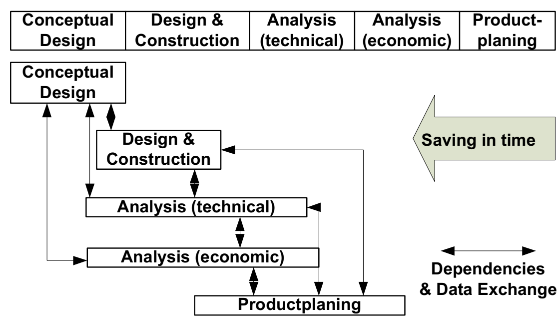


Figure 3: Concurrent Engineering & Management

By integrating both approaches into a common data schema, the main tasks of each domain should not be changed. Both need comfortable complete data storage with metadata management, data integrity, version management as well as the combined multi user interaction, consistency control, and specialized views. Those requirements for a new system are described below to enhance engineering and management at the same time, see Figure 3. The integrated data schema enables concurrent work processes, i.e. technical and economic analysis, where engineering steps can use “in-work” models to estimate their results.

One of the main challenges for databases is the (re-) definition of views. Since every cooperative partner has its own tools and data definition language as well as a concept of necessary information, the data integrated for one product has to be separated in many ways. These can be combined ones, for example the parts list, where products for the mechanical and electrical engineer are analyzed and enriched with additional information stored in databases as well as cost lists from PDM systems. The

user view on the data is the same as if he would work alone.
Comfortable complete data storage with metadata management is essential. This means, that different tools are used and it is not possible to have integration with manual re-definition of data and connections. Most work should be done automatically and every data should be storable and their metadata managed.
Information stored in the new data model should influence design decisions. Therefore dependencies and conductions have to be defined and used. Furthermore, those dependencies between information have to be held in a consistent state. An electrical engineer designing a connection between two elements which are not connected in mechanical model are inconsistent as well as constraints given by the design concept cannot be ignored by the engineers.
One of the most complicated tasks is the integration of different users, working on the same model with their own view, storing and updating the model at the same time. These changes have to be integrated in a persistent way.
Especially in heterogeneous production networks, consisting of many companies, the possibility to protect someone's information from the access is necessary. However, data security issues necessary in many forms even for each designer, who wants to work without the risk of changes in his area. This means, that he needs a warning for incompatible/inconsistent data instead of an automatic change routine is an essential requirement.

One adaptive Schema in VE
To build one adaptive schema both parts and their integration schema have to be considered:
In our scenario, three general data models are used: a CAD data model, a data model for mechatronic systems, and a FE data model. Typically, CAD systems use a feature-based, parameterizable, hierarchical data model. The construction is organized into assemblies, sub-assemblies, and parts. The properties of assemblies and parts are denoted as features and parameters. A part feature is its volume, i.e., its geometry. Additional features are for instance material or surface specifications. Assemblies group parts or other assemblies and assign positions like parameter values to them. The mechatronic data model consists of the mechanic model and the electric model. All components have parameters, e.g., inertias, center of gravity, mass. Components are classified into bodies and engines, while they are connected by ports. A port has a defined position on a component. The FE model is based on a mesh model of the possible simplified original geometry. This geometry can be derived from the CAD system, but can also be an abstract simplified geometry, e.g., 1D or 2D. The mesh model consists of mesh entities that are distinguished into elements, faces, edges, and nodes. Parameters are assigned to elements to describe materials, movement constraints, and masses.

Connections and hierarchies in the models are differently expressed and denote different real world concepts: for instance, in the CAD model, a hierarchy means a construction hierarchy and in the mechanic model connection corresponds to kinematic dependencies. Therefore, simple 1:1 correspondences between data model elements are only partly possible. Often complex conversions are necessary that also take actual models and instances into account.
Most ERP models are based either on the customer or product, but are just divided collections of reachable data without connections or interdisciplinary character. Thus the character of data models is in the structure, information area and data model are hard to characterize. To overcome this, current research is focused on the integration of different ERP solutions into one DW and also about the question why past research in this area is not used in current ERP systems. The integration of ERP systems in this paper assumes a DW, where different ERP solution databases are integrated into a common data structure which is developed and enlarged with the tasks of the integration database.

Integration-Architecture & Schema
The integration architecture can be divided into three basic components: the ERP data warehouse, the VE integrative DW and the multi database.
The ERP data warehouse consists of a basic structure, which holds product based information on products or product parts. The UID (unit id) represents the link to different material or part list databases. Further information, e.g., on cost, can be stored as well as customer information or delivery information. The information schema in the data warehouse is integrative, which means, that different tools are tested of consistency and can change data which is already defined. An adjustable feedback function translates changes on relevant data back to each related system, or just adds a new version with a warning to all inconsistent data variants. Furthermore an adaptable view function is included. The VE database is based on the solutions of the component based virtual family definition in [9, 10]. Here, several closely related information on technical descriptions are managed and held consistent. Feedback, views, and consistency are integrated, too. The Multi database integrates both approaches into one schema, as depicted in Figure 4.

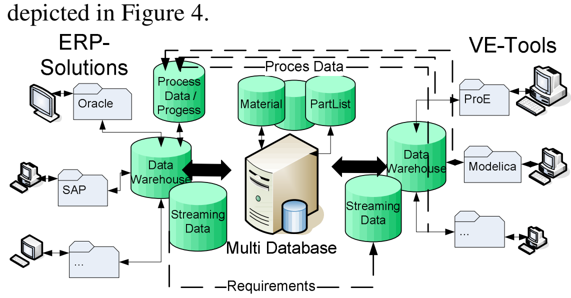


Figure 4: Integration Database (Basic Idea)

The Multi database is based on a component-oriented model that describes VPs in a multidisciplinary way. Figure 5 illustrates the main concepts of the model using UML notation.

A domain model describes the behavior, form, or function of a VP in one engineering domain, e.g., mechanic behavioral model or the CAD model. A domain model is composed by sub-models. Since ERP data models are combined and often not especially related to one basic model, they can be allocated to the component directly. Each model provides a set of parameters and of ports. Parameters are quantities that describe the characteristics of the model and the characteristics of a VP in a certain domain. A port is a connection point where models can be combined. Through connected ports signals, material, and forces are transferred. The libraries are integrated in the ERP data model as well as common ERP data with their constraints and specified processes. In summary, a model is represented by a 3-tuple M = (id, parameters, ports) while ERP models are represented by ERPM= (id, constraints, processes, parameter).

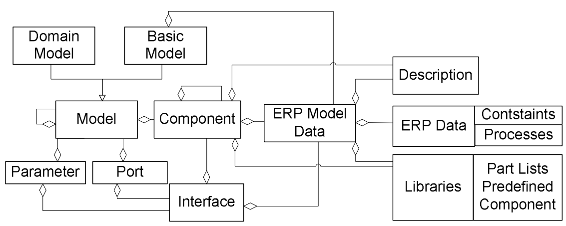


Figure 5: UML Schema

Each domain model and ERP model represents one view to a VP and does not support inter-domain correspondences. Therefore, components are introduced which are hierarchically organized. A component represents a conceptual part of a VP and encapsulates all models as well as their dependencies for this artifact. A component ensures the dependencies between different domains, while domain models combine sub-models within one domain. Furthermore, components can contain sub-components. Components provide interfaces for communication and parameterization. An interface consists of parameters as well as ports. Constraints and mappings ensure the consistency control within one component. Combining the concepts, a component is defined as a tuple C = (id, M, C, Pdept, PMdept, I) with an identifier id, M and C are sets of encapsulated models and components, respectively. Pdept represents a set of port mappings between different domain models and PMdept is a set of parameter dependencies. Constraints and processes can be translated into dependencies or parameter. Finally, the interface I, consisting of a set of external ports and parameters, describes the behavior of the component to the environment. The interface is mapped to internal models and components.

Figure 6 illustrates our exemplary component CAR. The component contains of two domain models, a CAD model as well as a combined ERP model. A set of dependencies describes the internal relations between the domain models. The external interface offers ports and parameter that are mapped internally to the domain models. The component can be distributed and instantiated in the coupled way allowing the usage directly in CAD as well as ERP models.

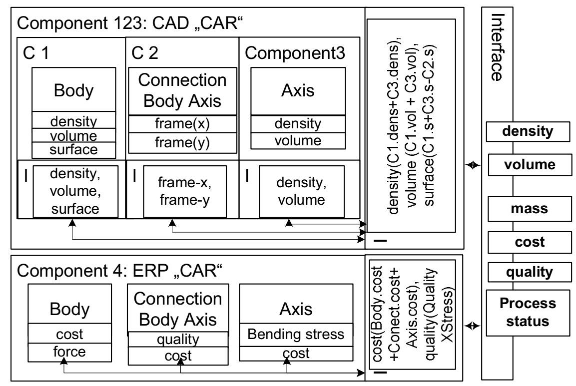


Figure 6: Component Construction of CAR

Views: An information overload results if too many details are used at the same time in the development process. Views filter the relevant information for each costumer, the engineer as well as the manager or initiator. The views can be integrated in the schema and processes as illustrated in Figure 7. The arrows illustrate the combination and integration of data representations as well as the influence of changed properties or parameters. Based on the (meta-) data which are stored in files and folders the integrated data are stored in one component based structure. Inconsistent information is eliminated and libraries as well as files are linked in the (meta-)data repository.

The global view contains all parameters and ports in each file and can be addressed directly. The economic interface contains all relevant DW information on ERP data and hides the technical information. The technical interface contains technical data from virtual engineering. Business and economic information are not viewable. Another defined view is the combined interface, which contains information on both or information which is necessary in both disciplines as well as new views on the data. Views are defined and planned to be supported by an adaptive view selection where dependencies are defined. Thus any combination, new sights and approaches are possible.

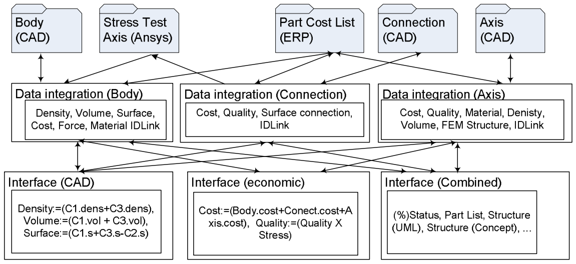


Figure 7: Views

4. Conclusion and Outlook

In this paper we present data driven and integrated engineering of prototypes in the product development domain, which is illustrated by an automotive example. Based on the description of a virtual engineering integration solution the ERP data warehouse integration into a common structure is focus in this paper. The new integrative architecture is based on the challenges in the field of domain expert integration. Especially, the view concept enables different approaches to use and specify the integrative data as well as dependencies and further control options. The next steps in this field are the integration of different ERP systems in one data DW and the implementation of the system. Further steps and integration areas, such as logistics or training, are planned to be done in the future.

Acknowledgements

The work of Stephan Vornholt is supported by the European Commission: European Regional Development Fund, COMO C1-3201201 and C3-3201201. The Work of Veit Köppen is funded by the German Ministry of Education and Science (BMBF), project 01IM08003C.

References

[1] Fenves SJ, Foufou S, Bock C, Sudarsan R, Bouillon N, Sriram RD, **"CPM 2: A Revised Core Product Model for Representing Design Information"** Technical Report, NISTIR 7185, 2004.

[2] Xue D., Yang H, **"A concurrent engineering-oriented design database representation model."** Computer-Aided Design, 36:947-965, 2004.

[3] Juhasz T, Schmucker U, **"Automatic Model Conversion to Modelica for Dymola-based Mechatronic Simulation."** In: Bachmann B (ed.) Proceedings of Modelica' 2008, The Modelica Association, Bielefeld, p. 719-726, 2008.

[4] Modelica Association, **"Modelica - A Unified Object-Oriented Language for Physical Systems Modeling -- Language Specification -- Version 3.0"**, 2007.

[5] Bettaieb S, Noel F, **"A generic architecture to synchronise design models issued from heterogeneous business tools: towards more interoperability between design expertises."** In: Engineering with Computers, 2007.

[6] Armstrong CG, **"Modeling Requirements for Finite-Element Analysis"**, Computer-Aided Design 26(7):573 – 578, 1994.

[7] ISO, **"ISO 10303:1994 - Industrial Automation Systems and Integration - Product Data Representation and Exchange"**, 1994.

[8] Pavez L., **"STEP Datenmodelle zur Simulation mechatronischer Systeme."** Abschlussbericht des Verbundprojekts MechaSTEP, Forschungszentrum Karlsruhe GmbH, 2001.

[9] Vornholt S, Geist I**, "Flexible Integration Model for Virtual Prototype Families."** In: Proceedings of PLM08 5th International Conference on Product Lifecycle Management, Seoul, Korea, 2008.

[10] Vornholt S, Geist I, **"Interface for Multidisciplinary Virtual Prototype Components."** In: 19th International Conference on Database and Expert Systems Application (DEXA), 1st International Workshop on Data Management in Virtual Engineering DMVE '08, 2008.

[11] Esteves J., Pastor J., **"Enterprise resource planning systems research: an annotated bibliography."** In: Communication of AIS 7(8), 2001.

[12] Silwood Technology Ltd, **"Saphir White Paper: Metadata management for Enterprise Applications."** Manual, 2008.

[13] CA, Transforming IT Management, **"Technology Brief: CA ERwin Saphir Option."** Manual, 2009.

[14] Simon, Elke, **"Implementation of an Enterprise Resource Planning System with Focus on End-User Training"**, Hamburg: Diplomica Verlag, 1999.

Fault Detection and Isolation for a bus suspension model using an Unknown Input Observer Design

Juan Anzurez Marin, Luis A. Torres Salomao
j.anzurez@ieee.org, janzurez@umich.mx, latsalomao@ieee.org
División de Estudios de Posgrado de la Facultad de Ingeniería Eléctrica-Universidad Michoacana de San Nicolás de Hidalgo
Morelia, Michoacán C.P. 58030, México

Isidro I. Lázaro
ilazaro@umich.mx
Facultad de Ingeniería Eléctrica-Universidad Michoacana de San Nicolás de Hidalgo
Morelia, Michoacán C.P. 58030, México

ABSTRACT

A fault diagnosis scheme based on unknown input observers design is presented in this paper. The used technique is a model based approach, being the main task the observation of error signals known as residuals. These also known as robust residuals are derived by comparison of the system output and the estimated output. The robust residuals are generated by the Unknown Input Observer (UIO) design, this have their state estimation error vector approaches zero asymptotically as their special feature, regardless of the presence disturbances or unknown inputs in the system. In this paper we use the ¼ of a standard bus suspension to prove the design of Unknown Inputs Observers applied to the fault diagnosis problem in sensors.

Keywords: Fault Diagnosis, Unknown Input Observers, Sensor fault.

1. INTRODUCTION

Modern control systems are becoming more and more complex and control algorithms more and more sophisticated. Consequently, the issues of availability, reliability, and operating safety are of major importance. For safety-critical systems, the consequences of faults can be extremely serious in terms of human mortality and economic loss. Therefore, there is a growing need for the supervision and fault diagnosis analysis to increase the reliability of such safety-critical systems due to early indications concerning which faults are developing can help avoid system breakdown, mission aborting and catastrophes. Since the beginning of the 1970s, research in fault diagnosis has been gaining increasing consideration world-wide in both theory and application, a strong impetus comes from the side of modern control theory that has brought powerful techniques of mathematical modeling, state estimation and parameter identification that have been made feasible by the progresses of the computer technology [1-3].

The fault diagnosis process basically consists of three levels; the first level is fault detection, which indicates when a fault occurs in the system. The second level corresponds to fault isolation, where the location of the fault is determined. Third and final level is fault identification that estimates the size and type or nature of the fault. By this reason this process is also known as Fault Detection and Isolation (FDI) [8], [9], [10].

The model based FDI approach requires a system's model, which is an idealized assumption. In practice this assumption is not fully met. Usually the parameters of the system are uncertain or varying in time [6]. However, Unknown Input Observer (UIO) design, aside from modeling system uncertainties, solves the FDI problem in an effective way. This technique surpasses the classical approach of hardware redundancy for a software redundancy, with an obvious cost-efficiency benefit. UIO design requires apart from the system model, detectable output stages as well. The residual is obtained by comparison of the actual system output and the estimated observers output. The purpose of UIO's is to produce an estimated output which approaches asymptotically to actual state meanwhile rejects noise effects and modeling errors of the system. In this approach the perturbations should be decoupled from the generated residuals. This is achieved by assuming the unknown inputs matrix (disturbances) as known. Based on this information, the disturbances can be decoupled.

In this work, to analyze the sensors fault detection problem we propose to combine the advantages of the UIO proposed by Chen [1] which has the ability to decouple the disturbances with the advantages of dedicated observers to achieve an adequate isolation, applied to a bus suspension, since it is a very important system in the automotive sector.

The paper is organized as follows: In section 2 some basic concepts about unknown input observers are shown. In section 3 the bus suspension model that we use in this paper as study case is presented. Section 4 depicts the results obtained using the unknown input observers applied to a bus suspension system. Finally, in section 5 we present some conclusions about the used technique for the fault diagnosis problem.

2. UNKNOWN INPUTS OBSERVER DESIGN

In the considered system class in this case, the uncertainties can be summarized as an additive term in the state space dynamic equation as:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) + Ed(t) \\ y(t) = Cx(t) \end{cases} \qquad (1)$$

Where $x(t) \in \mathbb{R}^n$ is the state vector, $y(t) \in \mathbb{R}^m$ is the output vector, $u(t) \in \mathbb{R}^r$ is the known input vector and $d(t) \in \mathbb{R}^q$ is the disturbance or unknown input vector. ***A***, ***B***, ***C*** and ***E*** are known matrices with appropriate dimensions.

Then the residual is obtained as follows:

$$r(t) = y(t) - C\hat{x}(t) \tag{2}$$

Therefore, an observer is UIO defined for the system (1) if its state estimation error vector *e(t)* approaches zero asymptotically, regardless of the presence of the unknown input (disturbance) in the system.

This paper presents a complete order UIO scheme proposed by Chen and Patton [1]. The structure for a complete order observer is described as:

$$\begin{cases} \dot{z}(t) = Fz(t) + TBu(t) + Ky(t) \\ \hat{x}(t) = z(t) + Hy(t) \end{cases} \tag{3}$$

Where $x(t) \in \mathbb{R}^n$ is the state vector, $z(t) \in \mathbb{R}^n$ is the state of this complete order observer, and ***F***, ***T***, ***K***, ***H*** are matrices to be designed for achieving unknown input decoupling and other design requirements. The observer described by equation (3) is illustrated in Figure 1.

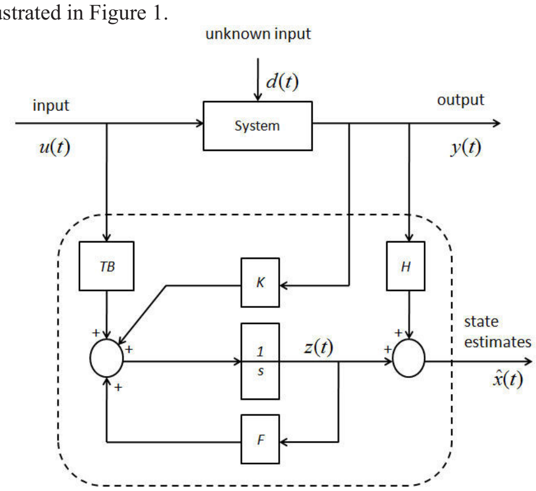


Figure 1. Unknown Input Observer Structure

When observer (3) is applied to the system (1), the estimation error (4) is governed by the equation (5).

$$e(t) = x(t) - \hat{x}(t) \tag{4}$$

$$\begin{aligned} \dot{e}(t) &= (A - HCA - K_1C)e(t) \\ &+ [F - (A - HCA - K_1C)]z(t) \\ &+ [K_2 - (A - HCA - K_1C)H]y(t) \\ &+ [T - (I - HC)]Bu(t) + (HC - I)Ed(t) \end{aligned} \tag{5}$$

Where:

$$K = K_1 + K_2 \tag{6}$$

If the following relations hold true:

$$\left.\begin{aligned} (HC - I)E &= 0 \\ T &= I - HC \\ F &= A - HCA - K_1C \\ K_2 &= FH \end{aligned}\right\} \tag{7}$$

The state estimation error will then be:

$$\dot{e}(t) = Fe(t) \tag{8}$$

If all eigenvalues of ***F*** are stable, $e(t)$ will approaches zero asymptotically, i.e. $\hat{x} \to x$. This makes the observer (3) UIO for the system (1).

In order to meet the design requirements presented above and obtain the matrices for achieving unknown input decoupling, the next simple algorithm should be followed:

1. Check rank condition for ***E*** and ***CE***: if $rank(CE) \neq rank(E)$, a UIO does not exist, go to 10.
2. Compute *H*, *T* and A_1:
   $$H = E[(CE)^T CE]^{-1}(CE)^T \tag{9}$$
   $$T = I - HC \tag{10}$$
   $$A_1 = TA \tag{11}$$
3. Check the observability: If (C, A_1) is observable, a UIO exists and $\boldsymbol{K_1}$ can be computed using pole placement, go to 9.
4. Construct a transformation matrix ***T*** for the observable canonical decomposition: To select independent $n_1 = rank(W_O)$ (W_O is the observability matrix of (C, A_1)) row vector $t_1^T, \ldots, t_{n_1}^T$, from W_O, together other $n - n_1$ row vector $t_{n_1+1}^T, \ldots, t_n^T$ to construct a non-singular matrix as:
   $$T = [t_1, \ldots, t_{n_o}; t_{n_o+1}, \ldots, t_n]^T \tag{12}$$
5. Perform an observable canonical decomposition on (C, A_1):
   $$TA_1T^{-1} = \begin{bmatrix} A_o & 0 \\ A_{12} & A_{no} \end{bmatrix} \tag{13}$$
   $$CT^{-1} = \begin{bmatrix} C_o & 0 \end{bmatrix} \tag{14}$$
6. Check the detectability of (C, A_1): If any one of the eigenvalues of A_{no} is unstable, a UIO does not exist, go to 10.
7. Select n_1 desirable eigenvalues and assign them to $A_o - K_t^1 C_o$ using pole placement.
8. Compute $K_1 = T^{-1}K_t = T^{-1}\left[(K_t^1)^T \quad (K_t^2)^T \right]^T$, where K_t^2 can be any $(n - n_1) \times m$ matrix.
9. Compute *F* and *K*:
   $$F = A_1 - K_1C \tag{15}$$
   $$K = K_1 + K_2 = K_1 + FH \tag{16}$$
10. STOP.

For detecting a fault in particular, an isolation scheme should be used; in this paper we use a dedicated observer scheme. It is possible for a system to present faults in both actuators and sensors. However in this paper only addresses the problem of detecting faults occurred in sensors. The isolation scheme for this purpose assumes that all actuators are fault free so the system equations can be written as:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) + Ed(t) \\ y^j(t) = C^j x(t) + f_s^j(t) \\ y_j(t) = c_j x(t) + f_{sj}(t) \end{cases} \tag{17}$$

For $j = 1, 2, ..., m$

Where $c_j \in \mathbb{R}^{1\times n}$ is the $j-th$ row of matrix ***C***, $C^j \in \mathbb{R}^{(m-1)\times n}$ is derived from matrix ***C*** by removing the $j-th$ row c_j, $y_j(t)$ is the $j-th$ component of $y(t)$ and $y_j(t) \in \mathbb{R}^{(m-1)}$ is derived from vector $y(t)$ without considering the $j-th$ component y_j. f_s^j represents the fault in sensor j. Accordingly, the *m* UIO – based residual generator can be constructed as:

$$\begin{cases} \dot{z}^j(t) = F^j z^j(t) + T^j Bu(t) + K^j y^j(t) \\ r^j(t) = (i - C^j H^j) y^j(t) - C^j z^j(t) \end{cases} \tag{18}$$

For $j = 1, 2, ..., m$

Where the matrices are due to satisfy the following equations:

$$\left. \begin{aligned} H^j C^j E^j &= E^j \\ T^j &= I - H^j C^j \\ F^j &= T^j A - K_1^j C^j \\ K_2^j &= F^j H^j \\ K^j &= K_1^j + K_2^j \end{aligned} \right\} \tag{19}$$

For $j = 1, 2, ..., m$

Each residual generator is driven by all inputs and only one output; this is shown in figure 2. When all actuators are fault-free and a fault occurs in the $j-th$ sensor, the residual satisfies the following isolation logic:

$$\left\| r^j(t) \right\| < T_{SFI}^j \tag{20}$$

For $j = 1, 2, ..., m$

Where T_{SFI}^j are isolation threshold.

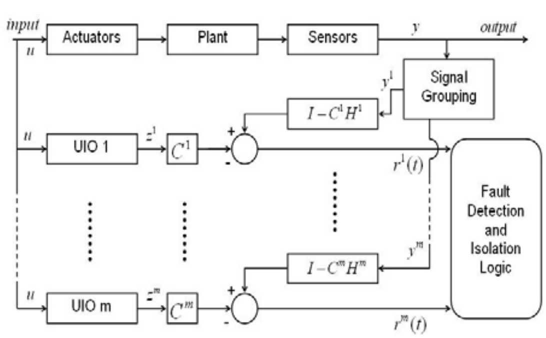


Figure 2. Sensor fault isolation scheme

3. BUS SUSPENSION MODEL

The simulation presented in this work is based on a dynamic system of a ¼ bus suspension model. The model diagram is shown in figure 3, and is described by the following equations [4].

$$\begin{cases} M_1 \ddot{x}_1 = -b_1(\dot{x}_1 - \dot{x}_2) - k_1(x_1 - x_2) \\ M_2 \ddot{x}_2 = b_1(\dot{x}_1 - \dot{x}_2) + k_1(x_1 - x_2) + \\ b_2(\dot{W} - \dot{x}_2) + k_2(W - x_2) \end{cases} \tag{21}$$

Where M_1 and M_2 are body mass and suspension mass respectively, k_1 is the spring constant of suspension system, k_2 is the spring constant of Wheel and tire, b_1 is the damping constant of the suspension system, b_2 the damping constant of the wheel and tire, x_1 and x_2 are distances and W is reference for any road disturbance.

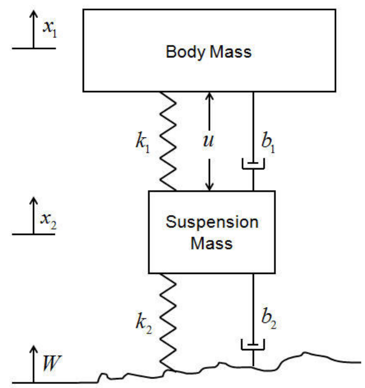


Figure 3. ¼ Bus Suspension model

For the UIO design procedure is necessary to have the state-space system model and assuming that you have the model of the plant (bus suspension) whose matrices are shown below:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{-b_1 b_2}{M_1 M_2} & 0 & \left(\frac{b_1}{M_1}\left(\frac{b_1}{M_1} + \frac{b_1}{M_2} + \frac{b_2}{M_2} \right) - \left(\frac{k_1}{M_1} \right) \right) & \frac{-b_1}{M_1} \\ \frac{b_2}{M_2} & 0 & \left(\frac{b_1}{M_1} + \frac{b_1}{M_2} + \frac{b_2}{M_2} \right) & 1 \\ \frac{k_2}{M_2} & 0 & \left(\frac{k_1}{M_1} + \frac{k_1}{M_2} + \frac{k_2}{M_2} \right) & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 \\ \frac{1}{M_1} \\ 0 \\ \left(\frac{1}{M_1} + \frac{1}{M_2} \right) \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad E = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

We begin the design process of the residual generators whose results are shown in the next section.

4. RESULTS

In order to design the UIO, the following magnitudes were substituted in the state space matrices of the bus suspension model. These values represent a bus suspension model.

Body mass (M_1) = 2500kg
Suspension mass (M_2) = 320kg
Spring suspension constant (k_1) = 80,000N/m
Spring wheel and tire constant (k_2) = 500,000N/m
Damping suspension constant (b_1) =350Ns/m
Damping wheel and tire constant (b_2) = 15,020Ns/m

Applying the design algorithm presented in Section 2, the first step in regard to range test, the ***rank (CE) = rank (E)***, so we go to step 2 and compute the matrices ***T*** and ***H*** using (9) and (10), obtaining the following:

$$H = \begin{bmatrix} 0.2500 & 0.2500 & 0.2500 & 0.2500 \\ 0.2500 & 0.2500 & 0.2500 & 0.2500 \\ 0.2500 & 0.2500 & 0.2500 & 0.2500 \\ 0.2500 & 0.2500 & 0.2500 & 0.2500 \end{bmatrix},$$

$$T = \begin{bmatrix} 0.7500 & -0.2500 & -0.2500 & -0.2500 \\ -0.2500 & 0.7500 & -0.2500 & -0.2500 \\ -0.2500 & -0.2500 & 0.7500 & -0.2500 \\ -0.2500 & -0.2500 & -0.2500 & 0.7500 \end{bmatrix}$$

Next, we compute the matrices F and K using (14) and (16), obtaining the following:

$$F = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -4 \end{bmatrix} \qquad K = \begin{bmatrix} -400 & 0.5 & 479.2 & -0.5 \\ -407.8 & 1.3 & 453.7 & -0.9 \\ -354.5 & -1 & 433.6 & 0 \\ 1160.8 & -1.3 & -1366 & 2.8 \end{bmatrix}$$

Substituting the matrices ***H***, ***T***, ***F*** and ***K*** in (3) we build a robust UIO to generate the residuals in order to detect faults in the sensors of the bus suspension system. The results from simulation are presented in figure 4, where a fault free system is shown first, then a fault in the state 2 is presented and finally we have a fault in state 3. Twenty seconds of simulation are presented; faults were introduced with a one second step input. However, fault isolation would not be possible with this approach, as shown in Figure 4, in graphs 2 and 3 are active residuals, but the change is not appreciated.

To achieve fault isolation the approach described in equations (18) was applied, ie a dedicated observer approach. Which is required to satisfy the conditions in (19), indicating a separation of the matrices in step two of the design algorithm, the results of these calculations are shown below. Note that matrices ***A***, ***B*** and ***E*** are not modified. However a new sub matrix ***C***, is obtained for each UIO designed, removing the respective row in the original matrix ***C***, as shown below.

$$C_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad C_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

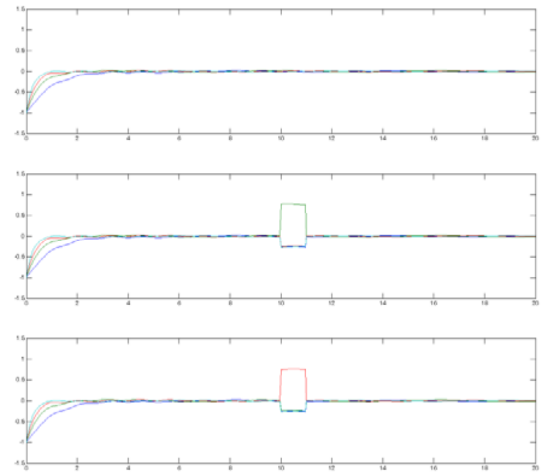


Figure 4. Robust UIO residuals

In this paper, we only show the results of the UIO 2 and 3, since the simulation results do not influence the result of fault isolation. However, it is important to note that all states have designable UIO's. As result from step 2 of the algorithm design we obtain the following values for the sub matrices ***H***.

$$H_2 = \begin{bmatrix} 0.3333 & 0.3333 & 0.3333 \\ 0.3333 & 0.3333 & 0.3333 \\ 0.3333 & 0.3333 & 0.3333 \\ 0.3333 & 0.3333 & 0.3333 \end{bmatrix} = H_3$$

Using (19) we obtain the values for sub matrices ***T***, ***F*** and ***K***, which result as follows:

$$T_2 = \begin{bmatrix} 0.6667 & 0 & -0.3333 & -0.3333 \\ -0.3333 & 1 & -0.3333 & -0.3333 \\ -0.3333 & 0 & 0.6667 & -0.3333 \\ -0.3333 & 0 & -0.3333 & 0.6667 \end{bmatrix}$$

$$T_3 = \begin{bmatrix} 0.6667 & -0.3333 & 0 & -0.3333 \\ -0.3333 & 0.6667 & 0 & -0.3333 \\ -0.3333 & -0.3333 & 1 & -0.3333 \\ -0.3333 & -0.3333 & 0 & 0.6667 \end{bmatrix}$$

For sub matrices ***F***:

$$F_2 = \begin{bmatrix} -4.7407 & 0.6667 & 0.9165 & 0.9165 \\ -6.0059 & -0.3333 & 3.1607 & 3.1607 \\ 0.8793 & -0.3333 & -2.4630 & -1.4630 \\ 0.8793 & -0.3333 & -1.4630 & -2.4630 \end{bmatrix}$$

$$F_3 = \begin{bmatrix} -31.2 & -141.7 & 623.3 & 226.1 \\ -13.1 & -156.6 & 598 & 215.4 \\ -14 & -136.2 & 575.1 & 213.7 \\ 30.8 & 280.1 & -1221.2 & -457.3 \end{bmatrix}$$

For the sub matrices ***K***:

$$K_2 = \begin{bmatrix} -509.8 & 617.9 & -13.3 \\ -181.4 & 424.6 & -181.5 \\ -502.9 & 594.2 & 2.2 \\ 1012.7 & -1212.1 & 11.2 \end{bmatrix}$$

$$K_3 = \begin{bmatrix} -261.9911 & 367.8684 & -0.5549 \\ -297.5605 & 370.8519 & -0.9363 \\ -244.8422 & 348.7216 & 0.2093 \\ 557.1806 & -736.3204 & 1.4622 \end{bmatrix}$$

Substituting these values in (18) we designed an observer and residual generator for each UIO. Proof of the fault diagnosis algorithm in this paper was conducted based on a fault simulation on the acceleration and offset of the bus damper sensors, states 2 and 3 ($\dot{x}_1$ and $y_1 = x_1 - x_2$). $\dot{x}_1$ would be easily accessible in a real situation by integrating the output of an accelerometer mounted on the bus.

Figure 5 shows the residuals generated for a fault free system. It is easy to see that after stabilization time the UIO's remain at zero, indicating that no fault has occurred in the system.

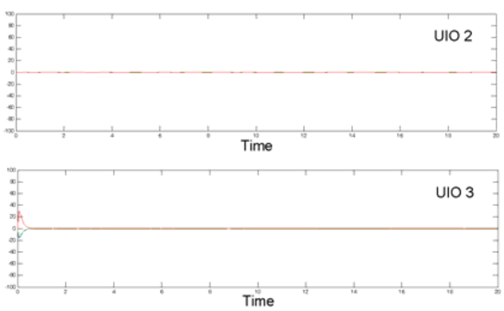


Figure 5. Fault free residuals

Figure 6 shows residuals for a fault occurred in state 2. We can see that the residual in UIO 3 is activated; i.e. does not satisfy condition (20).

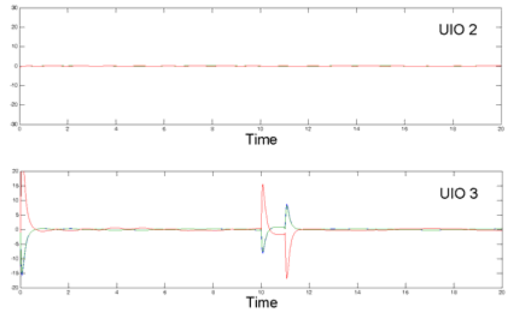


Figure 6. Sensor 2 fault residuals

Figure 7 shows residuals for a fault occurred in state 3. Residuals in UIO 2 is activated.

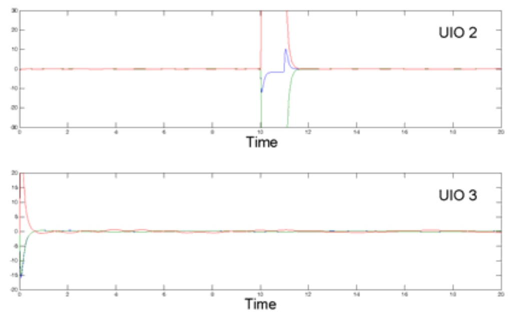


Figure 7. Sensor 3 fault residuals

5. CONCLUSIONS

This paper addresses fault detection and isolation via an Unknown Input Observers Fault Diagnosis Scheme applied to a ¼ bus suspension model; the robust fault diagnosis algorithm was successfully tested in simulation. According to the UIO's theory with modified isolation are driven by all states except those trying to detect. So that, for a fault w in a specific state will activate all UIO's residuals except for themselves. It can be appreciated in the simulations made that, for a fault occurred in state 2, the UIO 3 activates and vice versa. This corresponds with the expected behavior of an unknown input observer. From the obtained results it should be assumed that UIO design is a good tool in order to solve the fault diagnosis problem for systems with known mathematical model. It is also important to note that the design method presented can be applied to systems not completely observable by checking the observable part using a transformation matrix to obtain an observable canonical decomposition.

6. REFERENCES

[1] Jie Chen y Ron J. Patton, Robust Model-Based Fault Diagnosis for Dynamic Systems, Kluwer Academia Publishers, Boston, 1999.

[2] J. A. Marin, N. Pitalúa-Díaz, O. Cuevas-Silva and J. Villar-García, "Unknown Inputs Observers Design for Fault Detection in a Two-tank Hydraulic System", *2008 Electronics, Robotics and Automotive Mechanics Conference CERMA*, México, pp. 373-378, 2008.

[3] R. J. Patton, P. M. Frank and R. N. Clark, Issues of Fault Diagnosis for Dynamic Systems, Springer, 2000.

[4] Control Tutorials for MatLab and Simulink, bus suspension. April 2009, www.library.cmu.edu/ctms/

[5] Jason Ridenour y Stainislaw H. Zak, "Observer Based Fault Detection and Isolation", *Technical Report Advancing Technology Through Collaboration*, 2004.

[6] S. Nowakowski, M. Darouach and P. Pierrot, "A New Approach for Failure Detection in Uncertain Systems", *Proceedings of the 32nd Conference on Decision and Control*, San Antonio, Texas, pp. 257-258, 1993.

[7] C.S. Kallesoe, V. Cocquempot and R. Izadi-Zamanabadi, "Model Based Fault Detection in a Centrifugal Pump Application", *IEEE Transactions on Control Systems Technology*, Vol. 14, No. 2, pp. 204-215, 2006.

[8] J. Anzurez M., "Diagnóstico de fallas en sistemas no lineales usando lógica difusa y observadores con modos deslizantes", Tesis para obtener el grado de doctor en ciencias especialidad ingeniería eléctrica, CINVESTAV unidad Guadalajara, México, 2007.

[9] R. J. Patton, P.M. Frank and R.N. Clark, "Fault Diagnosis in Dynamic Systems: Theory and Applications", Prentice Hall, 1989.

[10] A. Akhenak, M. Chadli, J. Ragot and D. Maquin, "State estimation of uncertain multiple model with unknown inputs", *43rd IEEE Conference on Decision and Control*, Atlantic, Paradise Island, Bahamas, 2004.

BIOGRAPHIES:

Juan Anzurez Marin, received the B.E. degree in Electrical Engineering from the Universidad Michoacana de San Nicolas de Hidalgo (UMSNH), Mexico, in 1991; the M.Sc. degree in Electronic Engineering from the Instituto Tecnologico de Chihuahua, Mexico, in 1997 and Ph.D. degree in Electrical Engineering option Automatic Control from the Centro de Investigacion y Estudios Avanzados del IPN (CINVESTAV), Guadalajara Campus, Mexico, in 2007. He is an Associate Professor with Department of Electrical Engineering of UMSNH from 1987, Morelia, Mexico. His research interests include Instrumentation and Control Systems as well as the Fault Diagnosis Algorithms Design for the Nonlinear Systems and Energy Harvesting Applications. He is an IEEE member.

I. I. Lázaro is from Cordoba Veracruz, México. He holds the BSEE and MSEE degrees from the Facultad de Ingeniería Eléctrica Universidad of Michoacán. At the present he is a Professor at the same institute. His research interest is in electric power quality and power electronics. He is an IEEE member.

Luis Alberto Torres Salomao, He holds the B.S. degree in Electronic Engineering from the Universidad Michoacana de San Nicolas de Hidalgo (UMSNH), Mexico. His research interests are Control Systems, Fault Diagnosis, Microcontroller applications and Mecatronics. He is an IEEE member.

Engineering Technology and Gender: Enhancing Voice and Access for Minority groups by Curricular Design for Distance Learning

Kaninika BHATNAGAR
School of Technology, Eastern Illinois University
Charleston IL USA

ABSTRACT

Instructional framework for distance education has traditionally focused on teaching with technology. However, the notion of technology as a teaching tool may on occasions fail to take into account the learner characteristics, and the concept of learning with technology. This paper proposes a learning model focusing on learner criteria such as access, voice, and expectancy in addition to the online teaching triad of content, technology, and pedagogy. Gender and technology elective seminar course has been selected as a test case, due to its discussion oriented, participation intensive structure with a higher likelihood of a diverse student body.

Keywords: Engineering Technology, Curricular Design, Gender, Minority groups and Distance learning.

1. LITERATURE REVIEW

In general distance education learners are reported to have different characteristics and requirements from traditional learners, and the virtual classroom is found to differ significantly from a traditional classroom [1]. Online instruction can potentially alter classroom power dynamics, as it allows the hesitant, the introvert, those who are not as comfortable speaking out in English, and the shy students to participate without the need to speak up in class. Masters & Oberprieler [2] reported student participation as one of the main benefits of online education. They also pointed to the need to ensure access and to make sure that no one group of students tended to unduly dominate the virtual instructional landscape. The shift in classroom power dynamics affects the promotion of diversity in general, and women students in particular.

The facilitating nature of online courses, promoting participation from diverse groups, is a common theme that runs through the distance learning literature [2, 3, 4]. However, literature has tended to focus on online teaching more than on online learning. It has been suggested that good online teaching typically derives from the constant interaction between 3 components, namely content, pedagogy, and technology [5]. Kohler et al (2004) discuss this triad in the context of the role of technology in online teaching. They define ‘content’ as the set of core ideas, knowledge, procedures, resources etc. ‘Technology’ according to this model consists of presentation techniques, whether these be chalk and blackboard, or the Internet and digital video. The notion of technology is based upon interaction, and presentation. ‘Pedagogy’ according to Koehler et al (2004) forms the third leg of the triad, encompassing the process, practice, and methods of teaching. Pedagogy includes instructional, as well as evaluation methods.

2. TECHONOLOGY AND PEDAGOGY

This paper argues for an inclusion of three additional criteria to complement the aforementioned teaching triad. It is suggested that although the model in its present form fully informs ‘teaching with technology’, however it fails to capture the issues inherent in ‘learning with technology’. A set of learner criteria are identified and discussed in depth in the context of Gender and Technology seminar course, namely ‘voice’, ‘access’, and ‘expectancy’.

Voice represents a learner’s degree of participation in responding to, and collaborating with his or her classmates in a virtual classroom. Literature suggests that female participation tends to go up in online instructional situation [6].

Access represents the ability to use technology, a factor that could often operate as a gatekeeper to online education. It has been widely reported that minority groups continue to suffer to a greater degree from lack of access to technology, as in the Internet access [7]. Learners’ participation is meaningless without access. Thus although virtual classroom may serve to strengthen the participants’ ‘voice’, yet ‘access’ to technology may yet remain problematic.

Expectancy is the third leg of the learner triad, and forms the motivational backbone of this model. Expectation comprises a huge part of a learner’s response and experience of any course. Expectancy theory, and motivation theories have reported significant gender differences in expectation for success, and hence lead to differential participation and outputs [8, 9, 10].

The three criteria identified in this paper all address the learner. It is important to point out that merely enhancing technology instruction fails to address the complete picture. Teaching with technology is a function of the teacher who may do his or her job in the most effective manner possible, providing comprehensive content aided with technology, and supported by insightful pedagogical models. However, such instruction is still presented from the perspective of the broadcaster, and not the receiver. The most robust teaching practice is not effective, unless it is fully communicated to the learner. This is where the notion of a ‘learning triad’ comes in. Figure 1 illustrates the modified instructional framework after incorporating the factors for learner’s experience into the teaching triad.

The first learner criterion of ‘voice’, or participation determines who speaks up, and who doesn’t. In a virtual classroom online posts replace speaking up, and the traditional power dynamic can shift. The skill set needed to sit at one’s computer, think, and rethink one’s asynchronous post, is quite different from the skills entailed in attracting teacher’s attention in a classroom, and articulating a meaningful synchronous response. The experience can be intimidating for minority groups. Such bisection of participation can occur along ethnic as well as gender lines.

The second criterion of 'access' is problematic, in that technology access may in fact work to the detriment of minority groups. Though greatly improved from the last decade, the access to technology for minority groups is still relatively low [7]. It must however, be acknowledged as a critical factor when discussing 'learning with technology'. Recognizing potential problems with access may lead to creative rethinking. For example, the entire course reading material can be made available through online course reserve. Offering the use of lab machines to students who may not have a computer at home can enhance access. The very first virtual class discussion can be used to ensure that every student will be able to use a computer with internet access at least for a couple of hours for two to three times a week. Instructor can proactively help by pooling resources, and forming study groups to ensure that everyone will be able to go online.

.

The third learning criterion 'expectancy' is a behavioral trait. According to Vroom (1964) an individual's motivation is governed by his/her expectation of achieving a specific goal and the value that individual places on that goal. In the context of teaching and learning significant differences are reported in the expectancy values for diverse groups [9]. The level of expectancy is a major factor in determining a learner's motivation. In the context of a virtual classroom motivation becomes critical. Excellent technology, pedagogy, and content notwithstanding, the expectation of success, and the value the learner attaches to that success is a significant factor in the learning process. The inclusion of expectancy in the learner triad is an acknowledgment of motivational factors in the learning process. Although it is ultimately a learner characteristic, yet inclusion of the notion of expectancy can allow the instructor to be more cognizant of individual differences, and tailor the online conversation accordingly

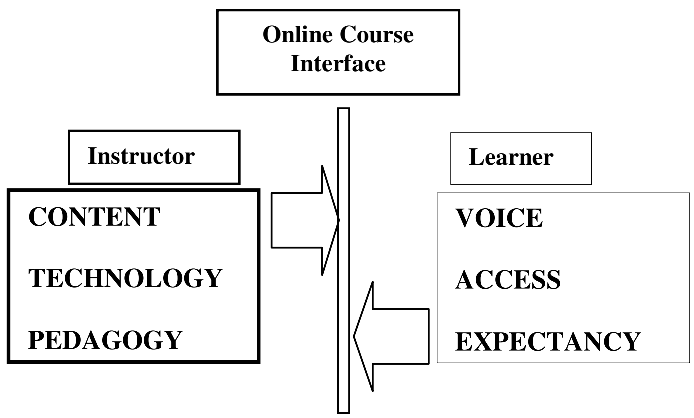


Figure 1: Pedagogical Model incorporating Instructor and Learner criteria

3. GENDER &TECHNOLOGY: COURSE STRUCTURE

The theme of gender and technology is addressed in numerous courses around the country in Women's Studies programs, and Technology departments. The course design outlined here specifically refers to WS4112 Gender and Technology Senior Seminar course that is currently offered by the School of Technology, Eastern Illinois University (EIU), Charleston. In its current form the course is taught in traditional classroom lecture format with a modest online component. This paper proposes to modify the course so as to adapt it for fully online delivery. In the process it is suggested that it will lose none of its content and analytical rigor, yet gain a degree of flexibility in affording greater voice to minority groups.

The course is designed to engage the students in an ongoing conversation about the various intersections of gender with technology. The students are exposed to scholarly literature about gender and education, with a focus on technology education. The course is a dialogue on the relationship between women and technology; it is grounded in historical context, navigates the current debate, and formulates a future prognosis. Key issues which have affected the mobility of women into technical occupational fields, are discussed.

Despite tremendous advances that have been made in the past twenty years, the gender gap in technology unfortunately, continues to persist [11]. Class discussion can focus on a number of arguments. These may include but are not limited to the gender differences in performance outcomes, the differential major choices in college, the difference in persistence in Science & Engineering (S&E) track, and finally the hiring differential in science and technology careers among male and female scientists, engineers, and technologists. Gender differences are persisting to this day in salaries, advancements [12]. as well as differential rates of attrition over the lifecourse [13].

The existing disparities will be discussed in the light of deep-seated belief systems, and prejudices in both genders, that have percolated down to a systemic level in our educational institutions. The course will be guided by an underlying theme of social construction of gender, where a constructivist argument will inform discussion and praxis.
The context for discussing women and technology education will be established by presenting introductory readings about the larger context of women and education. Reading material will include motivational theories as these relate to gender [8,

9], historical background necessary to inform the current debate, as well as a perspective on future outlook that turns the debate on its head in the so called 'boy crisis' [14].

Currently this course relies on face-to-face seminars and student presentations for vigorous debate. Porting the entire instructional experience online calls for sensitive course design in order to retain the robust discussion component of the face-to-face instruction. The following sections outline the pedagogical structure for the modified course. The course objectives are adapted from the existing course WS 4112G, the School of Technology, and the Women's Studies Minor, Student Learning Objectives at EIU, Charleston.

Course Objectives

Upon completion of the Women and Technology senior seminar, the students will be able to:

- Articulate the contributions of women to technology, science and engineering
- Identify personal assumptions concerning issues that affect women, men, and technology
- Articulate the influence of technology in altering the socio-economic landscape that has shaped gender role expectations.
- Critically analyze the context within which personal assumptions are formed that inform and/or cause prejudicial behaviors and actions.
- Develop the capacity to make informed, responsible individual and social judgments in the context of social responsibility
- Explore and imagine alternatives that exist to current ways of thinking about and acting on issues that affect women and men in a technological world.
- Demonstrate ability to research, and communicate with clarity and insight about gender issues in the context of technology

Instruction and Assessment

The course in its current form will be modified for fully porting to online instruction. The various activities, reading material, and assessment exercises for the virtual classroom will be presented in a weekly format in order to allow the students a sense of familiarity with the course routine. The lectures will be posted as PowerPoint and/or video files every week, and the students will be required to respond to specific questions, and discussion pointers, from these lectures. Each student will get a unique set of questions. Beyond academic honesty, individual questions can be customized, and used to enhance voice, opinion, and debate. Online format will enable periodical posting of grade points so the students can track their performance on an ongoing basis.

A list of tentative topics for this course is laid out in Table 1. Textbook can be limiting at best and a hindrance at worst in a constantly evolving landscape of scholarly debate The issues listed below will be explored by using current research literature, with readings from journals, book chapters, and /or electronic resources as determined by the instructor. The notion of social construction of gender will run through the course modules as its underlying theme. Table 2 illustrates the design of assessment activities from the types of classroom interactions, and learning objectives. Learning objectives are linked to learning goals according to Bloom's Taxonomy [15]. Each activity is outlined in sections that follow.

Table 1: Topics to be explored during the course of the seminar class

Theme	Issues
Identity	Ways in which gender intersects with other aspects of identity such as race, ethnicity, class, age, etc.
Education	Historical context, case studies, cultural and sub-cultural contexts of women's lives
Careers	Role of choice, opportunity, influences, and expectations (family, peer, society, media), recognition, credibility, accolades
Motivation and behavior	Attribution & expectancy theories, role of confidence, and self-esteem, gender role expectations
Contributions	Women's contributions to the disciplinary matrix of science, math, engineering, and technology
Stereotypes	Role of media, advertisements, image building, beauty vs. brains
Language	Obfuscation, masculinity in high tech literature
Body	Reproductive technologies: context, narrative, role, body image

Table 2: Learning interactions, objectives, and assessment activities

Type of interaction	Learning objective	Bloom's Taxonomy goal	Assessment Activity
Self	Reflection	Comprehension/ Analysis	Introduction/Adieu; Online Journal
Peer to Peer	Debate, Analysis	Comprehension/ Analysis	Chat/Discussion posts
Instructor to student	Critique, Discussion	Synthesis/ Evaluation	Online Journal Chat/Discussion posts
Instructor to group/class	Critique, Discussion	Synthesis/ Evaluation	Group Essay

Activity1-Introductions/Adieu: Each incoming student will be required to post a short introduction, along with a critical comment on any one idea from a given set of issues and/or themes. The students could choose from such words as

'stereotype', 'gender', 'minority', 'role', 'identity', 'language', and others. The students will essentially finish the sentence: "I think the word _____ means _______". At the conclusion of the course each student will be required to post a brief response to his or her own posting made at the beginning of the term. They could revise, modify, completely alter, or state that their opinions remain unchanged. In each case they must explain why, or why not. The exercise is designed to allow them to reflect and synthesize their learning experience during the course.

Activity 2-Online Journal: Online Journal is one of the current assessment activities. It is proposed to continue in the virtual classroom. During the course of the term each student will be required to prepare three "Reading response" assignments, where they will critically reflect on a self-selected reading, or a set of readings. Each journal (2-3 pages, double spaced) will be posted on the course site using WebCT. The journals should represent student's viewpoints about the readings, and their critique of the author or authors. The students will be encouraged to employ multiple readings for this assignment, so they could compare and contrast authors' arguments, and formulate their own position in the scholarly debate.

Activity 3-Chat/Discussion posts: Online chats or discussion sessions are designed as an integral part of the virtual classroom. The students will be required to post comments and questions, as well as post responses to a question posted by at least one other classmate. Their questions as well as comments will be based on the reading set for the particular week. Students will be required to use at least one external reading reference in their post. Extra credit will be awarded for the use of more than one external reference used in a meaningful way to support their argument.

Activity 4-Group Essays: Students will be required to form online groups of 3-4, and propose an essay topic based on their analysis of course material. The final deliverable of this project will be a term paper prepared as a group activity. The students will be required to complete this activity in three stages:

First stage will consist of formulating a topic of interest by internal online group discussions. The choice of topic is open, but it must relate to the course readings.

Second stage will require each group to work with the instructor, where they present an outline of the proposed essay to him or her. The outline will consist of at least four components, each presented by one team member:

- Introduction of topic, and contexualizing the problem
- Presenting existing scholarly work
- Presenting team's argument/ team's response/viewpoint/position
- Synthesizing a conclusion

The instructor will provide feedback, and critique at this stage in the project. The discussion session will be scheduled in half-hour long synchronous meetings conducted online with each team.

Third and the final stage will consist of the writing process, where each of the four sections will be addressed in detail. Particular attention will be paid to the use of referencing in presenting external material. The essay (10-15 pages excluding references, double spaced, APA format) will be posted online as the final deliverable for this activity.

Learning with Technology

Each assessment activity outlined above is adapted to work for two of three learner criteria outlined previously. The course structure is designed to encourage interaction and enhance 'voice'. The online journals, discussion postings, introduction/adieu, as well as group essays are all a means to allow diverse opinions, viewpoints, and debate to come through in a safe, proctored, collegiate environment. It is hoped that the 'voice' factor will not only be enhanced for women and minorities, but also see an overall improvement in everyone's participation.

'Access' forms the second leg of the learner triad. Admittedly, access is not a function of course structure or pedagogy. This is an external factor that cannot be controlled by the instructor. However, acknowledging it as a significant actor can help the instructor better understand some of the challenges his or her students may be facing in their online behavior.

'

'Expectancy' forms the third leg of the learner triad. Expectation of success is driven by motivation, and is not a function of pedagogy. However, as noted earlier, recognizing this variable is a necessary step in the instructor successfully reaching out to his or her diverse student body. An understanding of expectancy behavior can lead to more constructive instructor feedback to individuals' postings, discussions, and comments.

4. CONCLUSION

This paper has presented a virtual course design for Women and Technology seminar. An array of assessment activities is presented to allow for robust debate and discussion in online format. It is suggested that the 'teaching with technology' triad of content, technology, and pedagogy, must be informed by the 'learning with technology' set consisting of voice, access, and expectancy. Future research can empirically test the relative influence of each factor in this model. It is important to note that by considering only the teaching triad we might be overlooking a whole host of conditions that may impede a learner. The teaching triad can be uniformly applied to all learners, but the results may be far from uniform. The reasons for discrepancies are likely to be a function of the learner criteria set, which in turn would determine the success or failure of such a course. This paper proposes to adapt the learner criteria within the framework of a Gender and Technology senior seminar course.

Future research can focus on teasing out the learner triad, and look for possible additions and/or modifications within the learner dynamic. It is critical to acknowledge that the diverse learner characteristics will interact variably with the teaching input, and the course design must take this variability into account.

5. REFERENCES

[1] Rovai, Alfred P**. In search of higher persistence rates in distance education online programs**. Elsevier Science School of Education, Regent University, Virginia Beach, VA, 2002.

[2] Masters, K. and Oberprieler, G. **Encouraging equitable online participation through curriculum articulation**, Elsevier Ltd., 2003.

[3] Price, L. "Gender differences and similarities in online courses: challenging stereotypical views of women", **Journal of Computer Assisted Learning**, Vol. 22, Issue 5, 2006, pp.349-359

[4] Sullivan, P. (2002). It's easier to be yourself when you are invisible: Female college students discuss their online classroom experiences, *Innovative Higher Education, Vol 27, No 2, December 2002*. pp129-144.

[5] Koehler, M. J., Mishra, P., Hershey, K., and Peruski, L. "With a little help from your students: A new model for faculty development and online course design", **Journal of Technology and Teacher Education**, 12(1), 2004, pp.25-55.

[6] Caspi A., Chajut, E, and Saporta, K. **Participation in class and in online discussions: Gender differences**, Elsevier Ltd., 2006.

[7] Pachon, H. P., Macias, E. E., and Bagasao, P. Y. "Minority access to information technology: Lessons learned", **Michigan State University**, October 2000.

[8] Eccles, J. S., and Wigfield, A. "Motivational beliefs, values and goals", **Annual Review of Psychology,** *53*, 2002, pp.109-32.

[9] Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., and Midgley, C. "Expectancies, values, and academic behaviors" In J. T. Spence (Ed.), **Achievement and achievement motivation**, pp.75-146. San Francisco, CA: W. H. Freeman, 1983.

[10] Vroom, V. **Work and Motivation**. New York, NY: Wiley, 1964.

[11] Beizer, D. "Technology access equity and digital divides", **Technopolity**, accessed May 9, 2010 at http://technopolity.editme.com/technologyequity

[12] Lips, H. M. "The gender wage gap: Debunking the rationalizations", **Women's Media**, accessed Apr 5, 2010 at http://www.womensmedia.com/new/Lips-Hilary-gender-wage-gap.shtml

[13] Xie, Y. and Shauman, K. A. **Women in Science: Career Processes and Outcomes**, Cambridge, MA: Harvard University Press, 2003.

[14] Kohn, D. "The gender gap: Boys lagging", **CBS News, 60 Minutes**, 2003, accessed May 11, 2010 at http://www.cbsnews.com/stories/2002/10/31/60minutes/main527678.shtml

[15] Bloom, Benjamin S. **Taxonomy of Educational Objectives**, Allyn and Bacon, Boston MA, 1956.

New concepts in engineering education through e-learning

Vlasta RABE
Faculty of Education, University of Hradec Kralove,
Hradec Kralove, Czech Republic

and

Stepan HUBALOVSKY
Faculty of Education,University of Hradec Kralove,
Hradec Kralove, Czech Republic

ABSTRACT

This paper deals with new concepts in engineering education regarding with global trends in economical and information environment and their influence on engineering education, and also with possibilities of ICT deployment for integrating face-to-face and on-line learning. There is presented system approach in education and its impact on new chances of e-learning.

Keywords: Engineering education, role of teacher, role of experiments, e-learning, collaborative learning, quality of education.

1. INTRODUCTION

In the traditional approach not only to engineering education, the teacher lectures and assigns readings and convergent single-discipline problems, the students listen, take notes, and solve problems individually. Nowadays shifts demand on university education to new wave in tutorial, to ability work in team (e.g . at projected teaching), cope with changes, be flexible and innovate. In addition it is necessary to concentrate on scientific work and on quality them providing information and knowledge, which will be measured so, what students know, and how they can use it practically.

Nowadays modern information technologies are going to enhance development of new methods of searching, acquiring, organizing, processing, sharing and using of information from various sources and disseminate it to users. This paper is not aimed at giving a new method, a technic or a tool to solve current problems, but it is an experiment to look into the future. The applications enable automatic access to informations anytime and anywhere. By modern ICT it is able informations and knowledges share and effective use. Now it isn't problem access to computers and the Internet, but growing number of leaders in higher education see it more in necessity to increase of **information literacy**. It means the ability to find relevant informations accordance to needs.

2. CHANGING CONCEPTS IN ENGINEERING EDUCATION

Role of engineering education in the development of information and knowledge society is in active approach to learning and possibility to use own practical experiences in process of education.

In past time universities being asked especially to create and disseminate knowledges, now shapes demand on university education to new wave in tutorial, to ability work in team (e.g . at **projected teaching**), cope with changes, be flexible and innovate. Now these roles can be enhanced by promoting learning that ensures that people can take advantage of the information resources available to them. Such efforts can and should include both improving the ease of access to information and educating people to evaluate and use information effectively.

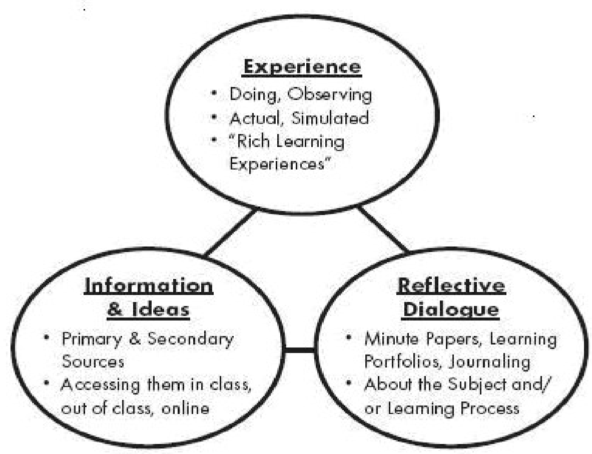


Fig. 1. A holistic view of active learning
(by John Wiley & Sons, Inc.)

Economic prosperity of univiersities will depend on quality of education and research and on their proactive behavior. Universities go to lifelong education, because in information society information processing and their sharing and presence takes part in everyday life. The focus in new concepts in European education system is changing **from learning to do (or know) towards learning to learn**. The concept lifelong learning also corresponds with the broad political efforts and initiatives on developing education related to ICT, in individual countries.

3. COLLABORATIVE LEARNING

Collaborative learning is based on the teachers help students respond to literature by taking a more active role in their own learning. The cooperative learning tradition tends to use quantitative methods which look at achievement, i.e., the product of learning. The collaborative tradition takes a **more qualitative approach**, analyzing student talk in response to a piece of literature or a primary source in history. Cooperative learning is defined by a set of processes which help people interact together in order to accomplish a specific goal or develop an end product which is usually content specific. It is more directive than a collaborative system of governance and closely controlled by the teacher. While there are many mechanisms for group analysis and introspection the fundamental approach is teacher centered whereas collaborative learning is more **student centered**. [5]

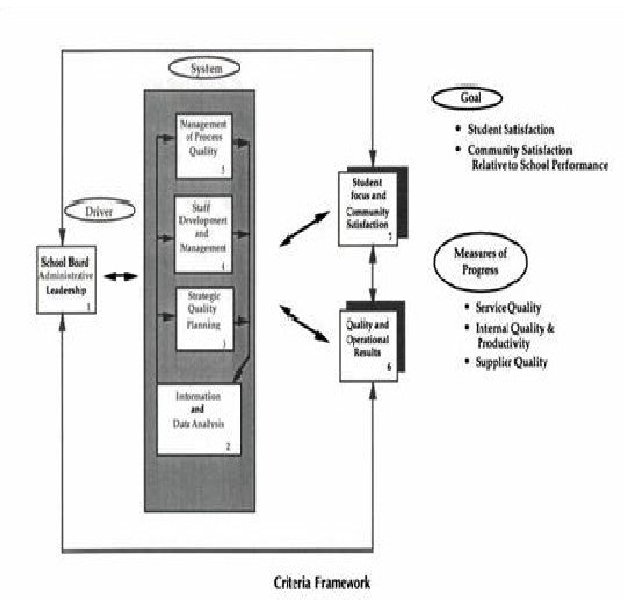


Fig. 2. Quality of education (by M. Baldrige award)

Active learning is classroom instruction that involves students in activities other than watching and listening to a lecturer. Working individually or in groups, the students may be called upon to answer questions, solve problems, discuss, debate, reflect, brainstorm, or formulate questions. Cooperative learning is instruction that involves students in team projects under conditions that meet several criteria, including positive interdependence (the team members must rely on one another to carry out their responsibilities) and individual accountability for every part of the project.[7]

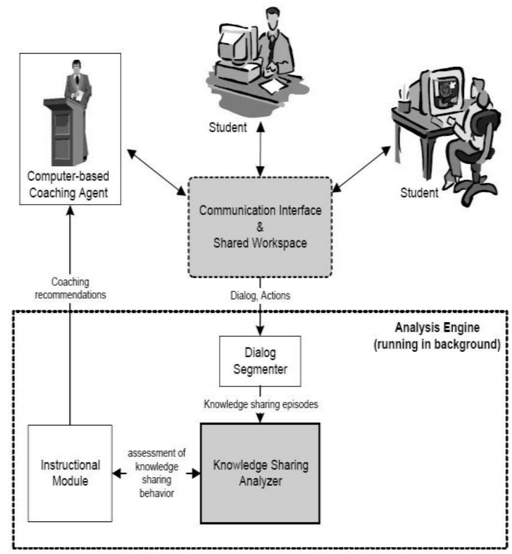


Fig. 3. Sharing knowledge by collaborative learning

4. DISTINCTION OF ENGINEERING EDUCATION

In the past century appears growing number of innovative programs and instructional methods and materials in engineering education. But the changes, that will move engineering education in the desired directions, may be grouped into four categories: [1]

- revisions in engineering curriculum and course structures
- implementation of alternative teaching methods and assessment of their effectiveness
- establishment of instructional development programs for faculty members and graduate students
- adoption of measures to raise the status of teaching in society and in institutional hiring, advancement, and reward policies

In addition important in engineering education are discussions about technical problems. It can be able by modern ICT, the informations and knowledges share and effective use. Now it isn't problem access to computers and the Internet, but growing number of leaders in higher education see it more in necessity to increase of information literacy. It means “the ability to know when information is needed, and to be able to identify, locate, and effectively use that information for lifelong learning and problem solving.”

System thinking

In connection with this fact system thinking together with system dynamism offers in education common frame for preservation cohesion, sense and motivation on all levels of education, at first. Second element is emphasis on active student cognition, which imbeds into tutorial new challenges and interests for learning, how it is common in experimental laboratories. These two innovation, gathered together, help to enhance creativity, inquisitiveness and life energy of young people.

There is necessary to systematically learn and use system thinking like common tool for everyday activities. Strong attribut of system thinking is there a connection philosophy, policy and culture with everyday work.

Brainstorming – making creative ideas

Brainstorming is a group technique for generating new, useful ideas and promoting **creative thinking**. It can be used to help define what project or problem to work on, to diagnose problems, remediate a project by coming up with possible solutions and to identify possible resistance to proposed solutions.

I tested this method in one new subject with students in distance form of study (teachers from practice). I gave them some material on web, and they together prepared lessons for whole course. We checked it out on tutorials. But, in my opinion, this way it is able only with experienced people.

Benchmarking in high education

Currently this method is used for measurement quality of education. We can find it e.g. on web page www.sibis-eu.org .

For some years now statistical indicators on Information Society have been central in the policy making process. This has been best demonstrated through the benchmarking exercise of eEurope Action Plan as a key activity. Having recognised this need and driven by the difficulties in obtaining reliable and appropriate statistics, the IST programme supported a pan-European research effort during Framework Programms. The prime objective has been to develop and make available methodologies, tools and new statistical indicators which can help remedy the deficit in this field. It is in this context that the SIBIS project was launched.There are at least two main reasons that make this document interesting. First, it is one of the few original attempts to have a coherent and comprehensive approach in measuring the information society. As such it is expected to stimulate further debate and research among the professional statistical community, leading to an improved statistical competence in Europe. Second, it provides a unique single source of data on real time which supports many of the new ICT research areas.

I see the advantage in application of benchmarking in collaboration among universities and learning from some best, like it is in business. To create a cooperative environment where full understanding of the performance and enablers of "**best in class**" processes can be obtained and shared at reasonable cost. [4]

5. INTERNET IN EDUCATION

Nowadays universities offer high quality opportunities not only in engineering education. ICT skills were obtained through formal education and training or more informally through using ICTs and gaining experience with them. High standards in mathematics, science, and informatics in most that countries were promising for the future supply of highly skilled professionals. The promising medium and longer-term supply of ICT professionals was associated with increasing numbers of young people in higher education and vocational schools.

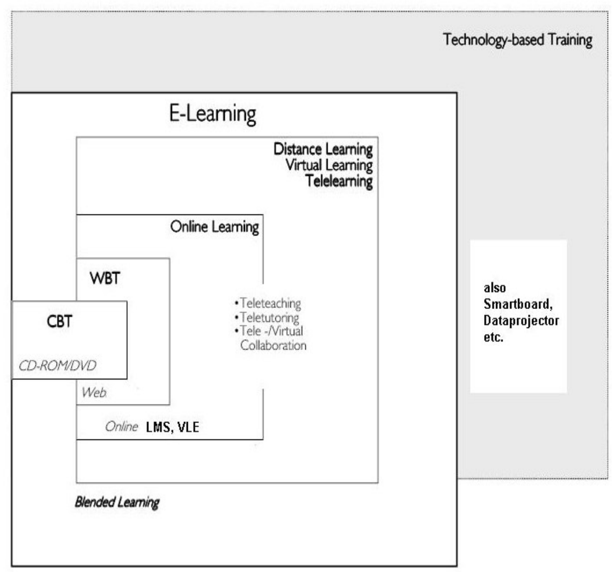


Fig. 4. Technology-based training (by Back, 2001)

One of modern learning methods is e-learning. All educators approach this new paradigm with varying degrees of enthusiasm and concern. It is important to consider both, the pros and cons, about on-line learning. So we can be better prepared to face the challenge of working in this new environment as well as embrace the new opportunities that it has to offer. In all countries rapidly changing information and communication technologies, and growth of ICT-related activities, in all sectors led to shortages of highly-qualified ICT professionals. Recruitment difficulties indicated imbalances between existing skills and company demands. ICT skills were obtained through formal education and training or more informally through using ICTs and gaining experience with them. Prospects for preserving and building ICT capabilities and skills was complicated by mobility of professionals. There are many valid reasons why online programs are rapidly becoming a popular form of distance learning in higher education today. The online environment offers unprecedented opportunities for people who would otherwise have limited access to education, as well as a new paradigm for educators in which dynamic courses of the highest quality can be developed. While online programs have significant strengths and offer unprecedented accessibility to quality education, there are also weaknesses that can pose potential threats to the success of any online program. Especially **in engineering education** I would prefer **blended learning**. [9]

In our university we are embarked in high-powered work on preparations choice subjects for combination forms studies in virtual educational environment WebCT, and for full-time forms of studies like supported materials. We also make up videoconference, but in my opinion not very suitable for education. [8]

In e-learning tutorials are a way, where teachers supplement on-line learning with a face-to-face component. Typically, a teacher will organise a time where students can come and see him or her, or arrange for students to work in a learning centre with assistance from a tutor. [10]

6. E-LEARNING - PROS AND CONS

It is important to consider both, the pros and cons, about on-line learning. So we can be better prepared to face the challenge of working in this new environment as well as embrace the new **opportunities** that it has to offer. First e-learning course was created in our university in the year 1998. Since the year 2001 our faculties embarked in high-powered work on preparations choice subjects for combination forms studies in virtual educational environment WebCT, and for full-time forms of studies like supported materials.

Inter – universities studies

University of Hradec Kralove has devoted its attention to the problems and issues associated with e-learning since the beginning of 1997. As early as then, voices could be heard calling for cooperative and collaborative ventures between university-level institutions in the creation of distance e-learning courses or unified study programs.

Interesting possibility for collaboration was identified at the e-learning in higher education on 2003 conference organized by university in Zlin, which involved the sharing of courses - including the relevant teaching staff, and providing these to the students of partner institutions, leading to the exchange of students through the intermediary of e-learning supported distance courses.

Since 2005 the first Czech virtual mobility has been realized in the project RIUS (the start-up of inter-university studies in a network of selected universities in the Czech Republic), in which cooperated three universities (Hradec Kralove, Plzen and Zlin). It involves the sharing of both the courses and teaching staff of the universities participating in this project possibility of absolving part of study programme at university. The courses were provided in a distance education form with e-learning support. This e-learning courses are organized in the form of an introductory tutorial in face-to-face of both teachers and students at the students` alma mater university, directed self-studies supported by a virtual educational environment and, as may be required, by further interim face-to-face meetings, combined with live examinations. The face-to-face meeting can be replaced by synchronous video-conferencing using the Internet. Depending upon prior agreement with a partner university, students can choose these courses within the context of their compulsory elective subjects. Through making the best use of the range of inter-university studies on offer to them, students can not only enrich their own study plans with topics attractive to them, but also get to know new educational methods and instruments, and have a share in the genesis of an expanded system of inter-university studies in the Czech Republic, enabling mutual sharing of study subjects and experts across university network.

Due to these projects much students were given chance to study at least one out of 164 subjects in a virtual learning space in academic year.

7. PROBLEMS IN ENGINEERING EDUCATION

There exists some specific problems especially in engineering scope:

- disparity between belived importand needs and amount of coverage in receives in the classroom
- lack of expertise

(Most people with a software engineering, managerial, and teaching skill are in industry or other mostly for financial incentives.)

- lack of relevat text and multimedial teaching materials

(Each book is different. It must be confusing to the technically oriented instructor to see such an incredible diversity of methods, approaches, and complex non - standard forms.)

- inadequate computer technique

(All computers, techniqe and materials should be closely associated with teaching of concrete subject)

8. Conclusions

In my opinion impact engineering education in information society depend on quality of lessons, learning material and especially on quality of teacher – each instructor must develop own style, using those techniques that best suits them and seem to achieve their objectives, but on-line learning cannot replace face to face education. In engineering education there is similar situation like in case education future teachers – there is needed direct dialog and practical experiences.

On base the experience it seems, that in full-time forms education, we must think about, if in all cases is necessary introduce e-learning.

In case of engineering education it seems, that blended learning, which combines on-line and face-to-face approaches (in accordance with our experiences), is more appropriate method of learning.

9. References

[1] Committee on Science, Engineering, and Public Policy: **Reshaping the Graduate Education of Scientists and Engineers**, National Academy Press, Washington, D.C., 1995

[2] **OECD: The measurement of scientific and technical activities**. "Frascati Manual 1980", OECD, Paris, 1981, p.26, 2.2.2.1 a "Frascati manual 1993", ibid, 1994, p.30, 2.2.2.1.

[3] **Commission on National Investment on Higher Education: Breaking the Social Contract. The Fiscal Crisis in Higher Education**. Council for Aid to Education, RAND Corporation, Santa Monica, CA, 1997

[4] Anderson, T. D. & Garrison, D. R. (1998). **Learning in a networked world: New roles and responsibilities**. In Gibson, C.C. (Ed.), Distance Learners in Higher Education: Institutional Responses for Quality Outcomes (pp. 1-8). Madison, WI: Atwood Publishing.

[5] J. Clifford: **Composing in stages: The effects of collaborative pedagogy**. Research in Teaching English 15 (1), 1981, pp. 37-44.

[6] V. Basili and A. Turner. **Iterative enhancement : a practical technique for software development .** IEEE Transactions on Software Engineering (1975).

[7] L.Newton, & L.Rogers: **Thinking frameworks for planning ICT in science lessons.** School Science Review, 84 (309), 113-120, 2003.

[8] V. Jehlicka, V. Rabe: **On-line courses of ICT training in WebCT environment.** ED-MEDIA, Vienna, 2008, ISBN: 1-880094-62-2.

[9] V. Rabe, V. Jehlicka: **E-learning and blended learning at the university.** ED-MEDIA, Vienna, 2008, ISBN: 1-880094-62-2.

[10] I. Cernak, E. Masek: **Possible approaches in implementation of electronic education at university.** 2009, ISBN 978-80-8084-431-8.

An approach to online support for effective Teaching and Learning, a study case

Eudes K Tshitshonu
Department of Mechanical Engineering

Faculty of Engineering and Technology
Vaal University of Technology
Vanderbijlpark 1900
South Africa

1. ABSTRACT

The gap opened by the brain drain due to global work force mobility, on the one hand and the speed at which new knowledge and technology are being developed on the other is growing wider every day. Paradigm shift is paramount; business models are being reviewed and revamped. Skilled hands are not the only feature in demand for the implementations of new technologies. Trained minds with sufficient background and the ability to investigate novel approaches are required even more. Stakes are higher today, as sustainability and awareness to the environment are top on the agenda in any new engineering and innovation endeavour.

South Africa has embarked on a venture to increase its delivery for more engineers, technologist and other essential career streams required to fill up the skill shortage gap. As key players, tertiary institutions extend themselves into enrolling massively with an objective to produce as many graduates as possible and contribute their bit to address the challenge.

The competency of candidate students in basic numeracy and literacy is a great debate topic. Mathematics, science and physics skills of applicant university students seem to leave a lot to be desired. An ongoing debate suggests the extension of the average three years stay for undergraduate university qualification to four years, which can be detrimental especially in an environment where the workforce is starved with skill shortage. As the relevance of the curriculum offered is being questioned, tools that are made available to assist the learning process are under investigation and their effectiveness assessed.

This paper discusses the interim results of an ongoing observation and study on relevance of the effective use of other teaching and learning modes afforded by technology, in the particular case of handling big class groups. A comparison of the pass rates between groups of students participating in traditional classrooms and those exposed to the blended learning mode that includes an online component.

Keywords: blended learning, embed questions, Moodle, online tutorial, pass rate

2. INTRODUCTION

Innovation, customisation, upgradeable business models, and new ways of organising work; these are some of the expected attributes required to keep competitive in the 21st-century economy. Individuals and organisations need to continually update their skills and find new ways of managing knowledge and information.

In 2000, the American Society for Training and Development (ASTD) and the National Governors Association (NGA) convened to form the Commission on Technology and Adult Learning [1], with a mandate to define and encourage a technology-enabled learning environment that would result in an engaged citizenry and a skilled workforce for the digital economy. The commission focused its attention on technology-enabled learning designed to increase workers' knowledge and skills so they could be more productive, find and keep high-quality jobs, advance in their careers, and have a positive impact on the success of their employers, their families and their communities. E-learning was spotted for its potential as a tool for reducing the costs of workplace-related education and training as well as its ability to offer universal access to the best-in-class learning content, as well as a wide variety of content available anywhere in the world. In traditional training delivery, the student needs to adjust to the training presentation schedule in order to grab the maximum benefit and is in most cases confined to the content as presented in class.

This study observes two groups of students registered for introductory applied mechanics. The first group is made up of fulltime students. The second is made of part-time students who can only afford two hours class a day after work. In an environment where students are exposed to labour intensive work prior to attending two hours class, in the afternoon, the maximum of their best performance is not to be expected. Human limitations, unresolved outstanding plants/workshops issues could take the toll out of them. In some instances they just cannot afford attending class, particularly in cases of unexpected shut downs or any other work related matter. Computer Based Learning (CBL) kicks in and reveals to be the alternative of choice. Wentling *et. al* present Karon's advocating for the convenience factor of well-designed computer-based training [2]. Karon argues that any well-designed computer-based training - whether it is local intranet based or delivered via the Internet - is more convenient than traditional instructor-led training or seminars, as self-paced Computer Based Training (CBT) courses are available when students are ready and able to complete them, not just when the seminar or class is scheduled or when the instructor is available. Class content becomes portable. The portability of the class content does not however overshadow the trainer's role though. Trainers need to redefine their role, as their work design and environment changes. Stemming from their traditional role trainers' responsibility, now span across instructional designer, through to instructional developer, trainer, and materials supporter. As an instructional designer, the trainer performs the initial analysis and instructional design tasks. They also include advice on course exercises and revision. As an instructional developer, the trainer writes course materials, exercises, auxiliary materials and develops overheads graphics, exercises, and so forth [3].

Modes of certifying the outcome or the learning process are required. Tests, examinations, quizzes and a host of different types of assessment methods are used to gauge and sanction the outcome as acceptable or not. While most education/training systems implement and agree that some form of assessment needs to take place, a paradigm shift is presented as to the very essence of assessments. Assess the learning or assess for learning? A dramatic change is happening in the assessment culture. Authors advocate for assessment as a tool for enhancing learning [4, 5, and 6] rather than a simple testing of learning outcome. This redefines the training experience as a project managed by the trainer in his/her role as the major facilitator connecting diverse stake holders. The student is one of the key stakeholders in the whole process and no actual training/learning is likely to happen without his/her involvement. It suits at this point to redefine the student's role in the updated setup.

3. ONLINE LEARNER AND TRAINING RESOURCES

As the training environment evolves, the traditional classroom contact is no longer the only and prime point of delivery. In distant/online learning environments, the trainer has less direct and sometimes nearly no real time interaction with the student. That makes it difficult to address issues where spontaneous classroom feedback would be of very much added value. In a classroom, the trainer can interact with students through a set of input/response sets that would at least serve an indication to whether the topic under discussion is accessible and understandable to students or not. Facial and body languages are part of the communication protocol and can facilitate the interaction particularly in small groups. Short improvised activities or spot assignments provide real time feedback and prompt an indication to grey areas where more emphasis might be required. Lack of involvement and/or motivation is likely to be addressed timely, making it easier to reach the learning objectives. The situation is completely different with out-sight students whose challenges are not discerned unless effective channels of communication are open and functioning. Students' attitude towards the training, the media of delivery (technology) their self motivation is of prime importance. Several researchers have identified individual characteristics that seem to describe a successful online student. Gibson [7] found that it is critical for distance students to be focused, better time managers, and able to work both independently and as group members, depending on the delivery mode and location of the distance course. Other studies suggested strong self-motivation, self-discipline, independence, and assertiveness as important characteristics of online students [8]. While, the onus seems to be entirely on the student to address their working discipline and attitude, the content and context need to be provided in a manner that will be conducive enough for effective delivery of the content. Clark [9] addresses the latter through multimedia, contiguity and modality principles. The multimedia principle suggests the use of graphics such as line drawings, charts, and photographs and motion graphics such as animation and video. Graphics need to be congruent with the learning content and add value to the learning content rather than overload and distract. Clark points out a common violation of the contiguity principle encountered in screen scrolling. Displaying graphics close or next to related text allows a better descriptive effect than seeing the graphics first and having to scroll further down for descriptive text of vice-versa. Learning occurs in individuals by way of working memory which is the active part of our memory system. Working memory capacity is needed for learning to occur. Learning becomes depressed when working memory becomes overloaded. If words and the visuals they describe are separated from each other, the learner needs to expend extra cognitive resources to integrate them. Contiguous display of visuals and words achieves the integration for students, leaving their focus to the learning content [10]. Audio inserts are praised for adding extra value to online materials and improves the learning experience as suggested by the modality principle [9].

4. TARGET GROUP

Two major groups of student were sampled for this study, fulltime and part-time working student all registered for Introductory Applied Mechanics. A blended delivery mode made of online resources and face-to-face delivery of class content was applied. Groups were subdivided as follows: groups 1 and 2 were made up part-time student combining study and work whereas groups 3 and 4 were made up of fulltime students. All groups had real time class presentations. Group 2 however, attended real time class simultaneously with group 1 but from a different location, from a different province in the country via video conference. Groups 1, 2 and 3 were exposed to online resources and group 4 had no online exposure at all.

Students in groups 1 and 2 freely chose to take their employer's offer to study and upgrade their qualification for a better position on the corporate ladder with the same employer after graduation. It was assumed for the purpose of this investigation that they all had a clear vision of the deliverable expected after the training, thus motivation and attitude towards the training were not considered to be a major concern. The same assumption was applied to groups 3 and 4, though most fulltime students were much younger compared to candidates in groups 1 and 2 and would show some lack of maturity as for their motivation to complete their academic programme. However, in a country where inequalities from the past are being redressed and a strong emphasis put on the benefit of education, the study was conducted under the assumption that fulltime students should also be motivated enough, with the right attitude to drive their success.

While the attitude towards the training might have not been a problem, the attitude towards the media of delivery might on the contrary have been a cause for concern. The digital divide is still a challenge for many today, especially as long as basic life can be conducted without much to do with computer devices. That is still true for a great deal of communities in Africa. Some students would still be intimidated in using computer. Our role was to create, foster and work as facilitator to make the platform operational in the most possibly optimal way.

5. BLENDED LEARNING: STUDY CASE

Students were introduced to theoretical concepts as well as respective practical aspects and implementations through normal, traditional face-to-face contact classes. A set of worked examples and tutorial activities were introduced. After a topic had been discussed during contact sessions, groups 1, 2 and 3 students would be offered the opportunity to revisit the content material and attempt tutorials online. Selected worked examples and additional tutorials were expanded in a way to dissuade students from regurgitating repetitious memorised routines and rather highlight fundamentals as well as the underlying theory. The process was facilitated through the Moodle (Modular Object-Oriented Dynamic Learning Environment) Course Management System.

The user friendly Moodle environment has been investigated as potential piece of the support system to address the shortfalls and enhance delivery of the teaching and improve the learning experience for large groups. The interface is easy to learn for both the student and the teacher/lecturer. Basic keyboard, typing and click of a mouse button are sufficient even for a student without prior computer exposure to start using the interface. A quick introduction and training session would be used to kick-start the process and take away the intimidation of lacking computer skills. Selected class materials could be made available through the “add resources” command. The facility was used as a virtual notice board where essential highlights were displayed for reference. Class notes, sample worked solutions, link to the internet are amongst features used in the implementation of the online content offered in mechanics classes. The command enables the design of web pages thus enhancing the delivery of learning content with multimedia materials. An assessment interface makes it possible to upload quizzes, and lesson paths.

A powerful learning environment is characterized by a good balance between discovery learning and personal exploration, on the one hand, and systematic instruction and guidance, on the other [11]. The activity functionality in Moodle was used to make that powerful learning environment possible. Use was made of embedded questions, which accommodate calculated and worded/phrased results as well as multiple choices in a single question format. Graphs/pictures could be added as required in most engineering problems and care was taken to format the graphs and descriptive pictures as recommended by the contiguity principle [9, 10]. Questions were designed to allow students to discuss theoretical contents that might have seemed not straightforward to discern through class presentations. Students were hence led to learn to appreciate working from fundamentals and check the way they approached problems against the backdrop theoretical essence, rather than just do meaningless engineering calculations.

Student would be allowed to attempt a tutorial activity with immediate feedback. Moodle questions can be programmed for automatic grading, allowing one to see cross and check marks showing successful and erroneous attempts. The feature was pointed out to students as an opportunity to engage into constructive debates and discussions. Cross (erroneous results) marks would foster the platform for students to revisit their fundamentals and theoretical contents rather than engage in a blind chasing of the satisfaction of check marks. A particular emphasis was placed to prevent students from the temptation of playing guesses especially in the cases where a feedback provided expected answers. It is common practice in situations where textbook provide final expected answers that students target answers instead of discussing their way to correct results and end-up missing out on the learning opportunity. Subsequent attempts could be allowed for student to check the outcome after their revisit on fundamentals, through additional readings, discussions with peers as well as consultation with the lecturer. Learning was expected to blossom out of the described iteration process.

6. CONCLUSION

This paper was intended to discuss the interim results of an observation being conducted on relevance of the effective use of other teaching and learning modes afforded by technology, in the particular case of handling big class groups. A comparison of the pass rate between groups of students participating in traditional classrooms (group 4) and those exposed to the blended learning mode that included an online component. Admission to the final examination, at the end of a study term is not automatic. Students are exposed to a set of tests and need to pass them to earn their admission to the final exam. The pass rate is calculated as the ratio of students passing out of the total admitted into the final exam. The following observations were made after the second semester 2009: 43 % of group 1 students were allowed to write the final exam and 91 % were allowed from group 2. Groups 1 and 2 produced 100% pass rate. 30% of students from group 3 qualified to write their final exam, while group 4 allowed 37 % of its population into the exam. However, 75% of group 3 students allowed into the final exam passed whereas 61% passed from group 4. It seems to transpire that students exposed to online content (groups 1, 2 and 3) beside normal contact classes seem relatively better prepared. Results from previous observations showed 31% of the population was exposed to additional online support and could generate 53% of the overall pass rate on the sampled class population. The result is even more interesting for group 2 that joined classes through video conference, where students had no physical contact at all. The speculation is to whether the lack of physical contact with the lecturer spurred them into more personal initiative, better time management and effective use of all resources made available online. More observations are still being conducted before significant conclusions could be validly drawn and set as an indication for conducting big population classes.

7. REFERENCES

[1] Stephen E. Merrill et. al, **A Vision of E-Learning for America’s Workforce**, Report of the Commission on Technology and Adult Learning, 2001

[2] Tim L. Wentling, Consuelo Waight, James Gallaher, Jason La Fleur, Christine Wang and Alaina anfer, **e-learning - A Review of Literature Knowledge and Learning Systems Group**, University of Illinois at Urbana-Champaign, 2000

[3] Abernathy, D. J. (1998). **The WWW of distance learning: Who does what and where?** Training and Development 52 (9) p. 29-30.

[4] Filip J.R.C. Dochy and Liz McDowell ,**Assessment as a tool for learning, Studies in Educational Evaluation**, Vol. 23, No. 4, pp. 279-298, 1997

[5] Filip Dochy, **A new assessment era: different needs, new challenges**, Research Dialogue in Learning and Instruction 2 (2001) 11–20

[6] M. Birenbaum, K. Breuer, E. Cascallar, F. Dochy, Y. Dori, J. Ridgway, R. Wiesemes (Editor), G. Nickmans (Editor), **A learning Integrated Assessment System**, Educational Research Review 1 (2006) 61–67

[7] Gibson, C. (1996). **Toward an understanding of self-concept in distance education**, American Journal of Distance Education, 10(1), 23-36.

[8] Hardy & Boaz, 1997; Baker, 1995), quoted by Tim L Wentling, 2000 – [2]

[9] Ruth Clark, **Six Principles of Effective e-Learning: What Works and Why, The eLearning Guild’s. Practical Applications of Technology for Learning**, LEARNING SOLUTIONS e-magazine, September 10, 2002

[10] Sutherland, G. **A curriculum framework for an Introduction programme in the National Diploma: Engineering at the Vaal University of Technology**. University of Stellenbosch (PhD Curriculum Studies), 2009

[11] Sabine Dieric and Filip Dochy, **New lines in edumetrics: New forms of assessment lead to new assessment criteria**, Studies in Educational Evaluation 27 (2001) 307-329

Planning and the Novice Programmer: How Grounded Theory Research Can Lead to Better Interventions

Jonathan Wellons, Julie Johnson
Department of Electrical Engineering and Computer Science
Vanderbilt University
Email: {jonathan.wellons, julie.l.johnson}@vanderbilt.edu

***Abstract*—Planning is a critical, early step on the path to successful program writing and a skill that is often lacking in novice programmers. As practitioners we are continually searching for or creating interventions to help our students, particularly those who struggle in the early stages of their computer science education. In this paper we report on our ongoing research of novice programming skills that utilizes the qualitative research method of grounded theory to develop theories and inform the construction of these interventions. We describe how grounded theory, a popular research method in the social sciences since the 1960's, can lend formality and structure to the common practice of simply asking students what they did and why they did it. Further, we aim to inform the reader not only about our emerging theories on interventions for planning but also how they might collect and analyze their own data in this and other areas that trouble novice programmers. In this way those who lecture and design CS1 interventions can do so from a more informed perspective.**

***Index Terms*—Novice Programmers, Planning, Qualitative Research, Grounded Theory**

I. Introduction

There is much research in the area of self-regulated learning and its effects on student performance. Students who report exercising such skills as goal setting, planning, self-monitoring and self-evaluation experience higher levels of success and satisfaction then students who do not [1]. For programmers, planning is one of the first critical self-regulatory skills they will need. Early programming experiences are defined by a novice's ability to engage the complex cognitive process of problem solving while employing the metacognitive processes of self-regulation. While advice in teaching problem solving abounds, only recently have we seen developments in software tools and learning modules that address the process of planning. The overall goal of our research is to design activities and scaffolds to teach and support the metacognitive skills that novice programmers need to achieve early success in programming. The focus of this qualitative study was to use a systematic approach to observe and explain the process of planning among novice programmers at an undergraduate university. We began by analyzing interview data with the aim of developing emergent theories. Such theories and the subsequent results of testing (not a part of this study) will lead to the development of tools to help students improve and refine this critical self-regulatory skill. More specifically we asked, "What is the theory that explains the process of planning among novice programmers with no formal instruction in making such plans?" and "How could such a theory inform the construction of scaffolds and learning tools for novice programmers?".

In order to address these questions we employ a grounded theory study. This methodology is common in the social sciences and is appropriate when attempting to develop the groundwork for theories while avoiding pre-existing biases. Our objective is to produce data which can be qualitatively examined for connections between novice programming planning, causes or effects of their planning skills and other habits or tendencies on the part of the students.

II. The Role of Planning in the Programming Process

Developers use plans in large-scale projects to model a program at a manageable level of abstraction [2]. It is universally accepted that programming successfully requires planning and many different strategies are currently in use. The general model first put into wide use is the waterfall, consisting of distinct stages of requirements, design, implementation, testing and maintenance [3]. More flexible models gradually came to be applied to software beginning with the idea of iterative and incremental development, a cyclic process that applies elements of the waterfall model multiple times allowing for adaptation [4]. In the past few decades many of the same principles reappeared in slightly different forms, such as the spiral model [5] and more recently extreme programming, agile programming and test-driven development. Extreme programming is characterized by frequent release cycles, pair programming, extensive unit testing and flexible schedules [6], [7]. Agile programming emphasizes open and frequent communication, adaptability, customer stakeholders and cross-functional teams [8], [9]. The principle of test-driven development is to first institute a unit test before coding each feature or bug fix [10]. Numerous other ideas have been practiced such as cleanroom software engineering which was developed to provide reliable and verifiable software [11] and lean software development which aims to eliminate waste in all forms (excess bureaucracy, requirements, code, delay, etc.). It is clear that in order to participate in the design, management, development and testing of large-scale projects, students must develop the requisite planning skills exhibited by productive computer scientists and engineers.

III. Research on Planning

In his 1986 article Soloway called for a redesign of Computer Science curriculum to include the explicit teaching of problem solving skills which included planning and goal setting. He noted that expert programmers drew from a library of canned solutions to form a template or starting point for their solutions. Soloway proposed using goal/plan language for teaching introductory programming thus making the role of plans and goals more explicit to the novice from the outset [12]. In [13] it was demonstrated that teaching programming strategies and planning was possible, did not increase time needed for instruction and could be measured through written assessment.

In response to these reports, several tools have been designed and tested in an effort to support planning and strategy selection using the expert programmer's approach as a model. In [14] 25 undergraduate students in two randomly assigned groups received training in planning by one of two methods. The treatment group used an intelligent tutoring system (ProPL) to implicitly scaffold planning through the use of prompts while the other group received "click-through" text describing planning and strategies for programming. Their results demonstrated the value of a scaffolded approach to teaching planning skills as the ProPL group fared better in the assigned programming tasks than did the other group. In [15], [16] the authors describe two object oriented programming languages (Visual Plan Construct Language and Web Plan Object Language) designed to teach programming to novices through plan management and integration. These languages facilitate the Plan-Object Paradigm, an approach that gives context to programming objects by allowing students to create a plan first and then use that plan to create working programs.

In a more broad approach, [17] explored scaffolds for scientific inquiry and the needs of learners engaging in new and complex work processes, an appropriate description for novice programming. The authors evaluated Symphony, a software tool meant to scaffold the learner's planning activities. They used the artifacts created by the tool to analyze the complex process of scientific inquiry, identifying the needs of learners that could be further scaffolded. In this way, the tool itself became a means by which further research could be pursued.

Despite these advances in supporting the novice programmer, seventeen years after Soloway's article, [18] reported that strategies and plans, though crucial to learning outcomes in introductory programming courses still receive much less attention than language related knowledge. They also noted that the questions of why and how different strategies emerge and how they are related to underlying knowledge was still an open question.

Introductory programming classes differ in their emphasis on planning. Those that encourage it often do so in different forms. One class may teach the writing of program comments in advance of any code and a different class might teach top-down modularization. Students exposed to a limited style of planning or not exposed to planning at all may experience deceptive initial success with small projects but have difficulties later. Basic programming tasks are notoriously difficult for students to master and many potential majors leave computer science or fail to gain essential skills (examples from Australia, the US and the UK in [19], [20]). Given that programming is an essential skill in many engineering disciplines (many with a shortage of employees) as well as advanced coursework and that an enormous diversity of educational approaches are employed, it is natural to ask if there is a better way to teach planning.

Kuhl and Goschke [21] proposed a model for self-regulated learning which includes the steps of goal setting and planning. Their model was recursive; learners would return to the tasks again and again as they experienced internal feedback from the products they generated. This recursion is made more overt in the programming experience as learners receive explicit feedback from the compiler or debugger used while building a program. Error messages and programs that fail to terminate are external signs to the learner that a change is required. With few experiences to draw from, reflection is often limited to the code itself and incremental changes are made in an effort to improve the outcome. Changes often lead to incremental success which in turn contributes to progress. When progress leads to program completion this abbreviated reflective cycle, fully contained within the coding exercise, becomes a template for future programming assignments. As programming problems become more complex progress may be significantly slowed or halted signaling the learner that reflection on the larger plan may be in order. Whether this indication is taken up or ignored, in the presence or absence of an initial plan constructed by the novice, may yield some insight into the initial planning process.

IV. The Use of Grounded Theory as a Method of Analysis

Grounded theory is a form of qualitative research based on the formation of theory from data. Open-ended interview questions are posed and data is collected in an effort to generate theories about the domain in question. This methodology can be described in five steps: 1) gathering of data, 2) open coding whereby researchers assign discrete codes to the qualitative data, 3) grouping of codes into concepts and identifying one or more important concepts that merit reexamination of existing data and possibly further data collection to facilitate the building of a model, 4) axial coding, meaning the construction of categories that highlight the relationship between concepts and 5) the suggestion of one or more theories which describe the relationships. Grounded theory researchers cast a wide net to capture a diverse and multi-dimensional data set which may be fertile ground for theories [22]. Making every attempt to avoid the influence of prior theories or other constructs, researchers allow the data to form the theory rather than using existing theories to code and categorize the data.

As practitioners in the university system, we routinely gather data from our students in an effort to gauge both the effectiveness of our teaching and the impression it leaves on the students themselves. Performance data coupled with course evaluations serve as a reliable record of the course outcomes. Open-ended questions are often posed to students in an effort to collect qualitative data to inform future modifications to the course ("What if anything would you change about this course?") or to elicit reaction to recent modifications ("How did you use the online planning modules when completing programming assignments?"). When we use this data to make course refinements we are using some of the techniques formalized in grounded theory.

We synthesize this data with performance data, anecdotal evidence and past experience to guide our subsequent steps. This type of approach is the basis for such research methods as design based research or action research. By contrast, grounded theory formalizes the process of data analysis and produces a theory or set of theories that can then be used to develop course modifications, controlled experiments or future research in a broader context.

Qualitative research has been used to investigate novice programming, although grounded theory has not been applied to planning specifically. Interviews and coding are used in [23] to investigate how non-major programmers conceptualize Java concepts. To investigate if some elementary programming tasks are more difficult than others, [24] searched for novice programmer bottlenecks in object oriented programming. Students were observed during labs and their affective states and behaviors were coded. Compiler error logs coupled with interviews were used in [25] to track the most common errors made by novice programmers.

Like those involved in the BRACElet Project and the others, we believe that challenges to CS curriculum should be approached as research problems requiring established research methods [26]. We chose to use the grounded theory approach for several reasons. First, we wanted to use a method of data collection that was already familiar to us and to many practitioners in the field of computer science education and would not require additional equipment or instrumentation. Second, the systematic approach to data analysis appealed to us as computer science educators, acting as a segue to other more unfamiliar qualitative research methods. And finally, we wanted to demonstrate how student interviews could be used for more than just the course on which they were commenting, but could yield insights into more general aspects of the programming experience.

V. Method

Our sample consisted of volunteers from three sections of an Introduction to Engineering class offered at a research university in the United States. The course, required for all freshmen engineering majors, was presented in one four-week module consisting of approximately 14 one-hour lectures. This course introduces the fundamentals of programming (using MATLAB) within the context of cryptography. The general idea behind the course is to familiarize students with problem solving techniques and tools (such as MATLAB and Excel) while giving them a snapshot of the field of Computer Science and some of its practical applications. The volunteers came from varying backgrounds and were not all necessarily declared Computer Science majors. They were compensated for their time and reflected approximately the same gender breakdown as the class. None who volunteered were denied inclusion in the study. Because the goal of initial data collection is to gather as many different stories and experiences as possible, thus saturating each category with explanations and examples, random sampling is not as critical in grounded theory. In fact, in our discussion section we describe future data collection that will involve theoretical sampling–the selection of data based on the potential to represent the core theoretical constructs being studied.

In order to gather accurate data, volunteers' names were removed from their interviews. We asked open-ended questions and students were encouraged to discuss any aspects of their programming experience that they deemed meaningful. The interviewer gathering the data, held office hours and guest lectured in two or three class sessions of each of the three sections to develop a rapport with students and introduce the study. All interviews were voice recorded for later analysis. Volunteers were encouraged to describe their experiences during several assignments, what type of plan they created, how detailed it was, and how it was adapted. Subjects were questioned about their background in programming, their hobbies and other details that could lead to a planning theory. Several hundred codes were derived from the data, which were in turn placed into 14 concepts. Finally, we organized the concepts into five categories which naturally suggested the three theories that comprise our results, as detailed in Sec. VI.

VI. Results

Concepts are derived organically, based only on the codes that are present. The interviews explicitly asked for the content and complexity of the student's plan. Examples include "I made a list of tasks," "I just started writing" and "I wrote the comments then filled in the code." These are clustered in the *Initial Plan* concept. In addition, we asked how the plan performed when the student attempted to implement it and whether the plan evolved. One subject reported "it all fell in pretty nicely." Another said "I realized I didn't need some things. I had difficulty making the alphabet [substitution cipher] work." These are coded in the *Plan Adequacy* concept.

We invited students to describe their programming process. One subject reported "I had a hard time keeping track of variables for rows, columns, indices and so on." Another subject "... relied on old programs and examples." Other students reported difficult language features and the mechanical details of their run-debug cycle. These codes are grouped into the *Coding Process* concept. Subjects were asked about sources of help during the programming assignments. Many reported seeking aid from classmates and friends who were advanced engineering majors or had previous programming experience. Many students used Google, the MATLAB online or built-in help, the professor or the textbook. These codes are clustered into the *Help Sources* concept. All subjects are asked about their debugging process. One student said that after working on a frustrating bug "I took a day off to clear my mind, then returned to see if it was right." Some students reran their program after every newly added line to check for errors. Others ran it only when they believe the programming was finished. A handful of students used debugging output statements. Considerable diversity was present in the types of test cases used. Some students used only the cases given in the assignment, others gave random test cases. These codes are in the *Testing* concept.

TABLE I
EXAMPLE CODES AND THE CORRESPONDING CONCEPT

Sample Code	Assigned Concept
Initial plan was a short, vague list of tasks	Initial Plan
Plan failed because MATLAB doesn't handle long strings	Plan Adequacy
Program was constructed of fragments adapted from in-class examples	Coding Process
Searched for help on Google, e.g., "how to write a for loop"	Help Sources
Wrote program in 3-4 parts which were tested separately and combined at the end	Testing
"My goal was to receive a B or an A"	Goals
It took 2 to 2.5 hours to finish. Student had thought it would take 1 hour	Time Needed
"First day was too much, second day I started to understand, it clicked on the third day"	Class Experience
Programming background consists of using the Starcraft map editor	Programming Background
Had some advanced Math, "not good at it" but "loves it"	Quantitative Background
Calls self a "number cruncher" in everyday life	Hobbies
"Why go for the extra credit when I don't understand the basics"	Ambition
Rather than write a time-consuming brute-force program, he solved the permutation puzzle visually out of 362,880 possibilities	Lateral Thinking
Influenced lab partners to use pseudocode in the future	Personality

Subjects were asked about the type of goals they set for the assignment. Some students focused on grades, one reported "a B or an A." Another said simply "to finish." Other goals were "to finish before the weekend" and "initially I just wanted it to work, but later I wanted to satisfy my intellectual curiosity." Several students reported they enjoyed the assignment and no external motivation was necessary, but only one reported that learning MATLAB was his goal. These codes were placed in the *Goals* concept. Student were asked how much time they took to complete the assignment. Most responses were between 1.5 and 3 hours. Subjects typically found it took longer than they had expected, although there were exceptions. All codes related to how long the assignment took and how this compared to the student's prior expectation were combined in the *Time Needed* concept. Students were prompted for information describing their educational experience in the course which we classified under the *Class Experience* concept. One reported, "the scavenger hunt [assignment] was fun because cleverness was required." Another response: "Oh the hash code, that was frustrating!" Other students reported that the class moved too fast or that the examples were not related. A subject said that he was lost on the first day, began to understand on the second day and "it clicked" on the third day. We developed another concept measuring ambition from the reactions to a cryptography assignment which required students to choose between four algorithms with differing degrees of difficulty. Each choice was accompanied by a maximum possible number of points, ranging from 110 for the most complex algorithm to 87 for the most straightforward. Several intermediate choices were also given (such as input restrictions or user interface affordances) that could increase the point value of the attempt. A student's choices could be perceived as a measure of their self-confidence. Many students aimed low. One reported, "why go for the extra credit when I don't understand the basics." Others chose combinations over 100, but not the maximum possible. These codes are placed under the *Ambition* concept.

The interviews included questions about each subject's background in programming to discover a relationship between a student's previous experience and planning or assignment success. Many students in the sample had little or no background in programming. The most extensive backgrounds came from AP classes in Java and toy problems on a educational website. Another subject had learned MATLAB and C++ over the summer. A subject had used a script-based map editor for the game Starcraft. These codes were grouped in the concept of *Programming Background*. Subjects also discussed their background in quantitative studies. Many students reported enjoying and excelling at Math. One said "Math is my best subject." A handful of students were ambivalent about Math, "I went up to AB Calculus because BC was like boot-camp." These codes were clustered in *Quantitative Background*.

Students were asked to describe their hobbies. Many students were interested in strategy, board, card and video games. Only one reported sports. One reported, "Piano, writing poetry and Chess." Guitar playing was mentioned. These codes were grouped in the *Hobbies* concept. In the course of describing their problem solving plans, students often revealed insightful solutions. One problem required decrypting a scrambled message with nine factorial $(362,880)$ possible keys. Rather than write the brute-force program suggested in class, several students were able to crack it with pencil and paper, or use creative shortcuts that reduced the complexity of the program to write. One student solved this puzzle in Excel using built-in functions. We label these codes with the concept *Lateral Thinking*. During the student's description of their problem-solving techniques or group work, aspects of their personalities were revealed. One strong proponent of using pseudocode reported that she had influenced her teammates to use it on the following individual assignments. When codes of this type became apparent, they were classified in the *Personality* concept.

At this phase of the project, 14 concepts emerged as natural partitions to the codes as shown in Table I. We then transition to the axial coding step and aggregate concepts into categories based on similarity. Five categories emerged as ideal clusterings as shown in Table II. The data organization is bottom-up and is reflected by placing the child data on the left and the parent data on the right.

VII. Emerging Theories

Our goal in grounded theory is to examine the apparent connections between categories to suggest theories which can explain the data before us. The process of selective coding requires the selection of a core category. Connections are then studied in an attempt to define the relationships between all other categories and the chosen core category. From this rich set of codes and concepts, many relationships are possible. For example, connecting the *Planning* and *Programming Methodology* categories we found that plans that included testing related to shorter total time spent on the program. Relationships between *Planning* and *Goal Setting and Achievement* included a connection between those who planned and the extent of their ambitions in the program. Other connections that arose include the relationship between students who use lateral thinking and a lower frustration level during their programming experience. To expand any one of these relationships into a working theory would require theoretical sampling followed by further coding in an effort to saturate the data concerning the categories being related. This exercise would allow us to strengthen the proposed theory.

Ideal theories will not only be supported by the data, but potentially lead to research to enhance pedagogy for novice programmers. In keeping with the conventions of grounded theory, we want to avoid the problem of existing theories or claims influencing our analysis of the data. We do however turn to existing literature to identify gaps that might be informed by our work.

Expert planners have amassed a library of templates that they can draw from and flexibly apply to the problem at hand [12]. Among novice programmers, for whom such a mental library most likely does not exist, we do not know what form the planning process takes. Investigating the data suggests that novice programmers borrow ideas from their areas of relative expertise. Students expressly referred to their math knowledge or experiences writing papers when describing the origin of their plans. This leads us to our first candidate theory:

Theory I: Novice programmers attempt to adapt problem-solving strategies from other domains, such as mathematics or essay composition.

To guide us to the next theory, we investigate how to build a scaffold that can serve in the place of this expert library of templates until it can be established. Based on the data, students who wrote pseudocode were more likely to perceive success on the assignment, have higher ambition and have an adequate plan. This is true regardless of whether the student had a programming background, leading us to consider that the traditional skill of pseudocode may be the only scaffold needed for first-time programmers, at least until they have successfully completed a few assignments. Furthermore, effectively using pseudocode is teachable which leads to the following theory.

Theory II: Planning by means of pseudocode is achievable for novice programmers.

Equally important to CS1 lecturers is how to develop the feeling of perceived success in novice programmers. Based on various comments about the quality of a student's learning experience which show a relationship between planning and perceived success in the class as a whole, we formulate the final theory:

Theory III: Students who planned their programs are more likely to report a positive experience in the class.

VIII. Limitations and Future Work

Within the scope of grounded theory which is the production rather than the testing of theories, this work has several areas for further development. The students were from a single class intended for students seeking an engineering major. Thus the results do not immediately apply to either majors or non-majors in computer science because it is not known if both of these sets are well represented in the data. Furthermore, the sample of students is somewhat small and self-selected. The next step is theoretical sampling, which deliberately chooses samples in order to diversify the data set. The authors are likely to theoretically sample different backgrounds and expected majors, as well as a class with a different style of teaching and programming language. During theoretical sampling, the original data will be added to rather than replaced.

Practitioners divide research on teaching into two kinds: "What is" and "What works". "What is" research focuses on observations about current conditions and processes in the learning environment. "What works" research tests and measures alternative teaching practices. This is a "What is" project attempting to suggest relationships between novice program planning and other concepts for further study.

Outside of using grounded theory to qualitatively produce theories, future work consists of two kinds: refinement of these candidate theories and the development of "What works" projects to realize the benefits of proving or disproving these theories. Having established three candidate theories, the next phase of our research is to collect more qualitative data to strengthen or deny each of them. The strengthening of Theories I and II could lead to guidelines for developing learning modules and subsequent scaffolding designed specifically for the novice programmer while further exploration of Theory III would naturally lead us to a quantitative study to verify the proposed relationship. Each candidate theory has the potential to affect how we teach the metacognitive skill of planning and the emphasis that we place on that exercise.

IX. Conclusion and Summary

In this paper, we have applied grounded theory to interviews with novice programmers about their first programs. Through the principles of grounded theory we have coded, conceptualized and categorized the interview data. We have elucidated the connections between categories to generate several plausible theories that explain the data. Three candidate theories are

TABLE II
CONCEPTS ASSIGNED TO EACH CATEGORY

Concepts	Assigned Category
Initial Planning, Plan Adequacy	Planning
Coding Process, Help Sources, Testing	Programming Methodology
Goals, Time Needed, Class Experience, Ambition	Goal Setting and Achievement
Programming Background, Quantitative Background	Background
Hobbies, Lateral Thinking, Personality	Personal

proposed in this project, based on the observed relationship between planning strategies and self-reports of programming experience. The theories are, 1) Novice programmers attempt to employ problem-solving strategies from other domains which are more familiar to them, 2) Pseudocode-based planning tends to be a relatively successful strategy for novices and 3) Program planning leads to a positive reported class experience.

X. Acknowledgments

We gratefully acknowledge the support of the Center for Integration of Research, Teaching and Learning project (NSF Grant No. DUE-0717768) and the Qualitative Research Methods Workshop (NSF Grant No. DUE CCLI 0923592). Without the support of these two projects, this research would not have been possible.

References

[1] B.J. Zimmerman and M. Martinez-Pons, "Development of a Structured Interview for Assessing Student use of Self-Regulated Learning Strategies," *American Educational Research Journal*, vol. 23, pp. 614–628, 1986.

[2] C. C. Yu and S. P. Robertson, "Plan-based Representations of Pascal and Fortran code," *Proceedings of the SIGCHI*, May 1988.

[3] Winston W. Royce, "Managing the Development of Large Software Systems," *Proceedings of IEEE WESCON*, pp. 1–9, 1970.

[4] Larman C. and V.R. Basili, "Iterative and Incremental Developments. a Brief History," *Computer*, vol. 36, no. 6, pp. 47 – 56, June 2003.

[5] B. Boehm, "A Spiral Model of Software Development and Enhancement," *SIGSOFT Softw. Eng. Notes*, vol. 11, no. 4, pp. 14–24, 1986.

[6] Mark C. Paulk, "Extreme Programming from a CMM Perspective," *IEEE Software*, vol. 18, pp. 19–26, 2001.

[7] Kent Beck, *Extreme Programming Explained: Embrace Change*, Addison-Wesley, 2000.

[8] Agile Alliance, "Manifesto for Agile Software Development," http://agilemanifesto.org, 2001.

[9] Orit Hazzan and Yael Dubinsky, "Why Software Engineering Programs should teach Agile Software Development," *SIGSOFT Softw. Eng. Notes*, vol. 32, no. 2, pp. 1–3, 2007.

[10] Kent Beck, *Test-Driven Development by Example*, Addison-Wesley, 2003.

[11] H.D. Mills, M. Dyer, and R.C. Linger, "Cleanroom Software Engineering," *Software, IEEE*, vol. 4, no. 5, pp. 19–25, 1987.

[12] E. Soloway, "Learning to Program = Learning to Construct Mechanisms and Explanations," *Communications of the ACM*, vol. 29, no. 9, pp. 850–858, 1986.

[13] Michael de Raadt, Richard Watson, and Mark Toleman, "Teaching and Assessing Programming Strategies Explicitly," *Eleventh Australasian Computing Education Conference (ACE2009)*, vol. 20–23, Jan 2009.

[14] Kurt Vanlehn and H. Chad Lane, "Teaching the Tacit Knowledge of Programming to Novices with Natural Language Tutoring," *Computer Science Education*, vol. 15, pp. 183–201, 2005.

[15] A. Ebrahimi and C. Schweikert, "Empirical Study of Novice Programming with Plans and Objects," *SIGCSE Bull.*, vol. 38, no. 4, pp. 52–54, 2006.

[16] Alireza Ebrahimi, "Novice Programmer Errors: Language Constructs and Plan Composition," *Int. J. Hum.-Comput. Stud.*, vol. 41, no. 4, pp. 457–480, 1994.

[17] Chris Quintana, Jim Eng, Andrew Carra, Hsin-Kai Wu, and Elliot Soloway, "Symphony: a Case Study in Extending Learner-Centered Design through Process Space Analysis," in *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, New York, NY, USA, 1999, pp. 473–480, ACM.

[18] Anthony Robins, Janet Rountree, and Nathan Rountree, "Learning and Teaching Programming: A Review and Discussion," *Computer Science Education*, vol. 13, pp. 137–172, 2003.

[19] Nghi Truong, Peter Bancroft, and Paul Roe, "A Web Based Environment for Learning to Program," in *ACSC '03: Proceedings of the 26th Australasian computer science conference*, Darlinghurst, Australia, Australia, 2003, pp. 255–264, Australian Computer Society, Inc.

[20] Michael McCracken, Vicki Almstrum, Danny Diaz, Mark Guzdial, Dianne Hagan, Yifat Ben-David Kolikant, Cary Laxer, Lynda Thomas, Ian Utting, and Tadeusz Wilusz, "A Multi-National, Multi-Institutional Study of Assessment of Programming Skills of First-Year CS Students," in *ITiCSE-WGR '01: Working group reports from ITiCSE on Innovation and technology in computer science education*, New York, NY, USA, 2001, pp. 125–180, ACM.

[21] J. Kuhl and T. Goschke, *Volition and personality*, Hogrefe and Huber, 1994.

[22] Patricia Yancey Martin and Barry A. Turner, "Grounded Theory and Organizational Research," *The Journal of Applied Behavioral Science*, vol. 22, no. 2, pp. 141, 1986.

[23] Anna Eckerdal, Anna Eckerdal, and Anna Eckerdal, "Novice students learning of object-oriented programming," 2006.

[24] J. O. Sugay M. M. T. Rodrigo, R. S. J. Baker and E. Tabanao, "Monitoring novice programmer affect and behaviors to identify learning bottlenecks," *Philippine Computing Society Congress*, March 2009.

[25] Suzanne Marie Thompson, "Exploratory study of novice programming experiences and errors," March 2006.

[26] Tony Clear, Jenny Edwards, Raymond Lister, Beth Simon, Errol Thompson, and Jacqueline Whalley, "The Teaching of Novice Computer Programmers: Bringing the Scholarly-Research Approach to Australia," in *ACE '08: Proceedings of the tenth conference on Australasian computing education*, Darlinghurst, Australia, 2008, pp. 63–68, Australian Computer Society, Inc.

Hardware Resources in Digital Systems Teaching

Yimin Xie, David Wong and Yinan Kong
Department of Physics and Engineering
Macquarie University
Sydney, NSW 2109, Australia

ABSTRACT

This paper provides an overview of all the hardware resources necessary to support the delivery of a sequence of courses in digital systems covering digital fundamentals to hierarchical design of complex digital systems. An example of an effective approach to engineering education is shown through the use of these resources.

Keywords: Digital Systems, Engineering Education, Hardware Design and Problem-Based Learning.

1. INTRODUCTION

The scope of the materials covered in a sequence of courses in digital systems is very wide, and therefore, adequate hardware resources to support the teaching are imperative. Hardware trainers have been developed to satisfy this requirement in the author's department. Some trainers have been designed with a focus on a specific digital concept while others have a much wider application and provide resources for problem-based learning. Trainers of both types are described in this paper.

The four courses involved in digital systems teaching consist of Digital Fundamentals (in the first year), Programmable Logic Design (in the second year) and Computer Hardware and Digital Systems Design (in the third year). The scope of the topics covered in the four courses is represented by the content of the textbooks and references shown below in Table 1.

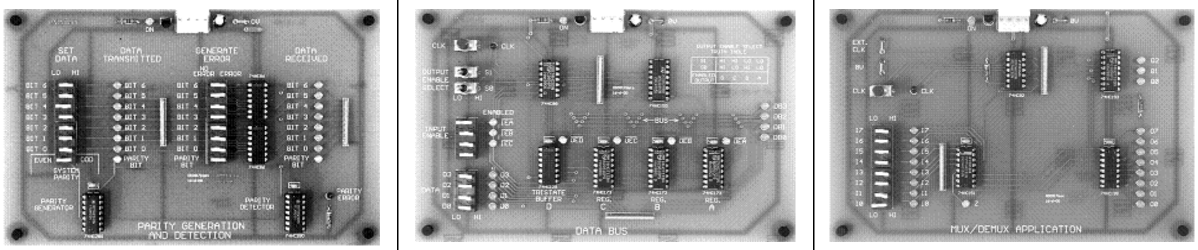
2. CONCEPT-SPECIFIC TRAINERS

This set of hardware is developed for the course of Digital Fundamentals. It also needs to be used for demonstrations in class as well as in large activities like university open days. All these require that the concept-specific trainers have to be easy to set up, and also that the PCB must be fully operational after the connection of the power supply (and clock). It must not require any additional wiring (or, in some cases, a minimum amount of wiring), and all the relevant signals must be labelled and monitored by light-emitting diodes (LEDs).

For this reason, a typical trainer of this type was designed to include small interactive printed-circuit boards (PCBs) that focus on digital concepts of parity, multiplexing, adder/accumulators, flip-flops/counters, data bus and shift registers. Then the advantage is realized that the PCB is exercised by setting switches or pressing pushbuttons (to provide a single clock pulse), and by monitoring signals using the LEDs or a logic analyser. This exactly satisfies the requirement of a simple class demo and also achieves the aim of spurring the interest of new undergraduates in digital systems design. Figure 1 shows three of these PCBs.

Table 1 Author(s) of Text/Reference Books

Course	Author	Text/Reference Book
Digital Fundamentals	T. Floyd	Text [1]
Digital Fundamentals	R. J. Tocci, N. S. Widmer and G. L. Moss	Reference [2]
Programmable Logic Design	T. Floyd	Reference [3]
Computer Hardware	P. Spasov	Text [4]
Computer Hardware	R. J. Dirkman and J. Leonard	Reference [5]
Digital Systems Design	C. H. Roth and L. K. John	Text [6]



(a) Parity | **(b) Data Bus** | **(c) Multiplexing**

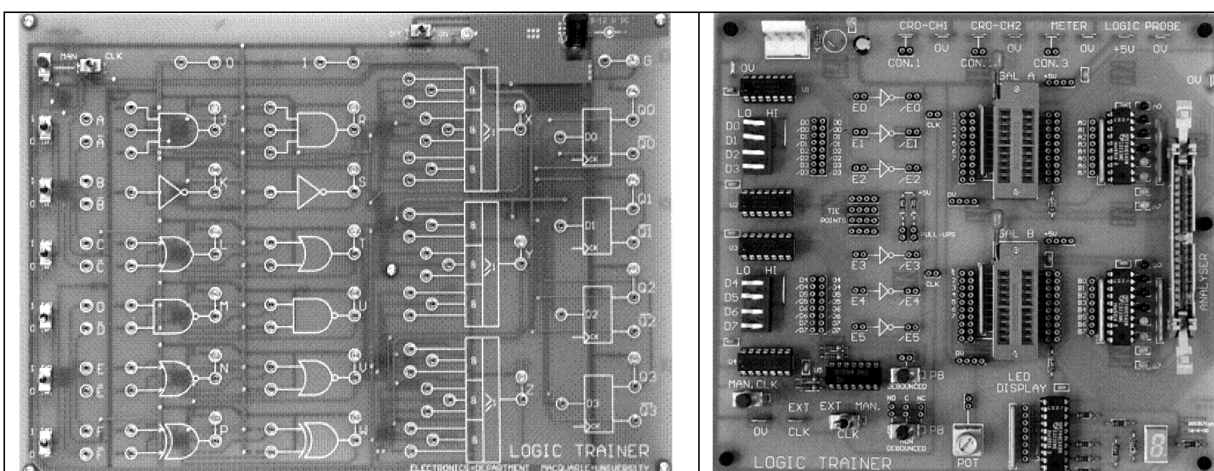
Figure 1: Examples of Concept-Specific Trainers

The present set of trainers has been implemented with medium-scale integrated circuits. Other concepts (such as error-correction codes) have been considered, and will probably have an implementation using complex programmable logic devices (CPLDs).

3. GENERAL-PURPOSE TRAINERS

The general-purpose trainers focus on logic circuits and small/medium-scale integrated circuits (ICs). This type of trainer provides basic building blocks that may be interconnected by patch leads to build combinational or sequential circuits of simple or modest complexity. This is exactly the kind of techniques students need to be equipped with in their second and third year. Two versions of this type have been used in the experiments to enrich the students' hands-on experience in circuit design.

The first consists of a PCB with top-over-layer printing showing distinctive shapes of gates and flip-flops with associated sockets for inputs and outputs that may be physically interconnected. Each output is monitored by an LED, and constructed circuits may be exercised using switches and pushbuttons. The second version is a PCB with several zero-insertion-force (ZIF) sockets for small/medium-scale integrated circuits. This version provides students with the experience of selecting and handling a wide range of ICs. Figure 2 shows the two versions of general-purpose trainers described as the Digital Fundamentals Trainer and the Integrated Circuit Trainer. The integrated-circuit trainer contains only two 20-pin ZIF sockets. Circuits of greater complexity may be handled by the GAL8 Trainer which has eight 24-pin ZIF sockets.



(a) Digital Fundamentals Trainer | **(b) Integrated Circuit Trainer**

Figure 2: General-Purpose Digital Trainers

4. PROGRAMMABLE LOGIC TRAINERS

Programmable Logic Trainers are trainers that exploit the flexibility and immense functional capability of programmable logic devices (PLDs). They have been used in the experiment sessions in the course of Programmable Logic Design and Digital Systems Design.

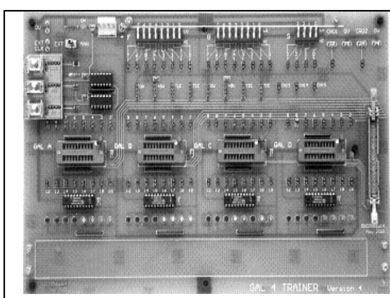
Programmable Logic Trainers have been developed using Generic Array Logic (GAL) devices and Field Programmable Gate Arrays (FPGAs). The GAL devices were chosen because their architecture, consisting of a programmable AND array followed by an OR gate and flip-flop, with feedback connections from the flip-flop to the AND array, provides students with a smooth transition from concepts of finite state machines (FSMs) studied in the earlier course. Modern FPGA devices have a huge functional capability, and this type of PLD is the logical choice for the implementation of complex digital systems. In addition, the mandatory use of computer-aided-design (CAD) software introduces students to state-of-the-art procedures used in industry for the implementation, simulation and synthesis of FPGA- based systems.

Figure 3 illustrates the programmable logic trainers. The GAL4 trainer consists of a PCB with four 20-pin ZIF sockets for four GAL16V8 devices. The GAL8 trainer consists of a PCB with eight 24-pin ZIF sockets for eight GAL22V10 devices, together with a microcontroller and a variable clock generator. An important feature of this type of trainer is that all outputs may be labelled and monitored with LEDs. In the case of the GAL8 trainer, there are 80 LEDs, and 32 signals may be monitored by a 32-channel logic analyser. The FPGA trainer consists of a FPGA (Spartan-3) development board interfaced to a desk-top computer with input and output boards. This system provides the capability to design and synthesise complex digital systems.

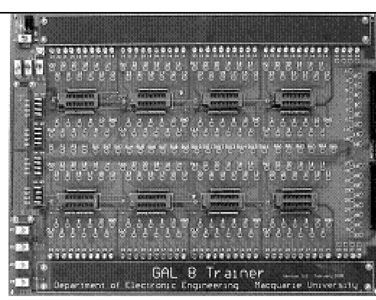
The Xilinx Spartan 3 FPGA has sufficient capacity to design and build complex systems such as those developed for final-year thesis projects (e.g. [7]).

5. MICROCONTROLLER TRAINERS

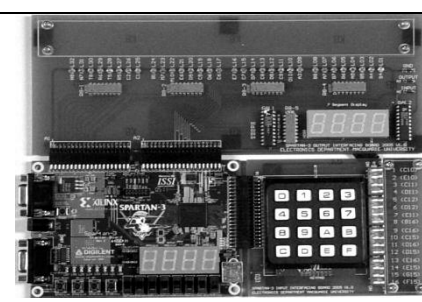
Microcontroller trainers have hardware and software resources that give students the experience of exercising and developing microcontroller-based systems. Microcontroller interfacing is important and experiments should cover parallel ports, serial ports, interrupts, timing and digital/analogue inputs/outputs. As bit-level manipulation is required, especially for a memory-mapped input/output architecture (such as that of the Motorola 68HC11 microcontroller), some experience with assembler coding is required. The microcontroller trainer consists of a microcontroller development board interfaced to a desktop computer and a specially designed input-output PCB. This supports a number of experiments on computer interfacing. the 68HC11 Evaluation Board and the Input/Output Board are shown in Figure 4.



(a) GAL4 Trainer

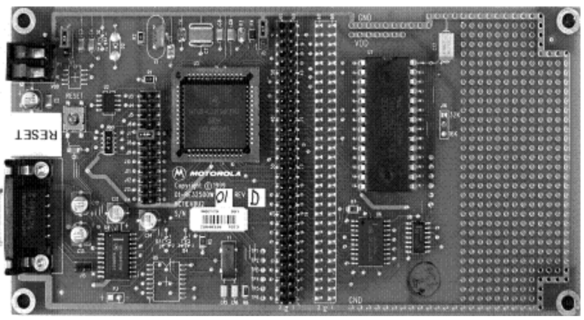


(b) GAL8 Trainer

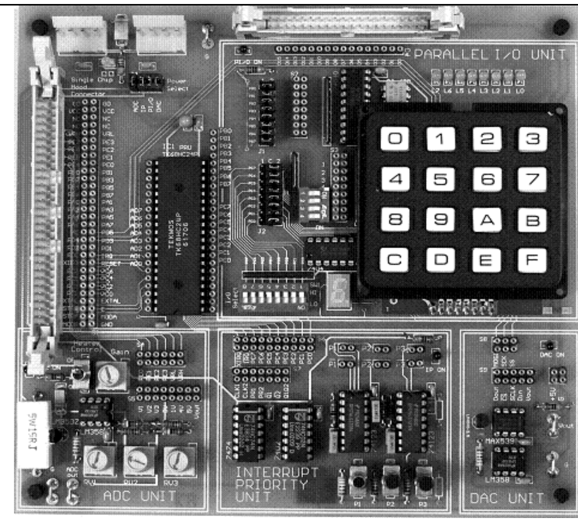


(c) FPGA Trainer with I/O boards

Figure 3: Programmable Logic Trainers



(a) 68HC11 Evaluation Board



(b) Input/Output Board

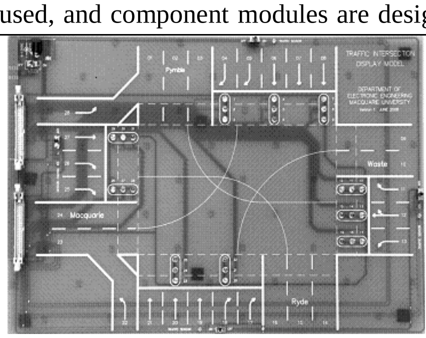
Figure 4: Microcontroller Trainer

6. HARDWARE RESOURCES FOR PROBLEM-BASED LEARNING

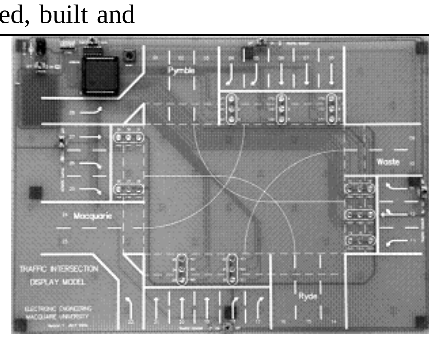
Problem-based learning (PBL) was a feature of a new course on Programmable Logic Design [8]. For this course, the main project involved the design, construction and testing of a digital controller to control the traffic lights of a complex traffic intersection. A PCB with LEDs in positions corresponding to those in the physical intersection was prepared to facilitate testing. This is shown in Figure 5 (a). To support those teams who may opt for VHDL implementation, a new PCB containing a CPLD and a microcontroller, together with a model of the intersection, has been developed. This is shown in Figure 5 (b). A PBL approach has also been used in the course Digital Systems Design, in which the main project is the development of a bus-structured computer. A hierarchical design approach has been used, and component modules are designed, built and tested using the GAL8 Trainer. To facilitate the development of a prototype, a PCB with a static RAM and tri-state buffers has been built. This is shown in Figure 5 (c). This PCB may also be used when a FPGA implementation using VHDL is chosen. The hardware resources for PBL have helped students gain the skills necessary to develop complex digital systems.

7. SOFTWARE RESOURCES

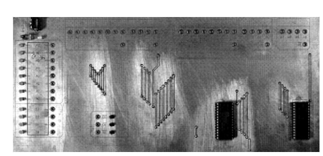
The GAL-based experiments use OPALjr. This allows the use of Boolean algebra for the specification of combinational circuits for GAL outputs and flip-flop inputs. The microcontroller software used is the AS11 cross-assembler. The FPGA experiments use Xilinx ISE Version 9.2.04i software (Xilinx, 2009) for schematic/VHDL entry, simulation and synthesis.



(a) Traffic Intersection



(b) CPLD & Traffic Intersection



(c) RAM & Tri-state Buffers

Figure 5: PBL Hardware Modules

8. CONCLUSION

Hardware resources developed for digital systems teaching have been presented. Concept-specific and general-purpose logic trainers were developed for the introductory courses, programmable-logic trainers were aiming at handling complex digital systems and microcontroller trainers were particularly designed for the course on microcontroller teaching. In addition, dedicated hardware was implemented to assist with problem-based learning.

All of these hardware resources have found their places and are playing an important role in the continuum of digital systems teaching. They help to integrate a systematical stream of digital systems in the education of Electronic Engineering in the author's department.

[1] T. Floyd, *Digital Fundamentals*, *10th ed.*, Pearson, 2009.

[2] R. J. Tocci, N. S. Widmer and G. L. Moss, *Digital Systems – Principles and Applications, 10th ed.,* Pearson, 2007

[3] T. Floyd, *Digital Fundamentals with PLD Programming*, Pearson, 2006.

[4] P. Spasov, *Microcontroller Technology – The 68HC11 and 68HC12, 5th ed.,* Pearson, 2004.

[5] R. J. Dirkman and J. Leonard, *68HC11 Microcontroller Laboratory Workbook,* Prentice Hall, 1996.

[6] C. H. Roth and L. K. John, *Digital Systems Design – Using VHDL*, Thomson, 2008.

[7] A. Pattison-Clarke, *The Design, Implementation and Operation of a Computer Peripheral Controlled by a Field Programmable Gate Array for the Acquisition, Processing and Display of Digital Data Streams.* BE Thesis, Macquarie University.

[8] Wong, Imrie and Xie, "Problem Based Learning Applied to a New Unit of Study on Programmable Logic Design", *Proceedings of Australasian Association for Engineering Education* (AAEE), 2008.

Innovations Required for Retail Beamed Power Transmission Over Short Range

Girish Chowdhary and Narayanan Komerath
Daniel Guggenheim School of Aerospace Engineering
Georgia Institute of Technology, Atlanta, GA

Abstract

Retail beamed power has many potential applications. This paper concentrates on the short range (within 100 km), low efficiency (less than 50%) applications of beamed power transfer. Such applications include connecting micro to medium sized electronic devices with distributed energy sources, transmission of power to mobile consumers, and rapid restructuring of conventional wired grid topology for damage mitigation. The innovations required for realizing these applications are discussed.

Keywords: Retail beamed power transfer, Power distribution and delivery, micro renewable energy

1. Introduction

Modern society has come to rely heavily on devices that are powered using electricity. The traditional method of electric power transmission consists of wired power grids that are widely implemented and well understood. Although modern wired power grids are efficient at high capacity, they require a significant, costly and relatively rigid physical infrastructure. This rigidity is not conducive to mobile computing and to distributed power generation. Applications such as power delivery to remote military and scientific outposts, disaster areas, miniature autonomous robots, and distribution on extra-terrestrial bases, are best served by a flexible method of power transmission. Finally, there is a perceived need for rapidly reconfigurable grids for damage mitigation.

Beamed Power Transmission Systems (BPTS) offer the required flexibility for power distribution. First demonstrated by Nikola Tesla in 1897, this method uses electromagnetic radiation. This paper analyzes the feasibility of BPTS for emerging short range (within 100 km), low efficiency (less than 50%) applications. The paper then points out the technological innovations and research directions required to realize these applications. We begin by describing the science of beamed power transfer. We then discuss the required innovations by outlining a spectrum of possible BPTS applications. The feasibility and cost-effectiveness of these applications is then explored.

2. The Science of Beamed Power Transfer

Wired energy transfer has various disadvantages:
1. Large infrastructure for transmission of power, including wires, poles, land mass (for siting poles), and transformers. Significant efforts and resources are required to set up the infrastructure, and once established it is hard to make changes to the grid topology. Hence, this method of power transfer can be considered rigid.
3. Large clear footprint across forests and mountains.
4. Vulnerable to attacks, accidents and natural disasters. Reliance on a rigid infrastructure inhibits restructuring the grid topology to mitigate damage.
6. High maintenance costs including remote locations.
5. Inhibits development of micro-renewable energy resource exploitation.

Beamed (wireless) power transmission uses electromagnetic radiation (microwave or lasers) for power transfer [2], [3], [11]. BPTS do not rely on a rigid infrastructure of wires and hence can bring great flexibility to power transmission. Wireless power transmission was first demonstrated in 1897 by Nikola Tesla using radio frequencies, and using microwaves in 1964 by William Brown [2]. BPTS were extended to tens of kilowatts by NASA in 1975. In the 1980s, beams of up to 1GW were considered under the Strategic Defense Initiative. BPTS have also been explored for bringing power generated using solar energy in space to the earth both in the USA and outside (e.g. [3], [4], [6], [10]), notably in Japan [13]. BPTS however, has not received much attention for mainstream power distribution.

On the other hand, there has been a revolution in the use of wireless information transmission in the past two decades. Satellite television, cellular telephones, and wireless internet connections reach billions of customers. Research to develop efficient low intensity information transfer over large distances has received a major boost with the advent of digital high-frequency transmission and reception, resulting in devices that require very low power for operation. Thus there are billions of low powered devices operating every day, that have the capability to rapidly decode the information contained in electromagnetic waves.

Wireless transfer of power (or BPTS) uses the same basic science as wireless information transfer. As shown in Figure 1, the direct current is converted to power in the microwave or millimeter wave regime, with efficiencies of 70-90% [2]. The process of forming the beam can be done with efficiencies of 70-97%, but transmission and reception efficiencies vary widely. Ref [2] cites efficiencies as low as 5% to as high as 95%. The final stage of converting the power back from microwave to DC current using rectennas has an efficiency range of about 85-92%. Brown [2] cites an overall efficiency of 52% achieved in DC to DC power transmission using microwave beams in laboratories using standard equipment. Brown also claims that this efficiency could be raised to about 76% using specifically designed components. It should be noted that Brown's results are mostly concerned with power transmission in the 2.4 – 2.5 GHz range. The classical relationship between the efficiency of free space to space transmission and an efficiency parameter τ:

$$\tau = \frac{\sqrt{A_t A_r}}{\lambda D} \qquad 1.$$

where A_t, and A_r represent the area of the transmitting and the receiving aperture respectively, λ denotes the wavelength of the signal, and D denotes the separation distance between the two apertures. Thus as the wavelength is decreased, by going to higher frequencies, the aperture areas required for a given efficiency can be brought down. This consideration offers substantial system improvements if conversion between DC and millimeter waves, especially in the atmospheric transmission windows near 140 and 220 GHz, can be made efficient.

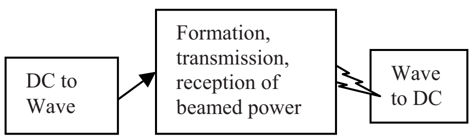


Figure 1: Schematic of a DC to DC beamed power transmission system

3. Feasibility of low range low efficiency BPTS

In the past, researchers have heavily concentrated on the use of BPTS as an enabling technology in Space Solar Power Systems (SPS) (see e.g. [2], [3], [4], [10], [13]). These systems are characterized by low efficiency power transmission from orbital distances (around 400 to 800 Km). With the notable exception of [19], little attention has been given to low intensity, short range (in 100s of Km) power transmission using BPTS for powering devices requiring low to medium range power. This potential area of application of BPTS holds tremendous potential considering the ubiquity of such devices. Power delivery in this context is characterized by low intensity, shorter range, broad coverage, consistency, and efficiency can be traded off for convenience and coverage. We will now use the Friis transmission equation to show that in this context BPTS can produce feasible designs.

Considering ideal conditions for microwave beam transmission, the Friis transmission equation is:

$$P_r = \frac{P_t G_t\, G_r}{(4\Pi R)^2}\lambda^2 \qquad 2.$$

where, P_t is the transmitted Power, P_r is the received power, G_t, G_r are the gains of the transmitting and receiving antenna respectively, and R is the distance over which the power is transferred. Assuming a conservative power efficiency of 20% (i.e. P_r / P_t = 0.2), the variation of the range of the BPTS system with the system frequency is given in Figure 2.

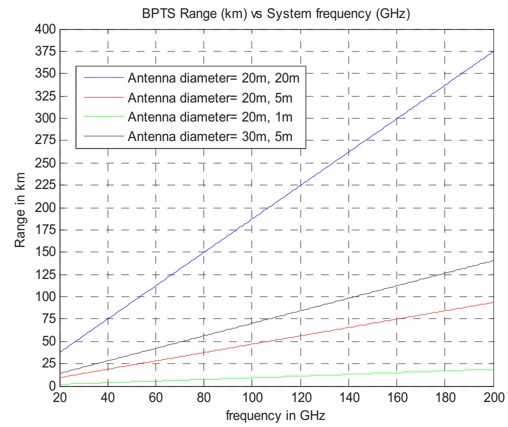


Figure 2 BPTS Range vs. System Frequency with different antenna diameters

The figure shows that for different combinations of antenna diameters the range varies linearly with the transmission frequency.

Figure 3 shows the variation in antenna diameter with different frequency ranges from 2GHz to 200GHz. It can be seen that at lower frequencies (near 2 GHz) larger antennas are required for achieving good transmission ranges. This plot, combined with the fact that the 2 GHz to 2.5 GHz frequency band is heavily used by wireless LAN and other electronic devices, suggests that optimal frequencies for BPTS might lie between 50GHz and 300GHz.

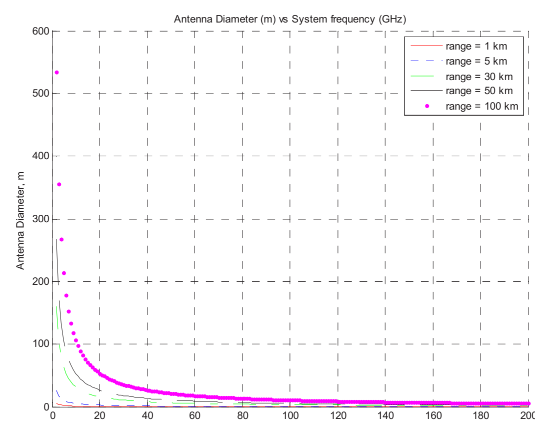


Figure 3 Receiver antenna diameter vs. transmission frequency for different system ranges.

Figure 4 shows the variation in power efficiency as a function of system frequency for a transmitter / receiver antenna combination of 20m and 5m. The plot indicates that higher efficiencies can be achieved at higher transmission frequencies. Hence analysis using the Friis equation strongly suggests use of high transmission frequency for BPTS. However, these results must be studied with caution as they do not account for the effect of atmospheric attenuation. Particularly, it is known that for frequencies greater than 10 GHz the attenuation due to rain is about 10dB/km, while that due to water vapor is greater [1]. Furthermore, in urban environments multi-path effects must also be taken into account. Finally, we note that Ball has raised concerns over the accuracy of atmospheric attenuation values at different frequency ranges [11]. Ball's comments suggest that the values might be excessively conservative, particularly for vertical transmission where the atmospheric density is no longer constant.

In summary, even with a low 20% efficiency the Friis equation indicates that a feasible BPTS system with reasonable antenna dimensions can be designed that can enable wireless power transmission over short ranges.

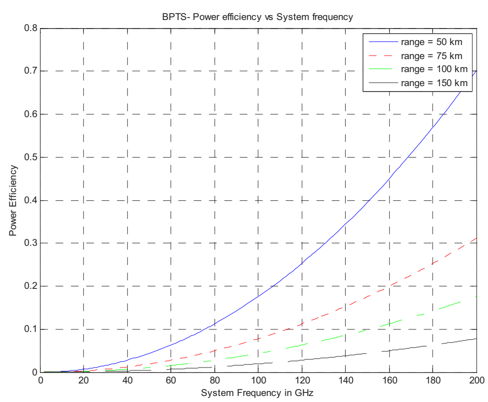


Figure 4 BPTS power efficiency vs. transmission frequency for a different system ranges, assuming transmitting antenna diameter of 20m and receiving antenna diameter of 5m.

This analysis supports the feasibility of BPTS intended to deliver small amounts of power to micro-to-medium sized devices for performing valuable functions over short ranges. In the simplest architecture, power will be spread over a significant area, inside which all devices can accept whatever power falls on their collectors, the rest being wasted. This will result in a significant loss of efficiency, however, if the energy used is sourced from micro-renewable sources, this becomes less of a concern.

Such BPTS can augment the established wired grid to satisfy requirements that are currently ill satisfied:
1. Rapidly provide power to remote consumers without having to setup expensive wired architecture or without having to carry power generation capabilities.
2. Manage seamlessly distributed power generation and consumption. This includes connecting mobile or distributed power sources with the main grid and connecting mobile power consumers with distributed sources without requiring elaborate wire infrastructure.
3. Rapidly restructure grid topology to enable fast and efficient mitigation of partial grid failures.

4. Potential Applications

In this section we discuss some potential applications of BPTS. This discussion serves to give perspective on what is required in terms of technical breakthroughs and pinpoint technical innovations required.

1. **Broad area low intensity power distribution**: Small scaled BPTS can be used in a mall or a café type environment to allow enabled devices to charge automatically. Due to the proximity of the power source, and low power requirements of modern portable electronics, the intensity of the transmitted power need not be high.
2. **Rapid power delivery to remote military or scientific outposts:** BPTS can be used to deliver targeted power to military outposts or scientific outposts operating in remote regions.
3. **High endurance miniature robots**: Battery power is one of the limiting factors in the design of miniature robotic vehicles such as Miniature Unmanned Aerial Systems (M-UAS). These vehicles are unable to carry significant amount of onboard power. Power delivery using trailing wires has been previously attempted [14]; however this method can prove impractical and limiting due to the infrastructure required. BPTS have the potential to revolutionize the capabilities afforded by M-UAS by using beamed power transmission to increasing their endurance. This capability can enable exciting new applications, including indoor arena exploration with a team of networked M-UAS operating in the vicinity of a mother-ship that harvests energy locally and transmits power wirelessly to recipient M-UAS.
4. **Distributed energy generation**: There is a strong drive in the market towards using smaller self sufficient units that generate sufficient electricity for local purposes. This paradigm is termed as distributed energy generation. It is postulated that distributed energy generation will not only be able to support our ever growing energy needs, but it has the potential to be extremely cost effective and sustainable since it is primarily based on exploiting renewable resources [14]. The power distribution systems that support distributed energy generation must remain flexible and highly adaptable. BPTS based distribution systems can provide this flexibility as they are not tied to a rigid ground based infrastructure.
5. **Remote area exploration**: Beamed power can be used to increase the distance a single unit can cover while exploring remote areas through targeted power delivery. On extraterrestrial bases such as the moon, beamed power is an extremely attractive option since the cost of transporting wires and other grid related components can be formidable. Furthermore, where there is no atmosphere, (such as the Moon) ideal efficiencies can be achieved. These benefits make the use of beamed power on lunar bases an extremely attractive option. Reference [16] suggests the use of mobile lunar surface power generation plants to ensure continued solar energy

supply throughout the lunar night and day. BPTS systems have a clear advantage over conventional wired power transmission systems in this case.

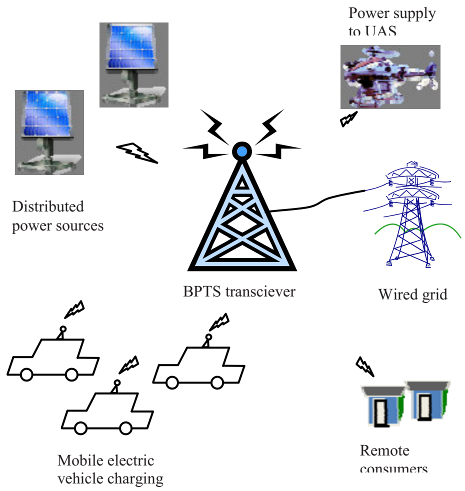


Figure 5: Possible applications of BPTS over short ranges (tens of kilometers)

6. **Increasing the range of electrically powered vehicles**: Current electrically powered vehicles are limited in range by battery weight and volume. Enabling metered power delivery on highways or in parking lots can aid greatly in increasing the range of electric cars and can provide new sources of revenue. By shielding the vehicles, or by delivering the power only when the vehicle is parked and unoccupied, safety concerns can be eliminated.
7. **Rapid restructuring of grid topology for damage mitigation**: Augmenting the wired power grid with BPTS transceivers can mitigate partial damage in conventional wired grids by enabling rapid rerouting of power to avoid areas of the grid that are damaged and incapable of transmission.

Figure 5 shows a schematic of the possible flexibility in applications afforded by BPTS. The figure visualizes the distribution of power generated from distributed energy sources and the conventional wired grid to remote consumers, flying vehicles, and electric vehicles.

5. Cost-effectiveness of BPTS for Potential Applications

The current cost per kilometer for wired power transmission is around USD 1 million (2010) (see for example [20],[21] for cost per mile). This cost includes material, siting (land cost), and environmental costs. It is traditionally computed by dividing the total cost of the project by the number of kilometers of wires used. Since traditional transmission projects range over hundreds (if not thousands) of kilometers, adjustments must be made for savings realized due to large scale production and operation. Due to this reason, this cost could be misleading over short distances. Therefore, we propose the following model for capturing the effective cost ϵ_w of wired power transmission:

$$\epsilon_w = \alpha R + e^{-\beta R}. \qquad 3.$$

In the above equation, R denotes the distance over which power is to be transmitted, the constant α captures the cost per kilometer, and the exponential term accounts for the effect of savings realized over large distances. Clearly, for a given β, if R is large, the effect of the exponential term is negligible.

On the other hand, the infrastructure cost for BPTS over short distances is significantly lower than wired transmission systems as siting and material costs are significantly reduced. Assuming that a BPTS implementation over 1 km is equivalent to a wired implementation with 10 poles, one way to approximate the cost of BPTS per kilometer is to divide the cost per kilometer of wired transmission by 10. With this assumption, the cost of BPTS can be approximated to be about USD 100,000 per km. However, this cost does not account for the reduced efficiency of power transmission over large distances. The power transmission efficiency (P_r/P_t), as modeled by the Friis equation (equation 2), is inversely proportional to the distance. The following model can be used to capture the effective cost ϵ_b of BPTS systems:

$$\epsilon_b = 0.1\alpha R + \frac{\gamma}{\frac{P_r}{P_t}}. \qquad 4.$$

In the above equation, for a scaling constant γ the last term captures the losses resulting from loss of efficiency over long distances. Figure 6 show the effective cost of wired and beamed power transmission for various frequencies. In that plot the receiving and transmitting antenna diameters are set to $D_r = 4\ m, D_t = 5\,m$, the antenna gains are approximated using the equation $G_r = \eta\left(\frac{\pi D_r}{\lambda}\right)^2$, where $\lambda = \frac{c}{\omega}$ is the wavelength and $\eta = 0.8$ is a transmission efficiency factor. From Figure 5 it can be seen that over a distance of 1 Km BPTS are competitive with wired power systems. The figure also shows that the cost of BPTS increases significantly over that of wired transmission with increasing transmission distance, and that the cost is inversely proportional to the transmission frequency. Particularly with a transmission frequency of around 20 GHz, BPTS can compete with wired transmission for distances up to 6 Km. The numerical values will depend heavily on the actual design of the BPTS, nonetheless, the trends should still remain the same.

The above preliminary analysis indicates that for applications requiring power transmission over short distances (within 100 km) BPTS can compete with wired power transmission in terms of cost effectiveness. There are a number of applications that fall into this category, including the ones mentioned in the previous section. On the other hand, when efficiency over long distances is considered, BPTS are ineffective. Transmission at higher frequencies, improved antenna design, and better understanding of health effects of high frequency transmission are required to make a case for BPTS over larger distances.

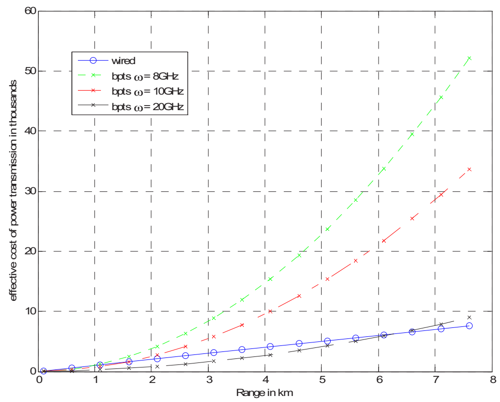


Figure 6 effective cost of wired and beamed power transmission for various frequencies.

6. Required Technological Innovations

In this section we provide a high level overview of the required technological innovations to realize some of the potential applications discussed in the previous section.

a) **Efficient Frequency Conversion:** One major hurdle in implementing BPTS is the inefficiency in converting microwave frequencies to the typical 50-60 Hz operating frequency voltages, and vice-versa. Advancements in Optical Rectennas which couple and rectify optical frequencies to DC provide encouraging news for beamed power. This direction of research needs to be further pursued.

b) **Advancements in Antennas:** Antennas operating efficiently at higher frequencies are very important for beamed power. Development of low cost antennas capable of transmitting at higher frequencies (20 to 200 GHz) with narrow frequency bands will be required. These antennas should be designed in the form of phased arrays, to allow real-time control for changing the direction of beam at high frequencies without physical actuation.

c) **Advancement in Electronics:** Passive and active electronic circuits operating at frequencies of the order of 90 and 250 GHz will require significant advancements in nano-scale fabrication technologies and RF engineering.

d) **Direct Conversion of Broadband Sunlight to Narrowband wave:** The implications of solving this breakthrough technological concept are enormous from a Retail Beamed Power perspective. The need to convert sunlight to DC is eliminated by this technology and hence the efficiency improves significantly not only for Space Solar Power (SPS, an application of BPTS that has been studied widely, see e.g. [3], [4], [10], [13]), but in general for Beamed Power.

e) **Innovations in sub 200GHz radiation monitoring**

f) Brown in [2] mentions that studies conducted by DOE/NASA have found no major issues that would hinder the deployment of BPTS, including environmental and biological considerations (Brown is referring to [18]). Similar studies are yet to be reported on the 200 GHz regime.

g) **Decentralized grid management through networked control:** A major capability brought about by the use of beamed power is the flexibility in the grid and the ability to support distributed power generation. Advances in decentralized management of grid structures are required. This includes thinking of each transceiver as an autonomous agent that must use locally available information to function synergistically with other networked transceivers to meet globally defined needs. The emerging field of decentralized control of networked systems (see for example [17]) promises the development of tools that will be essential for guaranteeing this. Some areas where technological advances are needed are: 1. Real time decentralized fault detection in grids 2. Efficient real time restructuring of grid topology to ensure uninterrupted delivery of power by bypassing nonfunctional units. 3. Decentralized grid voltage regulation.

h) **Graph based models of decentralized grids**: A power distribution system consists of independent power generation and consumption nodes that are connected through some kind of a network. Such systems are very well represented through the framework of Graph theory. Graph theory has excellent tools that can be used to model power distribution systems. For example, the notion of strong connectivity of a graph can be used to determine whether a network can distribute power over all its nodes. Research in bringing the tools of graph theory to modeling decentralized grids equipped with BPTS will be invaluable.

7. Conclusions

In this paper we described a number of innovations required for bringing Beamed Power Transmission (BPTS) from concept to reality. We noted that BPTS is an established concept that can bring immense flexibility to power distribution and generation. We pointed out several possible applications for BPTS that range from power delivery to remote outposts, to power delivery to mobile units. Furthermore, the feasibility of BPTS for these applications and its cost-effectiveness was also analyzed. We conclude that technological innovations needed include improved antenna design, efficient frequency conversion, conversion of broadband sunlight to narrowband wave, innovations in radiation monitoring, and control theoretic approaches in decentralized grid management. This list is by no means exhaustive. We strongly believe that current advances in electronics, control theory, and antenna design are pointing in the right direction to undertake these innovations.

REFERENCES

[1]. Girish Chowdhary, Rajeev Gadre, Narayanan Komerath, **Policy Issues for Retail Beamed Power Transmission**, Proceedings of the Atlanta Conference on Science and Innovation Policy, October 2009.

[2]. William C. Brown, E. Eugene Eves, **Beamed Microwave Power Transmission and its Application to Space**, IEEE transactions on Microwave Theory and Techniques, Vol. 40, No. 6 June 1992.

[3]. N.M.,Komerath, N. Boechler, S.S. Wanis, **Space Power Grid- Evolutionary Approach To Space Solar Power**, Proceedings of the ASCE Earth and Space 2006 conference, League City, Texas, April 06.

[4]. N. Komerath, V. Venkat, J. Fernandez, **Near-Millimeter Wave Issues for a Space Power Grid.** Proceedings of the SPESIF Conference, American Institute of Physics, March 2009.

[5]. Anon. US**- Climate Change Technology Program** http://climatetechnology.gov/library/2003/tech-options/tech-options-1-3-2.pdf

[6]. Goswami et al, **New and emerging developments in solar energy**; Solar Energy, Volume 76, Issues 1-3, January-March 2004, Pages 33-43.

[7]. Web: http://www.thespacereview.com/article/1210/1 , last visited October 16 2008

[8]. Web:http://www.wifiplanet.com/tutorials/article.php /3116531 , last visited October 31 2008

[9]. Web:http://www.cwaunion.org/issues/osh/articles/page.jsp?itemID=27339127. Microwave and Radio frequency radiation by CWA, Last visited October 31 2008.

[10]. Peter Koert, James T. Cha, **Millimeter wave technology for space power beaming**, IEEE Transactions on microwave theory and techniques, vol. 40. No. 6, June 1992.

[11]. John A Ball, **On Atmospheric Attenuation**, Haystack Laboratories, MA, Dec 3 1986.web link (http://www.haystack.mit.edu/obs/haystack/tech_guide/AtmAtt.pdf)

[12]. Constantine Balanis Antenna Theory: Analysis and Design, 3rd Edition, , John Wiley and Sons, USA, 2005.

[13]. Masahiro Mori, Hideshi Kagawa, Yuka Saito **Summary of studies on space solar power systems of Japan Aerospace Exploration Agency (JAXA)** , Acta Astronautica, Volume 59, Issues 1-5, pp 132-138.

[14]. Samuel A. Johnson, Justin M. Vallely, **A Portable Aerial Surveillance Robot**, Proceedings of SPIE--the international society for optical engineering, Vol 6201, 2006.

[15]. V.V. Vaitheeswaran, **Power to the People**, Farrar, Straus and Giroux, New York, 2003. ISBN 0-374-23675-5.

[16]. Paul D. Lowman Jr , **Moonbase Mons Malapert**, Aerospace America, AIAA oct 2008.

[17]. Magnus Egerstedt, Mehran Mesbahi, **Graph Theoretic Methods in Multiagent Networks**, Princeton Univ Pr, 2010.

[18]. Anon. **Final Proc. of the Solar Power Satellite Program Review**, Conf. Report 800491, Lincoln, NE, DOE/NASA Satellite Power System Concept Development and Evaluation Program, April 1980.

[19]. Andre Kurs, Robert Moffatt, and Marin Soljacic, **Simultaneous mid-range power transfer to multiple devices**. Appl. Phys. Lett. Vol.96, 044102, (2010).

[20]. Matthew H. Brown, Richard P. Sedano, **Electricity Transmission- A Primer**. Technical report to National Council on Electric Policy, June 2004

[21]. Anon., **Canada-Northwest-California Transmission Options Study**, Northwest Power Pool Northwest Transmission Assessment Committee NW- California Study Group, May 2006.

Analysis of Distributed Generation Influence on Voltage Sags in Electrical Networks

Sigridt García-Martínez and Elisa Espinosa-Juárez

Facultad de Ingeniería Eléctrica, Universidad Michoacana de San Nicolás de Hidalgo

Morelia, Michoacán, 58060, México

ABSTRACT

This paper presents an analysis of the influence of distributed generation in the occurrence frequency of voltage sags in electrical networks. It shows how the system characterization regarding voltage sags can be affected by the presence of distributed generation. The analysis is performed by using the fault positions method with a program that was implemented in *Matlab*. Several case studies in the IEEE 57-bus and 118-bus test systems taking into account different percentages of distributed generation penetration are presented.

Keywords: Voltage Sags, Distributed Generation (DG), Fault Position Method, Power Quality.

1. INTRODUCTION

In recent years, natural resource conservation and environmental protection issues have become of great importance. The electric power generation is one area of great concern to environmental problems. For this reason, the use of green energy has been promoted in order to reduce gas emissions that cause global warming, also to have a rational use of electricity and thus contribute in some way to improve the quality of the environment. Some countries have incentive systems for green energy generation. Some countries have set targets on percentage of electricity generation with renewable sources with respect to traditional generation sources; in the next ten or twenty years, these targets vary from 10% to over 40% in some countries[1][2].

In accordance with this, in recent years there has been a major boost for development and use of different generation technologies related to small-scale renewable energy generation, which has been called Distributed Generation. Distributed Generation (DG) is defined as the strategic use of modular power generation units, which are connected to the electrical networks. In general the power range is from 3 kW to 20 MW, however, some references indicate that the power range varies from 5kW to 300MW [3][4]. DG units are installed primarily to provide real power to the electric network, so their power factors are 1.0 or close to 1.0 [4].

There are several advantages when small generation units are installed; these can be economic, environmental, technical (voltage support) and even political (competition). Although the costs associated with DG are still high, DG could be a solution for those situations where high reliability in energy supply is needed [5]. However, the DG represents changes not in the network topology, but in the direction of flows and voltage levels [6]. Therefore, it is important to consider that the installation of DG units, in addition to economical benefit, must ensure the reliability, safety and quality of energy supply.

Moreover, it should be taken into account that the greater the DG introduced into a system is, the greater the influence in the power system behavior will be [7]. Generally, DG usually helps to mitigate voltage sags because it increases the voltage level of fault of network buses and also helps to maintain the voltages during the fault [3][5]. Also, the voltage sags features due to faults on electric networks can be affected by the DG presence; several factors such as the DG penetration level can change the system behavior regarding voltage sags. Therefore it is interesting to analyze the impact of DG on voltage sag probability at the system buses.

Voltage sags are one of the main disturbances that affect the power quality; they are also responsible for important economic losses in the industry because they can cause many devices to operate inadequately. Voltage sag in many occasions leads to complete pauses of the industrial process, pauses that increase even more the losses. For example, there are studies that estimate that in the European Union the losses in the industrial sector due to short duration interruptions and voltage sags were of 85 billion euros in 2007 [8].

Voltage sags are reductions in the *rms* voltage value, between 10% and 90% of the nominal voltage, with durations from 0.5 cycles to 1 minute [9], which are caused mainly by short circuits in the electrical system. Many of the equipment used in industries such as programmable logic controllers, computers, among others, are sensitive to voltage sags. Several studies have been realized to understand, quantify and reduce the voltage sag impact in electrical networks. There are diverse methods to predict and to analyze voltage sags; the number of voltage sags expected in a particular place of a system can be estimated through monitoring or by means of stochastic methods [10].

By monitoring, information about voltage sag magnitudes and duration can be obtained for particular system buses. Nevertheless, a very long period of monitoring may be required to obtain precise results in the system characterization [10].

Stochastic processes based on the model of the network and statistical data of fault are widely used to predict voltage sags at a bus of interest. The stochastic prediction is a very useful tool and it can be used in parallel with monitoring [10].

One of the stochastic methods more widely used for voltage sags estimation is the Fault Positions Method [11]-[13]. Many authors use this methodology to analyze voltage sags [12]-[15]; with this method the expected number of voltage sags can be obtained in any bus of the electrical system under study.

The impact on voltage sags in an electric network taking into account different DG penetration levels is analyzed in this paper by means of the Fault Position Method. In section I, a brief introduction about voltage sags and some DG characteristics are presented. In section II, the Fault Position Method and its implementation are described. In section III, the case studies are presented and an analysis of the DG impact on electric networks is performed. Finally in section IV, the conclusions of the analysis are presented.

2. FAULT POSITIONS METHOD IMPLEMENTATION

The Fault Positions Method allows making a stochastic prediction about the expected number of voltage sags in any bus of the electrical system under study [13]. The method consists of selecting fictitious fault positions in buses and lines of the network, and then it is possible to obtain voltage sags characteristics. To every preselected fault position a certain value of fault probability is assigned, taking into account statistical data about the faults in the electrical system. Through conventional techniques for short circuit calculation, voltage sag magnitudes and the frequency of occurrence are obtained by combining those results with historical data [10][16]. It is important to notice that when more fault positions are used in this method, the results improve but the time of calculation increases.

In this paper, the Fault Positions Method was implemented in MatLab®, modifying the systems admittance matrix in its original state, inserting fictitious buses that correspond to the preselected positions; then an additional procedure for faults calculation was applied. Pre-fault voltages were considered carrying out power flow studies by means of the software PSS/E [17].

3. CASE STUDIES

In order to analyze the influence of DG on voltage sags in electrical networks, studies using the Fault Positions Method have been performed on IEEE test systems of 57 buses and 118 buses. The number of fault positions for the case studies were selected taking into account the accuracy of numerical calculations and computational requirements aspect [19].

A. Studies in the IEEE 57-bus test system

The IEEE 57-buses test system consists of 57 buses which are interconnected by means of 63 lines, 15 transformers and 7 generating units [18]. In this case 50 fault positions in each line were considered. It was also considered that voltage sags are caused by three-phase balanced faults. A fault rate of 0.50 faults/year has been assumed for all system lines [11][14][19], and faults/year rate at buses has been assumed negligible.

For all the case studies, the magnitudes and angles of pre-fault voltages were obtained carrying out a power flow study by means of software PSS/E, as it was mentioned previously. For inclusion of DG in the system, synchronous generators of small power were considered. These generators were connected to load buses of the electrical system.

The next case studies were analyzed:

- A.1. Base case, the original system conditions according to [18] are considered.
- A.2 An increase of 10% of power with DG was considered.
- A.3 An increase of 20% of power with DG was considered.
- A.4 The 50% of the original generation was replaced with DG-
- A.5 The 21% of the original generation was replaced with DG.

In case A.1, the system is analyzed in its original conditions taking into account pre-fault voltages that were obtained using a power flow study.

In order to analyze the DG influence on the expected number of voltage sags at system buses, in cases A.2 and A.3 the DG is included in different buses from the system and different penetration levels have been assumed (10% and 20% with respect to the total real power of the system). The voltage sags number is calculated by means of the Fault Positions Method. It is important to mention that small generation units (DG) were placed at load buses selected in a random way.

Fig. 1 shows the influence of the DG penetration level on the frequency of occurrence of voltage sags. Generation units were connected in buses 23, 28, 31, 35, 43, 44 and 52. The Fig. 1 shows the voltage sags at each bus of the system, when a residual voltage threshold of 0.7 p.u. is considered for the cases A.1, A.2 and A.3, respectively.

The results indicate that increase of DG in an electrical network leads to a reduction of voltage sags in almost all of the buses.

Fig. 2 shows the results for an analogous situation to Fig. 1, but the generation units have been changed to 5, 17, 33, 38, 53, 54 and 56 buses. It can be observed that voltage sags have decreased again; also, it can be observed that there is a higher variation in those buses where a generation unit was connected or that are very near to one.

Fig. 3 and Fig. 4 present results in the same conditions as Fig. 1 and Fig. 2, respectively, but taking into account a residual voltage threshold of 0.8 p.u. In both figures, Fig. **3** and Fig. 4, voltage sags presented at each bus lightly decreased with respect to the base case, A.1.

Fig. 5shows graphics of voltage sags obtained by considering a voltage threshold of 0.9 p.u. It is observed that DG does not contribute to decrease voltage sags in most of the buses. Only buses 1, 2 and 17 present a slight decrease of voltage sags/year; for example, bus 2 has 31.42 voltage sags/year in base case, whereas in the case A.3 where the power was increased 20% with DG, voltage sags for the same bus are 30.16 voltage sags/year.

Table 1 and Table 2 show voltage sags at buses with higher variation considering a residual voltage threshold of 0.7 and 0.8 p.u., respectively. It can be observed that for a threshold of 0.7 p.u. the voltage sag differences among the three case studies are not greater than three voltage sags/year, while for the threshold of 0.9 p.u. variations do not exceed two voltage sags/year.

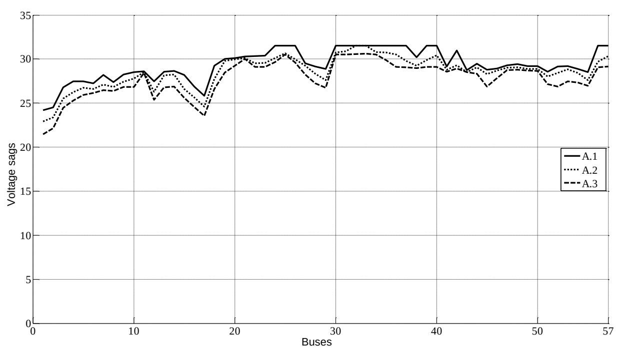


Fig. 1: Voltage sags considering a 0.7 p.u. voltage threshold for cases A.1, A.2 and A3. DG at buses 23, 28, 31, 35, 43, 44 and 52.

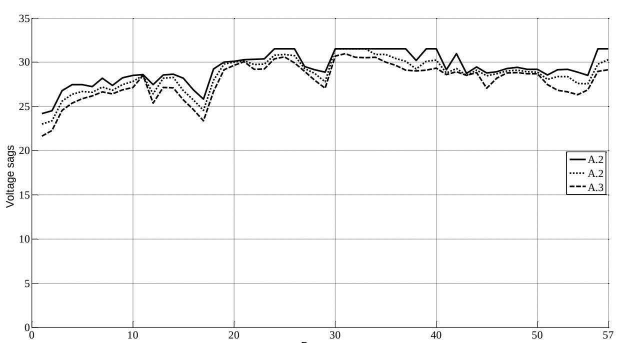


Fig. 2: Voltage sags considering a 0.7 p.u. voltage threshold for cases A.1, A.2 and A3. DG at buses 5, 17, 33, 38, 53, 54 y 56.

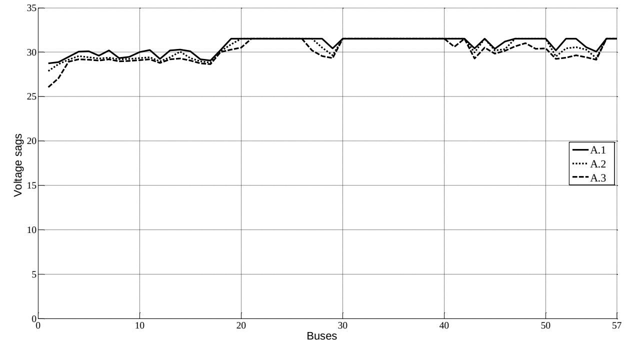


Fig. 3: Voltage sags considering a 0.8 p.u. voltage threshold for cases A.1, A.2 and A3. DG at buses 23, 28, 31, 35, 43, 44 and 52.

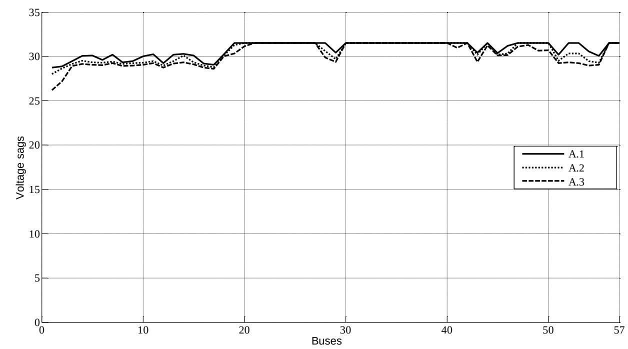


Fig. 4: Voltage sags considering a 0.8 p.u. voltage threshold for cases A.1, A.2 and A3. DG at buses 5, 17, 33, 38, 53, 54 y 56.

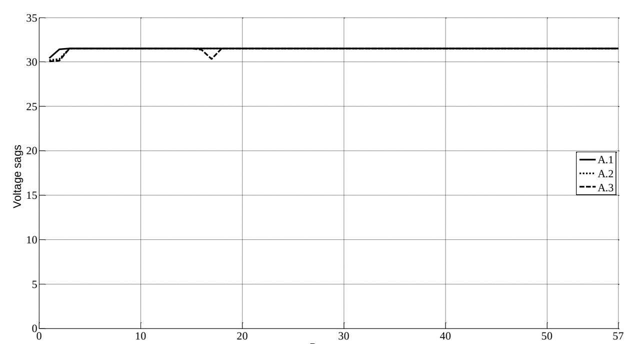


Fig. 5: Voltage sags considering a 0.9 p.u. voltage threshold for cases A.1, A.2 and A3. DG at buses 5, 17, 33, 38, 53, 54 y 56.

TABLE 1:
BUSES WITH HIGHER VARIATION ON VOLTAGE SAGS DUE TO DG, CONSIDERING A 0.7 P.U. VOLTAGE THRESHOLD

Buses	Voltage sags/year		
	A.1	A.2	A.3
1	24.20	22.91	21.46
7	28.20	27.09	26.48
15	28.20	26.62	25.59
29	28.88	27.61	26.74
37	31.50	29.77	29.08
45	28.81	28.3	26.89
54	28.9	27.61	26.33

TABLE 2:
BUSES WITH HIGHER VARIATION ON VOLTAGE SAGS DUE TO DG, CONSIDERING A 0.8 P.U. VOLTAGE THRESHOLD

Buses	Voltage sags/year		
	A.1	A.2	A.3
1	28.75	27.06	26.06
5	30.09	29.42	29.17
14	30.27	30.04	29.31
28	31.5	30.53	29.55
29	30.44	29.63	29.34
46	31.21	30.35	30.17
54	30.55	29.46	28.95

In A.4 and A.5, two of the original generation units, connected to buses 12 and 1 respectively, have been replaced by small generation units dispersed in the electrical network that represent a DG penetration of about 50% for case A.4 and 21% for case A.5.
Fig. 6 and Fig. 8 present results of voltage sags for case A.4, in which the generation unit at bus 12 was replaced by DG using small units at 20 buses of the system. It can be clearly observed that voltage sags at buses decrease for all the considered voltage thresholds, however, the variation is significantly lower for the voltage threshold of 0.9 p.u.

In an analogous way, Fig. 9 and Fig. 10 present results of voltage sags for case A.5, in which 21% of original generation is replaced by DG.
From Fig. 6 and Fig. 9, it is observed that for a voltage threshold of 0.7 p.u. when the penetration level is lower (about 21 % in case A.5), voltage sags/year decrease to a lesser extent at all of the buses with respect to the base case.
A similar situation is observed in Fig. 7 and Fig. 10, corresponding to a voltage threshold of 0.8 p.u. An analogous behavior occurs for the 0.9 p.u voltage threshold, in case A.5 the majority of the buses show no changes.

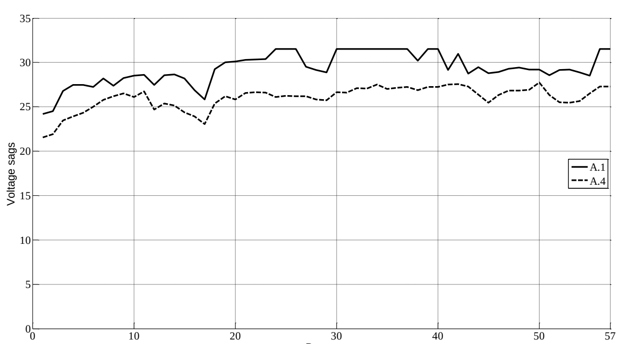


Fig. 6: Voltage sags considering a 0.7 p.u. voltage threshold for cases A.1 and A.4.

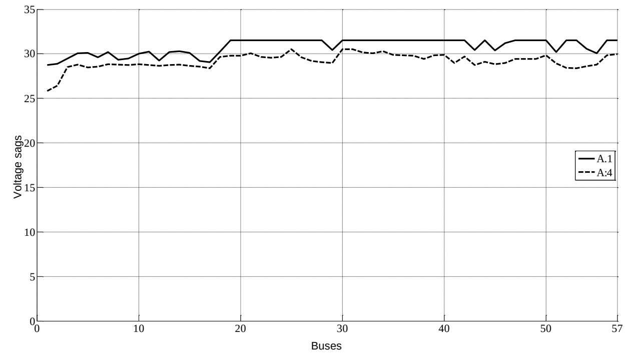


Fig. 7: Voltage sags considering a 0.8 p.u. voltage threshold for cases A.1 and A.4

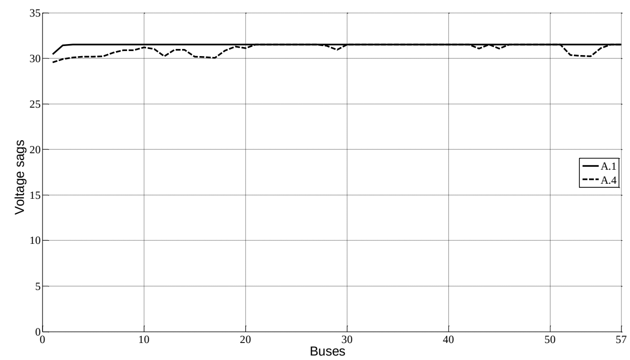


Fig. 8: Voltage sags considering a 0.9 p.u. voltage threshold for cases A.1 and A.4.

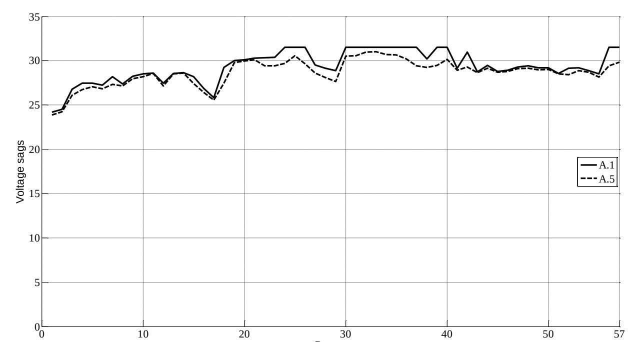


Fig. 9: Voltage sags considering a 0.7 p.u. voltage threshold for cases: A.1 and A.5.

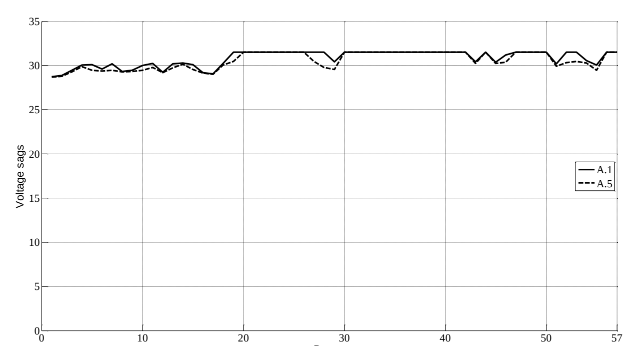


Fig. 10: Voltage sags considering a 0.8 p.u. voltage threshold for cases: A.1 and A.5.

In Table 3 the average of voltage sags/year estimation for all buses in the system considering voltage thresholds of 0.7, 0.8 and 0.9 p.u. are shown for the different analyzed cases. As it can be seen from cases A.2 and A.3, the average of voltage sags per year decreases when there is DG in the system under study. Furthermore, when DG is introduced without changes in the total power generation (case A.4 and A.5), the number of voltage sags decreases.

TABLE 3:
VOLTAGE SAGS AVERAGE FOR THE IEEE 57-BUS TEST SYSTEM

Range of Voltage sags (p.u.)	Average of voltage sags/year				
	A.1	A.2	A.3	A.4	A.5
0.0<v≤0.7	29.43	28.63	27.91	25.90	28.64
0.0<v≤0.8	30.80	30.52	30.22	28.84	30.56
0.0<v≤0.9	31.48	31.46	31.43	30.80	31.46

B. Studies in the IEEE 118-bus test system

The IEEE 118-bus test system consists of 118 nodes interconnected by 177 lines and 54 thermal units. The data system is provided in [20]. In this case 15 fault positions in each line were considered. It was also considered that faults causing voltage sags are single-phase faults. A fault rate of 2.58 faults/year has been assumed for all system lines; this value has been obtained by considering a random number between 0 and 5 of faults at each line; the total number of faults is then divided by the total number of lines in order to obtain the fault rate value [19]. The faults/year rate at buses has been assumed negligible. For the inclusion of DG in the system, synchronous generators of small power were considered to load buses of the electrical system.

The next case studies were analyzed:

B.1. Base case, the system is in original conditions according to [18].

B.2 An increase of 20% of power with DG was considered.

B.3 The 23% of the original generation was replaced with DG

In case B.1, the system is analyzed in its original conditions. In case B.2, the influence of DG in the occurrence and propagation of voltage sags is taken into account, when such generation is included in different buses of the system and when a penetration level has been assumed (20% with respect to the total of real power of the system). In case B.3 about 23% of the total generation was replaced by DG.

Fig. 11 and Fig. 13 show the influence of DG penetration level in the frequency of occurrence of voltage sags when a residual voltage threshold of 0.7 p.u., 0.8 p.u. and 0.9 p.u. are considered for cases B.1, B.2 and B.3. The results indicate that an increase of DG in an electrical network leads to a reduction of voltage sags in almost all of the buses.

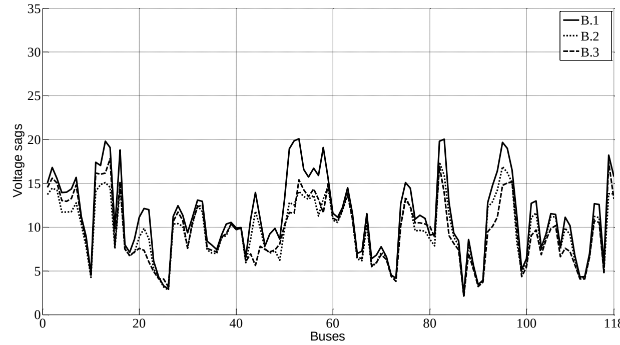


Fig. 11: Voltage sags considering a 0.7 p.u. voltage threshold for cases B.1, B2 and B.3.

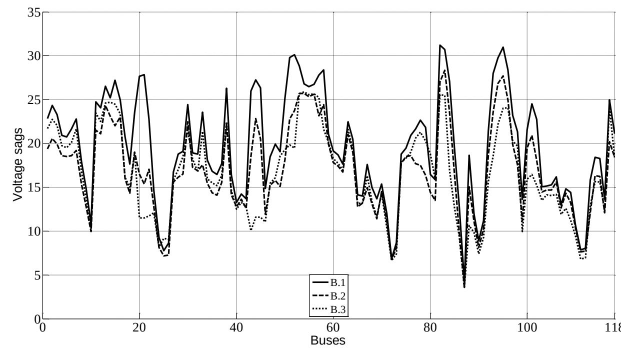


Fig. 12: Voltage sags considering a 0.8 p.u. voltage threshold for cases B.1, B2 and B.3.

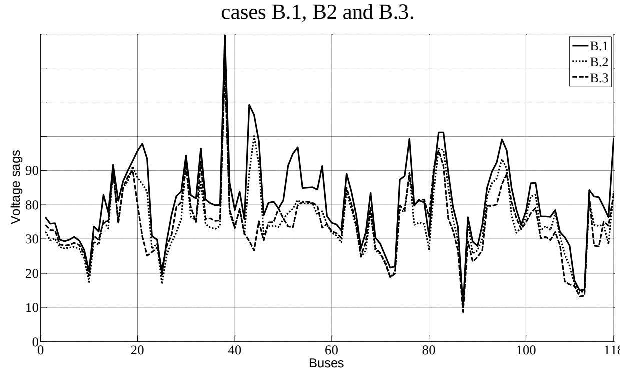


Fig. 13: Voltage sags considering a 0.9 p.u. voltage threshold for cases B.1, B2 and B.3.

The voltage sags changes (in percentage) corresponding to 0.7, 0.8 and 0.9 thresholds, from B.2 and B.3 cases with respect to the base case, are shown in Fig. 14-Fig. **16**. Notice that the greatest difference occurs when 23 % of the original power is replaced by DG. For example, from Fig. 16, at bus 44 voltage sags decrease about 60%. However, the results show that some buses experienced more sags due to the DG location. For example, from Fig. 16, at bus 80 voltage sags increase about 80%, this is because the generator that was connected at this bus was replaced by small units that were disperse on the network. As a consequence, bus 80 does not have the same voltage support.

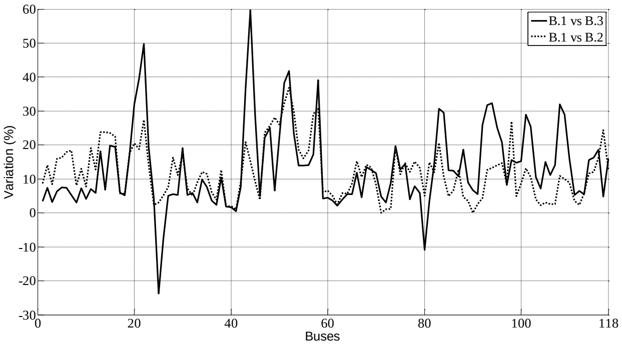


Fig. 14: Voltage sag variation considering a 0.7 p.u. voltage threshold for B.1 vs. B2 and B.1 vs. B.3.

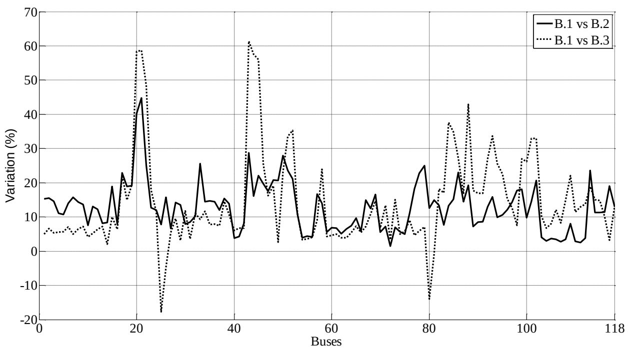


Fig. 15: Voltage sag variation considering a 0.8 p.u. voltage threshold for B.1 vs. B2 and B.1 vs. B.3.

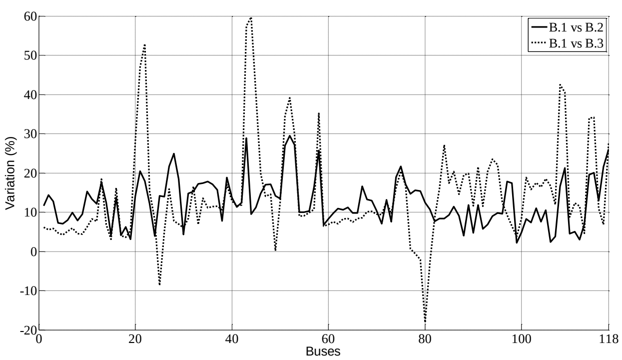


Fig. 16: Voltage sag variation considering a 0.9 p.u. voltage threshold for B.1 vs. B2 and B.1 vs. B.3.

4. CONCLUSIONS

The increasing penetration level of DG makes necessary to study the DG impact on electrical systems performance. In this paper an analysis of the influence of DG on voltage sags in electric networks has been presented. The Fault Positions Method has been implemented and applied to a 57-bus and a 118-bus test system. By these examples, it is shown that generally, the DG contributes to mitigate voltage sags. From all analyzed cases, the results presented have an important variation on voltage sags number for 0.7, 0.8 and 0.9 p.u. thresholds. Moreover, from the case studies analyzed, having DG with total power equivalent to the base case voltage sags decreased. However, some buses in the system have an increase of voltage sags because they do not have the same voltage support when a near generation unit is replaced with DG.

5. REFERENCES

[1] Europe's Energy Portal, Renewables. Brussels, Belgium (Apr. 2009). [Online]. Available in: http://www.energy.eu/#renewable.

[2] National Renewable Energy Laboratory, U.S. Department of Energy (2008, Jul.). 20%Wind Energy by 20030. Increasing Wind Energy's Contribution to U.S. Electricity Supply. [Online]. Available in: http://www1.eere.energy.gov/windandhydro/pdfs/41869.pdf.

[3] J. V. Milanovic, H. Ali, M. T. Aung, "Influence of distributed wind generation and load composition on voltage sags", Presented at International Conference On Power Systems Transients (IPST'07), Lyon, France, 2007.

[4] A. S. Yilmaz, E. Yanikoglu, "Behaviour of embedded generation during the voltage sags in distribution networks", Scientific Research and Essay, vol. 4, pp 112 – 119, March 2009

[5] J. A. Martínez-Velasco, J. Martin-Arnedo, "EMTP model for analysis for distributed generation impact on voltage sags", IET Gener. Transm. Distrib., vol. 1, no. 1, pp. 112-119, January 2007.

[6] J. A. Martínez-Velasco, J. Martin-Arnedo, "Distributed generation impact on voltage sags in distribution networks", 9th International Conference, Electrical Power Quality and Utilisation, Barcelona, 9-11, October 2007.

[7] J. G. Slootweg, W. L. Kling, "Impacts of distributed generation on power system transient stability", Proc. IEEE Power Engineering Society Summer Meeting, 2002, vol. 2, Jul. 2002, pp. 862-867.

[8] J. Manson, R. Targosz, "European Power Quality Survey Report", Leonardo Energy Initiative, Nov. 2008. Disponible en http://www.leonardo-energy.org/european-power-quality-survey-report

[9] *IEEE Recommended Practice for Monitoring Electric Power Quality*, IEEE Std. 1195-1995, Nov. 2, 1995.

[10] M. H. J. Bollen, "Understanding power quality problems. Voltage sags and interruptions", IEEE PRESS Series on Power Engineering, 2000.

[11] L. Conrad, K. Little, C. Grigg, "Predicting and preventing problems associated with remote fault-clearing voltage dips", *IEEE Trans. on Industry Applications*, vol. 27, no. 1, pp.167–172, Jan./Feb. 1991.

[12] M. R. Qader, M. H. J. Bollen, R. N. Allan, "Stochastic Prediction of Voltage Sags in a Large Transmission System", *IEEE Trans. Ind. Applic.*, vol. 35, no. 1, pp.152-162, Jan./Feb. 1999.

[13] G. Olguin, M. H. J. Bollen, "The method of fault positions for stochastic prediction of voltage sags: a case study", *Proceedings of 7th International conference on Probabilistic Methods Applied to Power Systems*, PMAPS *2002*, vol. 2, pp. 557-562

[14] M. F. Alves, R. C. Fonseca, "Voltage sag stochastic estimate", *Proc. Industry Applications Conference 2001*, vol. 3, Sep.30-Oct. 4, 2001, pp. 1665-1669.

[15] M. T. Aung, J. V. Milanovic, "Stochastic prediction of voltage sags by considering the probability of the failure of the protection system", *IEEE Trans. on Power Delivery*, vol. 21, no. 1, pp. 322-329 Jan. 2006.

[16] E. Espinosa-Juarez, A. Hernández, "Voltage sag state estimation: an approach based on the concept of fault positions", Proceedings of the International Conference on Harmonics and Quality of Power(ICHQP-2006), Cascais, Portugal, 1-5 Oct. 2006.

[17] Power System Simulator for Engineering (PSS/E), *Siemens Power Transmission & Distribution*, Inc., Power Technologies International, USA.

[18] R. Christie, "IEEE 57 Bust Test Case", College of Engineering, Electrical Engineering, University of Washington, Aug. 1993. Available in: http://www.ee.washington.edu/research/pstca

[19] E. Espinosa, A. Hernández, "An analytical approach for stochastic assessment of balanced and unbalanced voltage sags in large systems", IEEE Trans. on Power Delivery, vol. 21, no. 3, pp. 1493-1500, Jul. 2006.

[20] R. Christie, "IEEE 118 Bust Test Case", College of Engineering, Electrical Engineering, University of Washington, Aug. 1993. Available in: http://www.ee.washington.edu/research/pstca

Using DSS in an Industrial Context

Authors: Juan Pablo Musella, Gustavo Janezic, Diego Branca, Daniela López De Luise, James Stuart Milne, Germán Ricchinni, Francisco Milano, Santiago Bosio
Universidad de Palermo, Engineering School, Mario Bravo 1050, C1188AAB, Argentine,
aigroup@palermo.edu

Abstract

The purpose of this paper is to present the administration of socio-political aspects in the context of a help decision-making prototype, commonly known as DSS (Decision Support System).It was set to be consulted by industrial enterprises. It consists of a base of the company's own parameters that directly affect the business, an input of parameters associated industrial sectors and macro economic factors that directly affect the activity. In its architecture, there are two expert systems (one feeds the other) and a predefined set of data mining tools. The aim of the prototype presented here is to provide those responsible for decision-making information to minimize the risks faced in a given business. The main contribution of this study is the inclusion of socio-political variables as part of the information entered in the Knowledge Base and how these are derived from existing data in the community.

Keywords: Industry, Decisions, DSS, Expert System, Data Mining.

1.Introduction

DSS are specific information systems oriented to organizational decision making, and business. Their main purpose is to answer questions about things being difficult to evaluate, speculations about future events (typically where the context of the company is complex and mutable). They are commonly used for decision making of low and middle management. They usually provide information transactional and across departments. Data is typically collected from operational or or transactional data bases. One DSS definition describe it as computer-based interactive information systems whose models describe and predict production processes. It is also defined as systems designed to help decision makers to use data and models to identify and solve problems and make decisions. [1]

DSS has four major characteristics:
-integrate data and models
-the design is intended to assist managers in decision-making processes
-they do not replace (but support) human decision-making
-their purpose is to improve the effectiveness of decisions and the decision efficiency.

In the late seventies, DSS through computers was encouraged. According to Power [2] [3] they can process several input types and provide alternate decision. Therefore they can be classified in four categories as follows:

- Communication driven DSS: Allow concurrent support for many users at a specific shared task. Ex.: chats, instant messengers, collaborative systems, etc.
- Data driven DSS: focused on temporal series access and manipulation for certain organization. They are usually used for specific queries in databases or time mutable data Warehousing[1]. In this category lies Geographic information systems, that can be used for smart managing of geographical data from maps.
- Document driven DSS: arrange, retrieve and manage unstructured information into a variety of electronic formats. They are intended to a big number of users. Their main purpose is to crawler web pages to find documents. As example there are several text analysis software, named as Wikis (Web sites freely editable by volunteers).
- Knowledge driven DSS: provide accumulated experience from facts, rules, procedures, from the actual enterprise or imported similar structures. They intend to be applied for problem resolution. They are mainly used to provide advice for management about certain products or services. As example the processing of vast data volumes, hidden pattern recognition and smart representation of such discovered patterns.
- Model driven DSS: centered in statistical, financial model access and administration. It could be used for optimization or simulation. As case is the business process prediction using past event information in order to answer what if questions.

Some of the above described models exist in the market, as proprietary or open-source solutions as described in Table 3.

[1] Database with specific content, usually oriented to an specific field (mainly for certain enterprise, organization, or topic)

Market model				
Product	Remark	License type	Model type	Reference
Dicodess	For distributed DSS	Open-Source	Model driven	http://dicodess.sourceforge.net
Facilitator	Plain DSS	Open-Source	Communication driven	http://facilitator.sourceforge.net
EGADSS	Proof based DSS	Open-Source	Model driven	http://egadss.sourceforge.net
MicroStrategy	Business intelligence with DSS	Commercial	Optionally Data or model driven	http://www.microstrategy.com
Microsoft Office Groove	Collaborative software	Commercial	Communication driven	http://office.microsoft.com
Sybase	Business intelligence with DSS	Commercial	Optionally Data or model driven	http://www.sybase.com
SAP	Business intelligence with DSS	Commercial	Optionally Data or model driven	http://www.sap.com

Tabla 4. Typical DSS systems

As can be seen, both commercial and open source systems have several application fields. Any case, there is a strong tendency to business for commercial DSS.

This paper depict main architecture of the HECULES DSS prototype of the AIGroup research Lab. It handles business specific operative and account transactional information. It also has an statistical database form the business sector and one more database for macroeconomic of the country it belongs. From the logical perspective, all the data stored represents a large variable set classified as: business sector (sector for short), macro and microeconomic.

Selected macroeconomic variables arise from macroeconomic models rules. Among them GDP (Gross Domestic Product), CPI (Consumer Price Index), and occupation index. Even tough several variables are constantly used by the system, many other will depend on the specific query to be answered by the DSS.

Sector variables are mainly focused in the business activity, covering topics ranging from holding issues to competitors characteristics. Wherefore actual variable set changes with the specific analysis the DSS has to perform.

Among variables for microeconomic study, are balance information, since it provides measure of required topics such as enterprise Economic Activity normalized by country or region.

In HERCULES, decision maker perspectives are included as part of the solution search through the proces of variable set definition. A good selection enables more or less solution alternatives even when some variables represent conflicting information. Also expertise endow better definitions and as a consequence tuned answers.

From industrial point of view it could be stated that HERCULES constitutes a useful tool for a wide range of questions. In order to delimit implementation complexity, the type of possible questions was analyzed, defined and restricted. They are classified into two main types: open and restricted queries.

As questions are reusable for many studies, there are meta variables to act as binded wild-cards (see fig. 1). Questions are stated as a set of meta variables connected each other with certain specific words that force a special relationship between them. They are noted in squared brackets. In the figure, the question has two meta-variables: parameter1 and parameter2. These two meta-variables will eventually be replaced with specific variables (pre-defined by the system operator, of course). For instance parameter1 could be binded with total-screw-consumption variable, and parameter2 with screw-total-production-time.

What happens with *[parámetro1]* upon *[parámetro2]*

Fig. 1. Sample of query with two meta-variables

In the following, this papers organizes as follows: architecture (section II), interaction and query management approach (section III), economic model and specifications (section IV), socio-economic topics administration (section V), actual state of the implementation /section VI), conclusion and future work (section VI). [4] [5 [6]

2. HERCULES architecture

The prototype has two user interfaces (see fig.1) . The first for administrative duties over the problem domain. It enables data universe configuration and import activities through a special module named Parser. The Parser makes internal files to be used by an Expert System named Internal Expert System (IES). These files are processed automatically with alternate Data Mining techniques and the results formatted by the user GUI to be properly presented.

Parser follows three steps: select files and variables, explicit relationships between selected files (using foreign keys and primary keys, also with record filtering based on a set of variable values), export data to a plain text file able to be processed by any traditional Data Mining tool.

The second user interface is aimed to provide business consultant interaction and it is directly connected to a second Expert System named User Expert System (UES) fed by the IES upon user commands to show DSS results. UES is complemented with other minor modules to conform the Advisory System. The user interface provides a formalized way to input variable declarations and to ask for specific output formatting.

Figure 1 shows the connection between configuration interface and data conversion and processing. It builds one or more CSV files with processed data. Such files will be data mined for a set of tools as indicated by the IES. To do that, IES asses the best previously known DM approaches for subsets of data it recommends. This way the decision of which data is used and how it is processed remains on IES and its knowledge about the kind of query and the problem.

Results are recorded into Historical DB for further usage by the Advisory System and for self tuning of internal metrics.

Regarding man-machine interaction, this prototype provides a sophisticated data capture and conversion toolkit embedded in the user interfaces. It is important to note that one of the critic functionalities in HERCULES is the ability to upload information from a variety of sources (document files, plain files, worksheets, databases, etc.) into its internal database, and the possibility to process it successfully with a varying subset of tools. The user can define main variable subsets for the present business problem and afterwards import them considering standard units (set ISO/IEC 80000 is included) or custom ones.

The main menu (see Fig. 2), makes it possible to determine the type of study to be performed, define elementary or composed variables, configure Database connection and settle DM restrictions. It permits also to write a textual description of the test for future reference. For elementary variable creation (see Fig. 3 and 4) there are a set of fields to fulfill regarding domain, units, banned and allowed values. It is also possible to create variable groupings by selecting one or more elementary variables (see Fig. 5). [7] [8] [9]

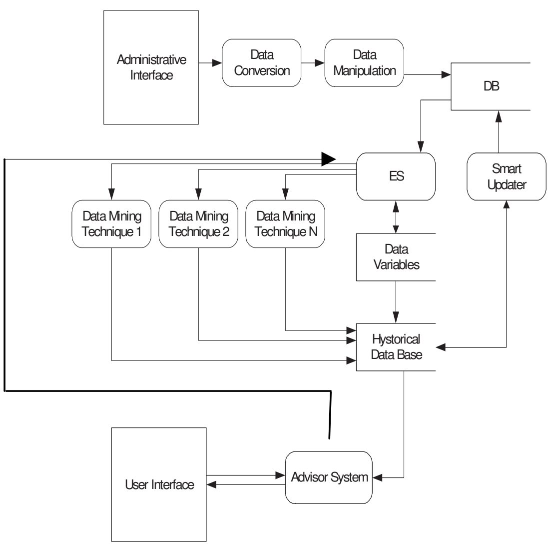


Fig. 1. Simplified Schematic Diagram

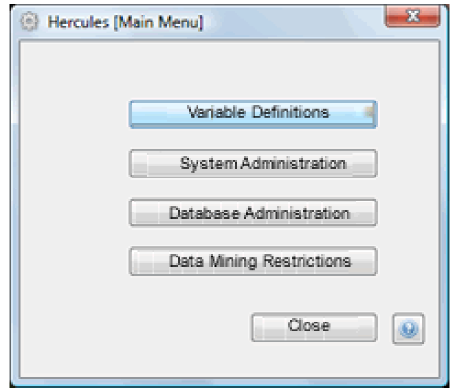


Fig. 2. HERCULES main menu

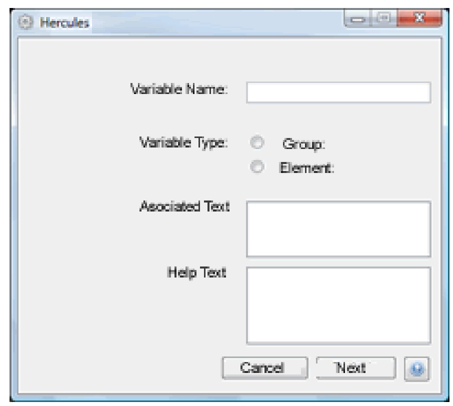


Fig. 3. Variable Creation

Data generator subsystem is another important module. Its main goal is information unification for later DM. It generates some output files from all inputs. Figure 6 depicts the three modules that implement the three main steps previously described here:

1) Importer – its main goal is to import from every source into a common repository. It reads a configuration file with the type of formats to be processed. It can perform temporal transformation over data during exportation, in order to keep temporary datafile information synchronized.
2) Processor and filterer – it merges data inputted into an unified database, joining, filtering and projecting sub-tables as needed. It performs intensive usage of dynamic SQL generated upon user choices and definitions.
3) Exporter – builds a set of files usable for internal DM toolkits.

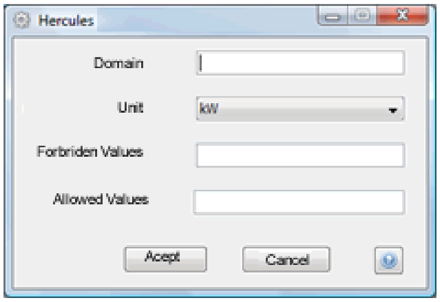


Fig. 4. Variable configuration dialog

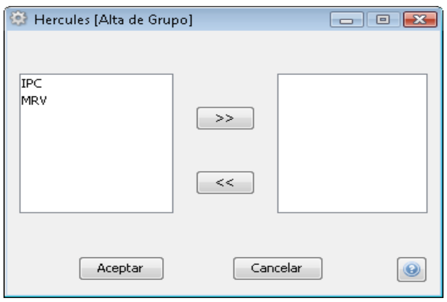


Fig. 5. Aggregated variables definition dialog

The functional prototype diagram is shown in figure 6.

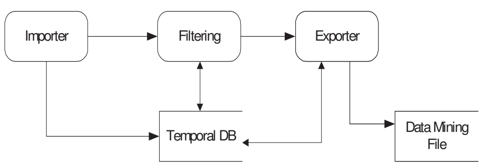


Fig. 6. Data Building process

3. User query administration

Due man-machine complexity, it is required a generic and flexible mechanism for querying. It is acquired by a set of meta-variable usage, composing a large number of predefined customizable queries named here as open queries. Meta-variables can be assigned to specific variables during consulting process. Table 1 shows some open queries, classified by its coverage.

Open Queries List	
Type	Question
Market questions	(1)What [sales-rate] is obtained upon [national-inflation+ change-rate+ export-amount] (2)What is [optimum-time] for [buy-raw-material=x] depending [international-cost+ sales-amount+ exchange-rate] (3) What factors influence on sales
Human Resouces	(4) How much [income-increase-rate] if [expected-profit+ exchange-rate+ truancy + productivity]

Tabla 1. Open Queries samples

It can be seen tha some cases replaces meta-variables with a specific variable (for instance sales-rate in sample 1). There is also a special type of question such as case 3. It does not has an explicit meta-variable but, in order to answer it, it is required a task called expansion before processing. During expansion the query is related and replaced by a set of expanded queries where there are no meta-variables. Both meta-variables and expansion procedure are parametric and can be defined by system administrator.
Although open queries are very flexible, there are many other questions that are frequently used and thus easily foreseen because they are within the set of classical questions. for such cases there is a set of restricted queries without any meta-variable nor variable. Table 2 shows the complete list designed for HERCULES. Some of them have been considered for future implementation.

RESTRICTED QUERIES
¿Is it convenient to invest in a new machine for production if it costs XX?
¿What is the minimum production of XX in order to avoid outsourcing?
¿Is it good to export product XX?
¿What is the annual revenue for sailing XX units of product YY?
¿How affects the enterprise an XX % of inflation?
¿How much lost will have the enterprise for producing product XX?
¿How increased sales marketing knowledge of product XX cost?
¿How much time is required for redeeming $XX spent on invertions?
¿What to do if demand decreases?
¿How to reduce fixed costs and variable costs?
¿How to keep final prices? (when costs are mutable)

Tabla 2. Restricted queries

4. Economic models and specifications

Generally speaking, models and specifications have been extracted for: macro-economy, micro-economy and socio-politics perspectives. In this section the first two will be briefly described and the next section is focused in socio-politics concerns.
Macro-economy deals with Economics aggregates, monetary expansion and Recession, total goods and services, Economic growth, inflation rate, unemployment, Balance of Payments and Exchange Rate. It also covers production growth and employment along time periods (since it reflects economy growth) and short term variations which constitutes business cycle. [10] [11] [12]

In spite of the big contrast between macro and micro-economy, there is no conflict between them. Macro-economy is just the aggregation of markets. The difference stays mainly in the focus and presentation. The scientific study of how to define, explain, anticipate economic phenomena using formalized tools to asses, model and structure information is named Econometrics.
Among main projects in the area, it worth to mention LINK project. It actually includes almost 80models representing a total of 73 national economies and 7 regional aggregation. One of them is the Wharton-UAM model. It defines about 30000 study variables. [13] [14] [15]

Some other tools and models are used un macro-economics field, mainly related to the evaluation of politics effects and some future projections. There are also certain regional models used by public entities from government or not, to make projections. Among them are XS21 model family, EDGE models (Stochastic General Equilibrium), ARIMA (Auto-Regressive Integrated Moving Average), REM (Run time Execution Monitoring), ARMA (Autoregressive moving average), etc. [16] [17] [18]

Macro-economic variables are defined in HERCULES to cover several kinds of analysis. All of them feed an internal module that selects important variables and models for the actual query to be solved. Typically selected variables are the ones that represent basic national equations, such as: GDP (Gross Domestic Product), C (Private consumption), G (public consumption), I (Internal Gross Investment, private or public) , X (Goods and services export), M (Goods and services Import). Other equally important are accounts related to Balance of Payments, Money Aggregates, Price variations, and Commercial exchange rates.

Due mentioned variables and their components are usually used for temporal series extraction, HERCULES design includes them along with additional ones as they result related with the problem. It is important to note that the result is derived from a set of previously known rules that model knowledge about database content and usage in every similar case. Data is updated periodically to improve precision. [19] [20] [21]

Other complementary perspective is the micro-economy. In this context specific metallurgy enterprises that are quoted companies (in Argentina Acindar, Siderar and Aluar). They were elicited because they are the most important at the present. Related information was extracted from trustworthy sources. Sales, production amount, statement of financial position, stock are variables considered here. All this information is required to be able to perform a deep analysis of the enterprise. All these variables derive in a big set of data useful to cross information, infer external changes and business cycle. From the resulting analysis a numeric and nominal set of values are fed to an Expert System.

5. Socio-Politics perspective variable management

The relationship between culture and socio-politics has been studied by sociology since its firsts stage as a discipline derived from industrial, economics, sociology and politics changes in the last century [22][23][24]. Society structures reflect consolidation of such relation, making them relevant for socio-economics assesment of a region, country or sector. There is a deep link between culture (values, beliefs, attitudes, regulations, etc.) and socio-economics development. (1) It has been observed also that social matter is crucial in determining enterprise performance [25] .

Henceforward, to relate economy with social sciences is a subjective, relative, complex and comparative concept that becomes one of the key recurrent social science topics. (2)

next there is a brief description of the considered variables in this field:

a) demographic variables: a preliminary study was performed with classic demographic variables. The provide information about regional context for the enterprise. To betray specific data for every case study, public well known source are referred. For instance, every State has a statistic center that promulgates official evaluations. In several cases there are also private entities that are good alternatives.

b) Socio-economics variables: this kind of variables are defined for addressing coverage issues. The first obstacle is to provide data faithfulness and credibility. Most of the studies and papers are written by sociologist and epidemiologist (3) (3). That biases the focus on, shifting the approaches in a way not totally compatible with an automatic processing software device.

c) Socio-politics variables: every variable is analyzed to split it into several measurable indicators. A good number of these indicators are covered by macro-economics, micro-economics and socio-economics variables. This point of view also focuses on the governability and legal security or each region, considering precedents.

In a general sense, variable are a demonstration of models built over reality. For instance, the human development index has been extended to cover several welfare extra topics. It can be measured as a collection of economics and sociologics indexes developed by United Nations Development Program(PNUD)(4). Particularly, human development index (HDI) is based on the individual income, health (life expectation) and culture (literacy and elementary, high school and university student attendance rates).

Variables are collected, protocoled and uploaded into the prototype. Data are organized into a set of tables and are pre-processed according to actual requirements to get them able to be processed by DM. This module also may perform conversions to different formats and a set of precoded adaptations.

Once prepared and converted, exploratory analysis is performed and the resulting data is used to feed the IES and get the best interaction and behavioral model. It is important to note that typically the classical variables in this field lack of a probed model for integrate them into an equation.

6. Actual implementation status

The user interface for input data and configure variables are implemented as long with adaptation modules, database and the IES with a few rules. Actually the full set of rules are being deployed to model DM initial results. Rules have been developed for modeling part of macro-economy, micro-economy and socio-politics.

7. Conclusion and future work

HERCULES prototype has been presented, its overall design and main architectural characteristics. It remains to study self tuning approach of internal working parameters and further open and restricted queries. Alternates pending to be covered are problem automatic break down into several subproblems, Expert systems full adjustment and results GUI.

8. References

[1]. E. Viglizzo. Sustentabilidad ecológica y económica de la ganadería. (1999). Argentinean Symp. on Animal Productivity.

[2]. J.C. Ascough, G.H. Dunn, G.S. McMaster, L.R. Ahuja, A.A. Andales. Producers, Decision Support Systems, and GPFARM: Lessons Learned From a Decade of Development (1997). MODSIM 2005.

[3]. INTA, Instituto Nacional de Tecnología Agropecuaria. (2009) URL: www.inta.gov.ar/bariloche/ssd/rn.htm

[4]. J. Ascough, M. Shaffer. D. Hoag, et al. GPFARM: An integrated Decision Support System for Sustainable Great Plains Agriculture. (1998) 10th International Soil Conservation Organization Meeting. Purdue Univ.

[5]. V. A. Ferreira, D. G. De Coursey, B. Faber, L. Knapp, R. Woodmansee. Terra Tools and Techniques for Ecosystem Mng. (1995) Cong. Res. Serv. Report to Congress, Library of Congress, Washington, 95-430 SPR: 47-59.

[6]. J. Milne, A. Sibbald. Modeling of grazing systems at the farm level. (1998) An. de Zootechnie, Vol. 47: 407-417.

[7]. L. Silva, B. Revello. Construcción de un sistema de apoyo en la toma de decisiones para el área gerencial del hospital de clínicas (2000). Computation Institute. School of Engineering. Universidad de la República.

[8]. J. Watson, M. Rainer, Koh. Critical review of EIS in the context of intelligence support. (1991) In Managing strategic intelligence: techniques and technologies. Mark Xu. ISBN. 1599042436

[9]. J.Wetherbe, Executive Information Requirements: Getting It Right, (1991). MIS Quarterly, pp. 51-65.

[10]. R. Dornbusch, S. Fischer. Macroeconomía. (2004) 5th edition. Ed. Mc Graw Hill, pp.3. ISBN 8448120558.

[11]. P. Capros et al. Applied Econometrics Association. (1991). Conf. on Int. Energy Market Modelling. France.

[12]. A. Pulido. Proyecto Link. http://www.un.org/esa/policy/link/, Madrid Coll. – School of Economy.

[13]. S. Martínez Vicente, M. Blanco Losada, E. López Díaz-Delgado. Modelo de Planificación regional: la familia de modelos XS21 (X-Siglo XXI). (2009). URL: www.salle.url.edu.

[14]. G. Janezic, D. Branca, D. López De Luise, J. Azcurra, J.Musella, F Milano, S.Bosio. Sistema de Soporte de Decisiones en Contextos Industriales. I CONGRESO PERUANO DE INVESTIGACIÓN DE OPERACIONES Y DE SISTEMAS (First Peruvian Congress on Research Operation and systems). COPIOS 2009. Nov. 03-05. Lima, Perú

[15]. J. Escudé, G. J. Progresos en Economía Computacional. (2008) Argentinean Central Bank. Argentine 42. ARGEMmy.

[16]. Ministerio de Economía de la República Argentina (Argentinean Economy Department). (2009) URL: www.mecon.gov.ar.

[17]. Dirección Nacional de Cuentas Internacionales de la República Argentina (International Accounting Center of Argentinean). (2009). URL:www.mecon.gov.ar.

[18]. Instituto Nacional de Estadística y Censos de la República Argentina (National Institute of Statistics and Polling); INDEC (2009). URL: www.indec.gov.ar

[19]. Grupo De Monitoreo Macroeconómico de la República Argentina (Argentinean macroeconomic monitor Group); GMM (2009). URL: gmm.mecon.gov.ar

[20]. M. Ahijado. Microeconomía. Teoría. (1994); Ramon Areces statistical Center. 01 ed. ISBN: 978-84-8004-112-6.

[21]. G. Brunetti, U. Collesei, T. Vescovi, U.Sòstero. La librería como negocio: economía y administración (Library as a business: economy and management). (2004). Fondo de Cultura Económica Ed. ISBN: 978-968-16-7383-3.

[22]. La cultura empresarial en Andalucía. Un estudio sociológico de la pequeña empresa. *Pablo Galindo Calvo, Universidad de Granada. Departamento de Sociología, Campus de Cartuja. 18071 Granada galindo@ugr.es*

[23]. Galindo Calvo, Pablo. *Estudio sociológico del pequeño empresario granadino: la cultura empresarial. Granada: Editorial Universidad de Granada 2006*

[24]. Noemí Guillamón Cano. *Variables socioeconómicas y problemas interiorizados y exteriorizados en niños y adolescentes*

[25]. MANUEL ALFONSO DE LA VEGA TOLEDO. *Análisis geográfico de la relación entre variables biofísicas y socioeconómicas y su influencia en categorías de pobreza en el territorio nacional. Universidad Rafael Landívar. Facultad de Ciencias Ambientales y Agrícolas*

OPTICALLY TUNING TC IN ANY SUPERCONDUCTOR

S.Curatolo**

CZT Inc. Lawrence Kansas 66044 USA

Abstract

The Excitonic Enhancement Mechanism [1] has been demonstrated theoretically and experimentally by CZT Inc. as the theory for superconductivity in hole carrier systems like YBCO and BISCCO[5][6], and novel ones like boron doped silicon[and [2][3][4], because excitons are like hydrogen atoms and possess excited energy levels, these Cooper pair exchange particles can be excited by suitable incoherent infrared light [5][6], resulting in large enhancement of the superconductivity-transition temperature [5][6]. It has been shown experimentally that EEM Tc enhancement for bulk and thin film YBCO, with optimum YBCO bulk system, and by observing the diamagnetism before and after IR light was applied, that the transition temperature increased ~80°K, from exactly as predicted by the EEM theory [5][6]. These original measurements were independently verified on an oxygen deficient YBCO thin film (dirty superconductor-bose-eintein condensate) by measuring the transition temperature using AC voltage versus temperature measurement before and after IR light was applied. The transition temperature was independently verified to increase by ~17°K, from 69.5°K to 86.5°K [5]. Because, the increase in Tc has been shown to be exactly as the value calculated by the EEM theory [6], it makes the EEM theory an invaluable tool for studying and predicting Tc enhancement in any cuprate based HTS with incoherent IR stimulation [5]. As a result of these experimental confirmations in YBCO, in this study we use EEM to predict and compare Tc enhancement in 8-PHASES of TBCCO and Hg-TBCCO and Pb-Sr-TBCCO-HTS-(TBCCO-2122,2212,2213, and 2223,1212,1223,1234,1245) against the proven YBCO with EEM. Since High-critical-temperature (HTc) superconductors like TBBCO, have important applications in many electronic and electrical devices due to no-resistance, low energy loss, exclusion of magnetic fields and special quantum electronic characteristics like the Josephson effect, this study shows an enhanced Tc of Near Room Temperature for all TBCCO 4 phases and for Hg-TBCCO. The observations of data behavior near the Dirty Limit are of much importance to the shortcomings of the SI phase. This result points to the ability to bridge the gap of the obvious well known limitations of Tl-Cu-based-HTS. A Near-Room Temperature TBBCO will enable the less-optimum thin film processes to be available affording a more economical superconductor with higher Tc performance and less power demand, having more phase stability, easier manufacturability and wider applications.

** Corresponding author: Susana Curatolo, President CZT Inc. cztinc@hughes.net

1. Introduction

Superconducting systems that are hole carriers and considered to be BCS like boron doped silicon [4], are in fact EEM-Excitonic Enhancement Mechanism, as their charge carriers are predominantly p-state holes, and both have an empty s-state conduction band above the valence p-band. Their relatively low Tc values, according to EEM, were a result of smaller band gap and excitonic binding energies. These latter systems have some common features with the Cuprates, such as their charge carriers are mainly ‘p’ orbital valence band holes, and that their conduction band is an empty ‘s’ orbital state. It is worthy to note that the main reasons pointed to their differences are first, the substantially higher Tc value for some cuprates, and secondly, the presence of the‘d’ orbitals in cuprates, which is absent in the other systems. In the cuprates, it has been found also that the superconducting Tc value is correlated to the number of CuO_2 layers in its crystal structure, with a maximum reached at 3 layers. On top of this feature, Tc varies according with hole carrier density in an inverse parabolic form. The Excitonic Enhancement Mechanism or EEM [1] for the cuprates, also possesses these features [9]. Bismuth and thallium superconducting cuprates with structures related to the Aurivillius family of oxides with Tc's in the 100K region also possess these features. Thallium cuprates of the Tl-Ba-Cu-O-2122,2212,2213,2223 systems show Tc onset in the 100-120K region, with resistivity susceptibility behavior as high as 130K for Tl-Ba-Cu-O-2223(optimum thin film) sample[10]. The basic EEM is based on a two band model, with a partially filled valence 'p' band, and an empty 's' conduction band, separated by a positive band gap G. In fact this theory model is satisfied by the actual band structures of all the above-mentioned novel superconductors. Apart from Tc, the EEM model also predicts other novel properties in both the normal and superconducting phases, such as the sign changes that were observed in the mixed-state quantum Hall effect [10], and the thermal electric power [11].It has already been determined experimentally and theoretically that this specific geometric model of EEM is able to produce the following properties known to be associated with cuprates like YBCO[1][6] and now to other cuprates Tl-Ba-Cu-O-2122, 2212, 2213, 2223,1212, 1223, 1234, 124 5. First, Tc correlation to carrier density and the number of CuO_2 layers [1][6], as mentioned earlier. It is, therefore, our intention in this paper to predict that the presence of excitons in Thallium based cuprate HTS behavior and that it is the same pattern as YBCO which was proven experimentally[5][6]. The exciton contribution to superconductivity has been verified experimentally through infrared light irradiation [5] [6]. I present the analysis and calculation for the TBCCO system using the EEM YBCO model which predicts an enhancement to near room temperature for some of the TBCCO systems. I explain the positive consequences of the enhancement in terms of the system's known phase instabilities in the deficient range, the dirty limit and in the dirty range before system collapse.

2.Theoretical analysis and calculation of TBCCO using EEM Theory

Our earlier work provided detailed discussion of the EEM Theory [5] [6] [1], particularly the two dimensional model [1] [5] [6] developed specifically for the cuprate systems. Since EEM has been proven for YBCO system theoretically and experimentally [1] [5] [6], and the aim of the current paper is to use this theory's validity for YBCO and apply it to the TBCCO system, including Hg-TBCCOs and Pb-Sr-TBCCOs. The implications from the behavior to BCS will be discussed. We will discuss our complete analysis based on the YBCO model, and calculate the TCU enhancement for the various TBCCO systems. All in all, we consider 19 superconductors using their optimum and deficient published values [6] [11] [12] [13] [14] [15] [16].This model is transportable across HTS-Cuprates like BiSCCO [5] [6], and theTl-Ba-Cu-O-2122, 2212, 2213, 2223, 1212, 1223, 1234, 1245 we are about to present since they are based on the same cuprate-geometry. The EEM theory is based essentially on the dipolar interband interaction between an intrinsic partially filled valence 'p' band from the Oxygen ions and the empty conduction 's ' band of the Ba/Sr ions, which lay above the CuO_2 layer. Since excitons have excited levels similar to phonons, we therefore expect that if a large enough population of such excited excitons are maintained, Tc can be enhanced [5][6]. So, the renormalized band structure results as given in reference [1] [5] [6]. The KEY Equations [1] [5] [6] are used in this study to perform the calculations and analysis on the TBCCO systems. The excitons are stable if and only if, its binding energy exceeds the thermal fluctuation energy, which creates them in the first place. Should this condition be satisfied, global formation of localized excitons will occur in the structure, quite similar to the phonon lattice. Because excitons are Bosons, when coupled to the mobile holes can replace the phonons in the BCS theory, and lead to hole-hole Cooper pairing, thus producing superconductivity [5] [6]. Consequently, to obtain Tc based on exchange of excitons, we first obtain the exciton's excitation energy, [5] [6]], which is given by λ_3 minus the exciton binding energy to the lattice $\hbar\Omega$. It is important to note that the excitonic binding to the lattice vanishes along the cell diagonals. The space average value of which is of order eV, and is quite stable in the superconducting phase. Hence the lateral thermal fluctuation that destabilizes the bound exciton occurs in the normal phase. Because the excitonic levels of the system are inside the band gap, it implies such exciton is localized or bounded to the crystal lattice. Hence, when the exciton is delocalized, it will decompose into a free electron and a free hole. Such a picture for the normal phase gives raise to the two carrier's model [5] [6].]This bound excitonic excitation energy $\hbar\Omega$ now replaces the Debye phonon energy in the Cooper pair BCS energy gap Δ_{BCS} . Since the superconducting Tc is directly proportional to the BCS gap Δ_{BCS} , it follows for the cuprate, Tc is proportional to the exciton excitation energy. It is precisely this feature that allows EEM to explain the Tc dependence on the hole density[5][6]. The photo-excitation necessary to excite the excitons into the first excited level is of energy [5][6] of $\hbar\gamma_1 = \frac{8}{9}E_o$. Since, this photon energy is the same for exciting the system from either above or below Tc, as $G^* \cong 0$ and $\hbar\gamma_1$ is of order 1 eV, it follows the photon wavelength is much larger than the crystal cell dimensions. Hence when we pump the system with such IR photons [5][6], the populated electrons with the highest energy level are excited. There is a difference where these electrons are located depending on the system's physical phase. Given that the proportionality of Tc on the exchange exciton excitation energy $\hbar\Omega_1$ and assuming J and N_F^* remain relatively constant, we can estimate the enhancement of Tc to Tc* from the ratio[5][6] of $\frac{T_c^*}{T_c} \cong \frac{\hbar\Omega_1}{\hbar\Omega}$. It follows that, how much Tc is enhanced to Tc* depends critically on the sample's original Tc value obtained from experimental sources for our study, because Tc of the sample is dependent on the 'p' hole density n [5][6], through the excitonic excitation energy shown here [5][6] $\hbar\Omega$, with $\overline{E}_o$ as optimum excitonic ground state binding energy [5] given as $\overline{E}_o \equiv \frac{16Rm^*}{\varepsilon^2}n_o^2$.

In our previous work in YBCO [1][5][6], the $\Delta^{(o)} = 2.07\ eV \cong 2\overline{E}_o$ . .So, in order to prepare at this point, to perform TBCCO Tc enhancement calculations and compare against our previous and proven study of infrared light enhancement on YBCO [5][6] on the critical temperature, namely Tc*, we consider two separate cases. These two cases Optimum and Deficient will govern the subsequent calculations for TBCCO exactly, we just use the corresponding Tc to the corresponding superconductor. First, with the optimum YBCO sample [5][6], and the optimum TBCCO sample, and $n = n_o$, then

$$Tc^*\left(n_o\right)_{HTS_{OPT}} = optTc \times \frac{\left(\frac{26}{9} - \frac{\pi-2}{\pi}\frac{18}{13}\right)}{4/\pi} = optTc^* ,$$

and we get, regardless of system and phase:

$$HTS_{OPT}\frac{T_c^*}{T_c} \cong \frac{\hbar\Omega_1}{\hbar\Omega} = 1.870967742 .$$ For $YBCO_{OPT}\Delta\eta = 0.1$ we obtained a maximum enhancement of ~90% of Tc[5][6], see Fig.2, and Fig.1 for NOS Tc behavior only. The Second case (see Fig.1): For the O2-deficient range (BoxII), Dirty Limit(DL) and Dirty(BoxIII), (dirty-HTS = maximal-defects-BCS like) YBCO-def $Tc = 70^\circ K$ [5][6] and for the O2-deficient TBCCOs see Fig.1. Now need to find the oxygen deficient factor $\Delta\eta \equiv \left(n_o - n\right)/n_o$ for EACH of these samples. First. $\Delta\eta$ is obtained from our equation using the values for optTc and defTc from the sources obtained for each superconducting system [6][10][11][12][13][14][15][16],and-by-comparing $T_c\left(n\right)\ to\ T_c\left(n_o\right)$ -that is, defTc to optTc, and solving for $\Delta\eta$ we get all the corresponding values listed in Table.1.(See Fig.1). Then substituting $\Delta\eta$ into E_o [1][5][6], we obtain

$$\frac{T_c^*(n)}{T_c(n_o)} = \frac{\left\{2-(\Delta\eta)^2+\frac{8}{9}(1-\Delta\eta)^2-\frac{\pi-2}{\pi}\frac{4}{2-(\Delta\eta)^2+\frac{8}{9}(1-\Delta\eta)^2}\right\}}{4/\pi} .$$

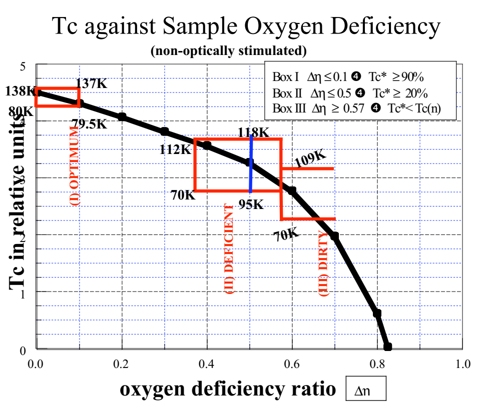
Thus, using this equation [1][5][6] we get $\Delta\eta$ for all 19 superconductors. Fig.1a&1b shows that NOS Tc behavior for under-doped and over-doped sample conditions, and they are mirror-images. BoxI-II&III are the same for - $\Delta\eta$ **.** Figure.2.compares NOS-&-OS Tc Enhancement ratios vs $\Delta\eta$ **,** and we clearly observe the linearization under OS conditions, following the same pattern observed for the YBCO system and BISCCO[1][5][6], and that for ovedoping it is a mirro-image. In fact, the non-coherent-infrared light(not a laser)[5][6] is able to stimulate Tc enhancement even when sample is under-doped or over-doped in BoxII and DL, enhancing Tc ≥ 20% within BoxII, < 20% within DL(dirty limit: underdoped or overdoped), and losing enhancement in BoxIII. Because we have normalized by dividing over optTc, Fig2a,b is universal for all cuprates! Then,

$$Tc^*_{DL} \le Tc^*_{DEF(\Delta\eta=0.38 to 0.53)} \le Tc^*_{MID(\Delta\eta=0.1 to 0.38)} \le Tc^*_{OPT(\Delta\eta=0 to 0.1)}$$

$$Tc^*_{DL(\Delta\eta=-0.53 to -0.57)} \le Tc^*_{MID(\Delta\eta=-0.1 to -0.38)} \le Tc^*_{OPT(\Delta\eta=0 to -0.1)}$$

Fig.1.a Mirror-Image NOS Tc behavior vs $\Delta\eta$

Fig.1.b NOS Tc Behavior vs $\Delta\eta$



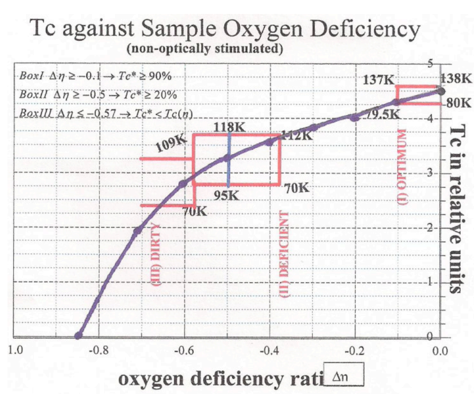
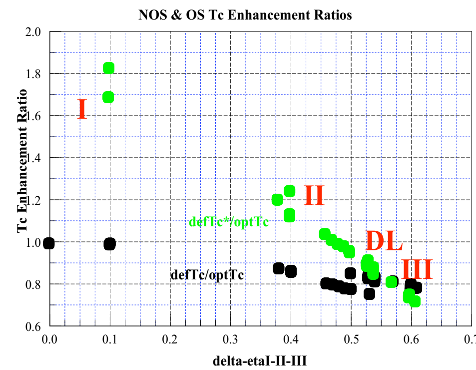
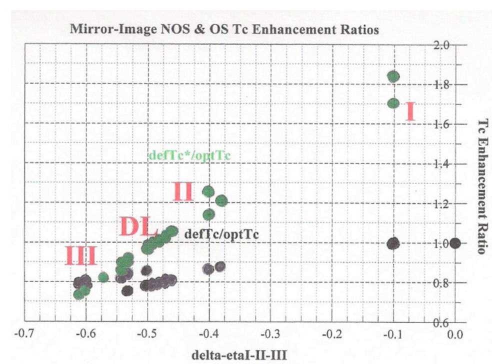


Fig.2a Mirror- NOS & OS Tc Enhanced Ratio vs $\Delta\eta$

Fig.2b NOS & OS Tc Enhanced Ratio vs $\Delta\eta$





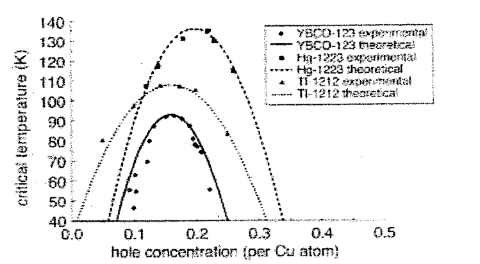


Fig.3 Phase diagram showing the under-doped and over-doped regions of various Cuprate systems[1]

Summary Discussion and Conclusions

NOS Behavior In Fig.1 a & b, we start first with the Tc vs $\Delta\eta$ I-II-III. This plot is intended to show the behavior of Tc for the standard normalized for optimum at $\Delta\eta$ = 0.0-0.1 (Box I), under doped and over doped deficient range : $\Delta\eta \pm$=0.38-0.5-0.53 (Box II), dirty limit range : $\Delta\eta \pm$= 0.53-0.57 (DL) , dirty range : $\Delta\eta \pm$= 0.57-0.7. This curve is also valid for over-doping (-$\Delta\eta$). So the shape is the same, but a mirror image from 0.0. The Tc values are for highest and lowest for each range, for the Tc and defTc values from the literature sources consulted. We can observe in Fig.(1) very clearly the non-enhanced [high,low] limits of Tc for each range. For Box-I : 138°K,80°K-137°K,79.5°K; Box-II : 112°K,70°K-118°K,95°K; 118°K,95°K < DL < 109°K.70°K; Box III > 109°K,70°K . The defTc/optTc vs $\Delta\eta$ in Fig.2a,b demonstrate the Tc behavior deficient/optimum ratio versus sample oxygen deficiency ratio. For Box I the ratio is between 0.98-1.0, for BoxII the ratio is between 0.775-0.875, for DL the ratio is between 0.845-0.825, and for BoxIII the ratio is 0.815-0.775. Notice the splitting at $\Delta\eta$ = 0.5-0.53 where the deficient range ends and the Dirty Limit begins. This is critical because it is where we observe SI (superconductor-insulator phase) behavior causing Tc to drop for both underdoped and over-doped. It is an area where the material behaves like a BCS-Bose-Einstein Condensate, and superconductivity is lost, particularly true for metal-insulator type metal oxides. Oxides are perceived in the industry as having inherent unstable nature[12]. There are manufacturing issues too, intrinsic local stoichometric -defects arise in HTS from the insertion of cations in the wrong layer and defects of the oxygen sublattice, forcing manufacturing to adjust oxygen content in compound-specific off stoichiometric ratio in order to optimize the superconducting properties[10]. This we shall see, is very much mitigated using incoherent-Near-IR to achieve superconductivity in the SI range as opposed to far-IR-stimulated Cu-O plane[17].Over-doping and Under-doping mirror each other based on this method. The manufacturing advantage is obvious from a production and quality control point of view. OS Behavior No parameter adjustments are necessary, for the cuprate geometry of the band structure provides the means for calculating the deficiency ratio and the Tc*, given a Tc for that superconductor. The incoherent IR λ is calculated based on equation (29) of [5][6] . As excitons can be excited, they can also be destroyed. To destroy the excitons we irradiate them with incoherent-near-IR photon of energy at least equal to E . For the YBCO system, E is approximately 1 eV. Since the natural relaxation of the excited state is independent of the wavelength intensity, and Tc* is wavelength intensity dependent, then the key is to maintain equilibrium. Since the optical dependent excitation rate and the density of excited excitons are related to one another, and with enough IR photons uniformly absorbed over the entire surface of the HTS material film Tc* is realized. since the superconducting Tc is directly proportional to the BCS gap Δ_{BCS} , it follows for the cuprate sytem, that Tc is proportional to the exciton excitation energy [5][6]. So, the defTc*/optTc ratios against the oxygen deficiency ratio $\Delta\eta$ show basically a linear relationship. For BoxI the ratio is between 1.7-1.85-0.98, for BoxII the ratio is between 1.25-0.98, for DL the ratio is between 0.92-0.82, and for BoxIII the ratio is 0.82-0.73. The difference in ratios between NOS defTc/optTc and OS defTc*/optc is on the order of a factor of 2 BoxI, whereas the difference is about 1.5 for BoxII, and it drops to 1.25 for DL and the difference in ratios for Box-III is 0.9 where the Tc* is below Tc. This indicates that the enhancement responds within the predicted enhanced limits and the proportionality of the enhancement to the exciton energy. The difference between NOS and OS ratios for the different $\Delta\eta$ ranges can be clearly seen in that the OS linearizes the response and we have the corresponding enhancements of 20% for def under-doped and over-doped, 10% for DL under-doped and over-doped as opposed to 90% for the optimum. These represent a significant broadening of the phase diagram of the experimental data sources seen in Fig.3 by as much as $n_o(1 \pm 0.89)$, and thereby exceeding the industrial averages substantially. We seek industrial deployment of these films given the obvious advantages in the SI as compared to literature.

References

[1] K.W.Wong,W.Y.Ching,Physica C, 416 (2004) 47
[2]I.N.Makarenko,D.V.Niforov,A.B.Kykov,O.K. Mel'nikov,S.M.Stishov,'
Anisotropy of the electrical resistance of single crystals of Htc YBCO', Pis'ma
Zh. Eksp. Theor. Fiz. No.1, 52-56, 1988, USSR.
[3] J.G.Bednorz and K. A.M ulller,'Perovskite – type oxides-the new approach to

HTcSC',Rev. Mod. Phys.Vol.60,No.3,July 1988
[4]Cava Lab: Intermetallic Superconductor research
[5] Susana Curatolo, Kai Wai Wong,-USPatent No.20080293577,IPC8 Class:AH02H700FI,USPC Class: 505850;Title: Method for Operating a Superconductor at a Temperature Higher than the Superconductive Temperature Tc.
[6] S.Curatolo,K.W.Wong,Progress in Superconductivity-EEM:The Exciton Enhancement Mechanism Theory and Experimental Evidence of Optically Enhanced Tc in High Tc Superconductors'.ISBN 978-1-60021-694-7, 2008. Nova Science Publishers Inc. Oliver Chang- Editor
[10] Hott, Roland; Narlikar,A.V.-High Superconductivity 1-Materials;Table-1, page 4, Springer, September 2004, Berlin. Forschungszentram Karlsruhe,Institut fur Festkorperphysik; P.O.Box 3640, 76021 Karlsruhe Germany
[11]Rao, C.N. High Temperature Ceramic Oxide Superconductors, Sadhana,Vol 13, Parts-1&2, July 1988, pp.19-35.Table 10, p.29.
[12] S. A. Sunshine and T.A. Vanderah. Preparation of Bismuth- and Thallium-Based Cuprate Superconductors.Chemestry Division Research department,Naval Weapons Center, China Lake, CA 93555 Office of Naval Research. Grant Contract N00014-91WX-24155, R&T CODE 4134053,TECHNICAL REPORT NO.2 Published in 'Chemestry of Oxide Superconductors' by T.A. Vanderah. Noyes Publications, 1991. Table 1, p.5
[13] DuPont Superconductivity : Development of High temperature Superconducting Josephson Junction Device Technology-Final Report January 1996-January 1998 Program Manager Kirsten E Myers; In situ deposition Thallium-containing oxides. K.E.Myers Du Pont CR&D Experimental Station-E304/C110 Wilmington DE, USA;page168/SPIE Vol.2697:Sections 4.3 and 4.4.
[14]-B.R.Xu, Y.Xin, G.F.Sun, K.W.Wong, XII-International.Conference.onThermoelectrics; Phys.Lett. A 192 (1994)p122-124
[15] F. Gouternoire et al, 'Substitution of Hg for Tl in Hg-Tl-BCCO-2223',Solid State Communications 90(1,)April-1994, p.47-51
[16] Hur, Yong Hi Park, K J Park,Jong C.(Daejeon) US Patent No. 5492885, 1995
[17] Vaglio, Ruggero 'The New Italian CNR Institute SPIN', CNR-SPIN/IEEE-CSC&ESASEuropean-Superconductivity-News-Forum-(ESNF), No.12, April 2010,p.7-9

Wet and Dry Abrasion behavior of AISI 8620 Steel Boriding.

I. Hilerio C.[1], M.A. Barron M.[2], R.T. Hernandez L[3], A. Altamirano T[4].
Universidad Autónoma Metropolitana Unidad Azcapotzalco Departamento de Materiales.
Av. San Pablo 180 Col. Reynosa-Tamaulipas 02200 México, D.F.
Tel. 55 53 18 90 00 ext. 2131. Email[1] ihc@correo.azc.uam.mx E. Mail[2] bmma@correo.azc.uam.mx
E. Mail[3] hlrt@correo.azc.uam.mx, E. mail[4]aat@correo.azc.uam.mx

ABSTRACT

Wear tests were developed in two different equipments, designed by Tribology group UAM A in Mexico under the standards ASTM G-65 to dry conditions and ASTM G-105 to wet conditions. Because the standards use different parameters, it was necessary to equal these for both tests, as: load, sliding distance, lineal velocity in the contact point, particle size and hardness of the wheel. Steel used in this paper is employed in machine's components. With the objective of observing the benefit of boriding process on the AISI 8620 steel, was carried out the comparison of this steel with the basic material and the same with the boriding process. The wear rate is significantly less in wet condition than in dry condition; this for the reason that the water acts as a lubricant or refrigerant and hide the abrasion mechanisms. The surface of the boriding sample is very hard and consequently with little loss of the mass; to both tests presented a good behavior to abrasion.

Key words: Abrasive wear, dry abrasion, wet abrasion, boriding.

INTRODUCTION

Abrasive wear is the result of three mechanisms; ploughing, cutting or wedge forming; this phenomenon occurs when hard asperities or abrasive particles, presents an sliding process and interacts with a surface.[1-2]. Wear by abrasion is approximately 50% of all wear in industry; due to the interaction at least by two bodies and in many occasions also the presence of a fluid. For this reason, the tests has been evaluated in dry and wet conditions. The standard ASTM G-65 with the title "Standard test method for measuring abrasion using the dry sand/rubber wheel apparatus"; where it describes the configuration and establishes the operation parameters of the tribometer in dry condition. On other hand, with respect to the test under wet condition is described by the standard ASTM G-105, with the title "Standard test method for conducting wet sand/rubber wheel abrasion test". When the operation parameters have been similars, is possible to make a comparison between both tribosystems.

AISI 8620 steel is a hardenable nickel-chromium-molybdenum low alloy steel; used widely in machine's components. Admit different surface treatment and forming process; such as, carburizing, tempering, forming, welding, boriding, etc.[3]. Thermal hardening is produced by heating a material, followed by a rapid quenching in oil or water. With this operation, the surface of the sample has been hardened. There are five used methods to heating the steel; by electrical induction, resistance, flame, laser and electron beam. [4,5]. The boriding process is a thermochemical treatment, when the boron is diffused in a metallic surface, but this process can also be applied to nonferrous metals such as nickel, cobalt alloys and -refractory metals. This can carried out in pack, liquid, or gas; with a temperature between 900 and 1100°C; The boriding steel exhibit extremely high hardness, commonly from 1500 to 2300HV, obtaining a thin layer between 50 to 150µm. The morphology of boriding layer formed on iron has two types: a superior layer FeB and an inferior layer Fe_2B. The second layer is distributed principally in the grain boundary. Both present similar orientation and they have structural characteristics in columns towards to surface [4,6].

EXPERIMENTAL PROCESS

Material preparation

There were three groups of test specimens in AISI 8620 steel; base, thermal hardening and boriding. The first group didn't have any process; there were cut from a bar into rectangular shape, 25.4mm wide by 57.2mm long by 12.5mm thick. The dimensions are equal to all specimens.

Thermal hardening was processed by induction at 900°C, during 30 minutes with a nitrogen flow at $10ft^3/hr$ and quenching in oil. After this, was realized a tempered treatment at 200°C, during 2 hours.

Boriding process was made by a salt bath, where the specimens were placed into a pot and covered by boron salt; this complex was heat up to 950°C, during four hours. Hardness after this process is 850 HV with a load 100gf, during 10s.

Abrasive wear tests

The tribometers have been developed by group in UAM A, by agreement to standards. Figure 1 shows both type of the machines in schematic form. This equipment was used to study the abrasive wear behavior in dry and wet conditions. The next parameters were used to both tests. The abrasive that was used in this work, is called grain quartz sand with size 230-270µm. Load used 200N between wheel and specimen. Sliding distance of 5586m. With a rotation speed from 250 to 250rpm. Dry test required a flow of abrasive from 0.3 to 0.4 Kg/min; by other hand, wet test used a mixture of 1.5kg. of abrasive and 0.940 Kg. of water. Before having started the tests, the specimens were cleaned and weighed by a analytical balance whit a accuracy of 0.0001g. During the entire test there were 5 times to weigh the difference of the mass.

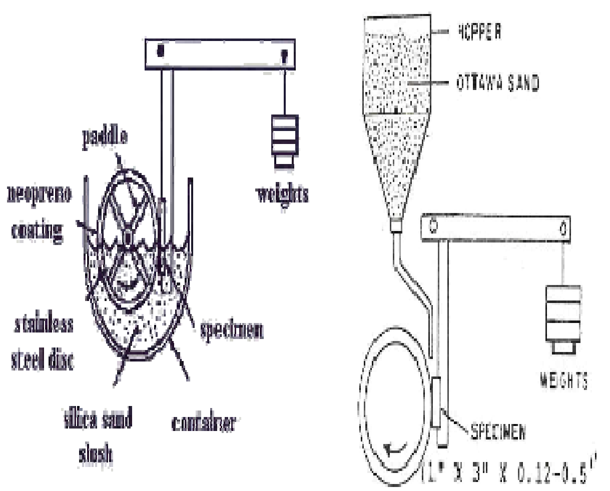


Fig. 1 Schematic of the tribometers. a) under wet condition; b) under dry condition.

Fig. 2: Micrograph of 8620 steel quenched and tempered.

The morphology of boriding layer formed by iron has two layers types: a superior layer made by FeB and an inferior layer made by Fe2B. The second layer is distributed principally in the grain boundary. Both of them, present similar orientation and they have a structural characteristic in columns towards the surface. A metallography in cross section is realised with a chemical reactive (nital 2%) during 6 seconds, it is shown in figure 3.

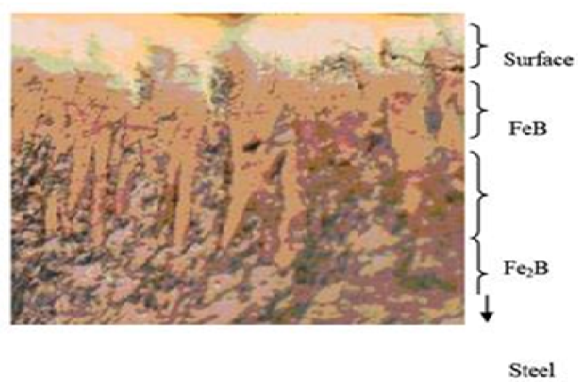


Fig. 3: Micrograph of boriding steel 8620.

RESULTS

The structure of AISI 8620 after hardness, it's shown in figure 2, through a metallurgical preparation for transversal section with a nital reactive (2%) during 6 seconds. Where it has presented a hardness of 400HV; with a load 100gf during 10s.

It is shown the evolution of hardness curve obtained with this process, as shown in figure 4. In comparison with the literature date, this material is very soft, due the sample have had a heat treatment after boriding.

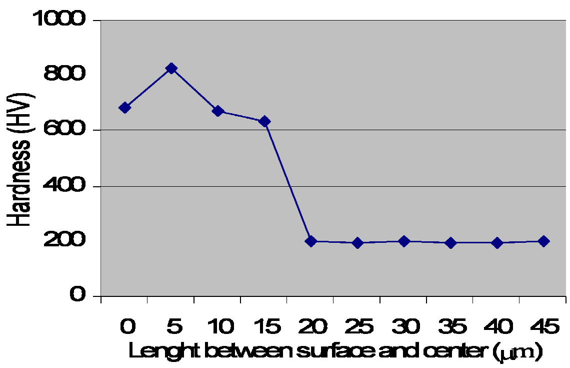


Fig 4. Hardness profile for boriding steel 8620.

Pictures 5 show wear scars obtained with metallic specimen for all steels: base (a), heat treatment (b) and boriding (c).

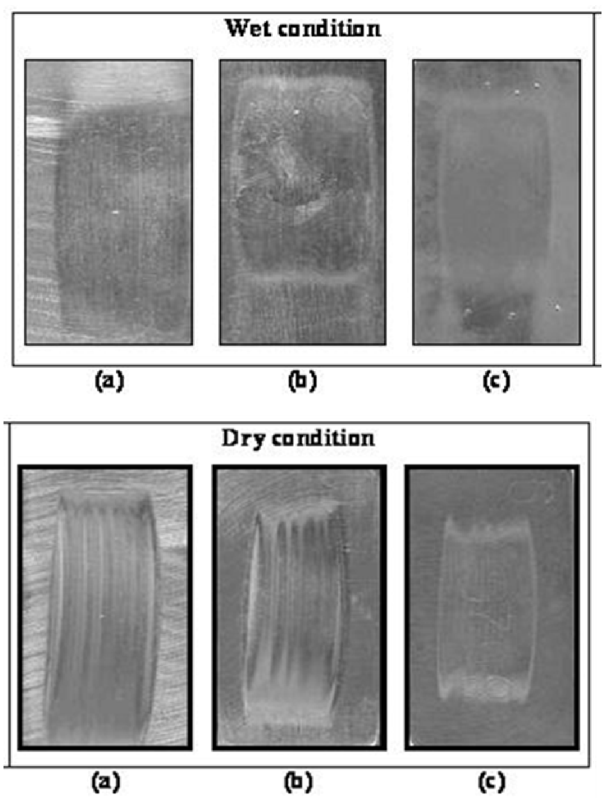


Figure 5: Wear abrasive scars for both conditions.

Abrasion behaviour of boriding steel AISI 8620 for both conditions presents a higher abrasion resistance, as shown in figure 6. It's possible to note a benefit with thermal treatment.

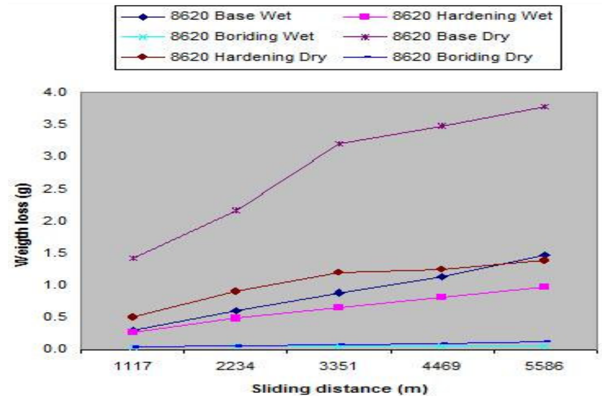


Fig 6. Comparison of behaviour for both wear test condition.

In wet conditions, effect is less severe than those measured for dry condition. This result is explained because in the aqueous tribosystem is formed a water film protection. By consequence, that induce firstly, a difficulty of penetration for abrasive silica sand particles and secondly, an easier slide for sample on metallic surface. For this reason, scars wear are less severe.

CONCLUSIONS

On the basis of the result obtained, it is possible to note that:

- Wear behaviour for a metallic material is modified with a change of its superficial structure. It has been clearly shown this relation in this work. In this study, it has been shown steel behaviours obtained by different surface treatments in function of structure.

- It has been hierarchised the wear resistance property of AISI 8620 steel in this order:
  1) AISI 8620 steel base as the less resistant
  2) AISI 8620 steel hardened by quenched and tempered
  3) AISI 8620 borided as the most resistant treatment

- The borided 8620 steel has excellent properties for both conditions of test, but it obtains better results in wet conditions. So it's possible to establish that borided steel is a useful tool to work in this kind of conditions and resistance against wear.

REFERENCES

1. K.Kato, Wear mechanisms, plenary and invited paper in American Society Mechanical Engineers publications, 39-55 (1997).
2. J-L. Bucaille et E. Felder, L'essai de rayure sur polymères et métaux. Modélisation et approches expérimentales, in Matériaux et Techniques, n° 3-4 (2001).
3. T.G.Mathia, M.Ouadou, F.Saucez, Etat des surfaces et sclérométrie, Ingénieurs et industrie, n°25, 45-48, (1991).
4. E.Bergmann, H.Matthieu, R.Gras, Analyse et technologies des surfaces, vol.4, Presse polytechnique et Universitaire Romande, Ed. Eyrolles (2003).

5 Y.Ono, C-Q.Zheng, F.Hamel, R.Charron, C-A.Long, Experimental investigations on monitoring ans control of induction heating process for semi-solid alloys using the heating coil as sensor, Measurement science and technology, 13, 1359-1365, (2002)

6 H.Bhushan, E.Gupta, Handbook of tribology:materials, coating and surface treatments, Ed. Mc Graw – Hill, (1991).

Study on the Influence of Cutting Parameters on Cutting Forces and Chip Shape of Austempered Ductile Iron (ADI)

Chihong Wang, Xuhong Guo, Wei Wang, Qu Dong

School of Mechanical & Electrical Engineering, Suzhou University

Suzhou, 215021, China

ABSTRACT

Precise dry cutting experiment was done on Austempered Ductile Iron (ADI) with CC650 ceramic cutting tools. A tri-axial piezoelectric dynamometer coupled to a multi-channel charge-amplifier was used to measure and acquire the cutting forces. The effects of cutting parameters on the main cutting force were analyzed separately by means of the Orthogonal test and the Fuzzy logic tool in Matlab Toolboxes, including optmizing the cutting parameters and establishing the empirical formulae of the cutting forces. Further exploration was done on the chip shapes as cutting parameters varies. The results showed that: the depth of cut had the main influence on cutting force followed by feed rate and cutting speed in turn. Cutting force fluctuated by the formation and fracture of the erratic built up edge (BUE) as cutting speed increased, decreasd as BUE grew and increased when it disappeared, almost keeping constant finally. Cutting force increased linearly with the growth of depth of cut and feed rate. The chip shape was mainly determined by cutting speed and feed rate, turnning from crack-like and C-shaped chips to coiled chip with increasing of cutting speed, while varying from coiled chip to C-shaped and rewind chip as feed rate increased.

Keywords: Austempered Ductile Iron (ADI), Cutting parameters, Cutting force, Chip shape

1. INTRODUCTION

Austempered Ductile Iron (ADI) can obtain excellent synthesis mechanical property attributed to its unique microstructure consisting of high carbon austenite and ferrite, due to the application of austempering heat treatment. ADI has became one of the most concerned hot material and technology in 21st century.

A lot of studies have been done to investigate the machining of ADI. Panasiewicz [1] investigated that machine spindle made by ADI has worse mechanical property than that made by common iron. Chang [2] from Advanced Technology Center in Michigan of USA further studied that the bad mechanical property is attribute to the retained austenite in the microstructure, and it changed into the martensite which decreased mechanical property when cutting force exerted on ADI. This point was also approved by four researchers Gundlach, Pashby, Berry and Seah.

Chen Ping and Keishima [3,4] from Japan studied high-speed machining of ADI, discussed the materials and cutting characteristic of a series of cutting tools, and figured out that the cemented carbide tool was not fit for cutting ADI materials, and ceramic and CBN tool was not sensitive to cutting speed.

K.Katuku [5] investigated the wear, cutting forces and chip characteristics when dry turning ASTM Grade 2 austempered ductile iron with PCBN cutting tools under finishing condition (depth of cut: 0.2 mm; feed rate: 0.05 mm/r; cutting speed: 50~80 m/min), pointed out that flank wear and crater wear were the main wear modes within the range of cutting speeds, abrasion wear and thermally activated wear were the main wear mechanisms.

Hongtao Zhang [6] did cutting force experiment with four kinds of polycrystalline cubic boron nitride compact tools on ADI for the purpose of getting the curves of cutting force, friction force and friction coefficient when changing cutting speed (depth of cut: 0.3 mm; feed rate: 0.16 mm/r; cutting speed: 57, 89, 141 m/min). They observed that the friction force and friction coefficient of low-content CBN compact were smaller than those of high-content CBN compact and those of the ceramic binder compact were smaller than those of the metallic binder compact. The PCBN tools was not suit for cutting ADI with massive Al in the binders, since the friction force and friction coefficient unceasingly increased along with the speeding up of the cutting.

This paper will focus on experimental studies of cutting forces and chip shape of ADI with ceramic tool under finishing condition, establish the empirical formulae between cutting parameters and cutting forces, optimize cutting parameters, discuss the influence of cutting parameters on cutting forces and observe the chip shape with the changing of cutting parameters.

2. CUTTING EXPERIMENT

The Preparation and Performance of Experiment Materials

All the experiment material were made in the CSR Qishuyan Locomotive & Rolling Stock Technology Research Institute. The Ductile Iron used in this experiment was prepared in the form of cylindrical barrel with Φ600mm×400mm. The chemical composition(mass percent) of ADI before the austempering treatment is given in Table 1. The heat treatment schedule involved Austenitizing at 890℃ for 120min followed by Austempering in $NaNO_3$ salt bath at 350 ℃ for 60min. Microstructure consisting of ferrite needles and stringer-like retained austenite can be seen with the help of Electron Microscope(EM). The mechanical behavior of ADI material are listed in Table 2.

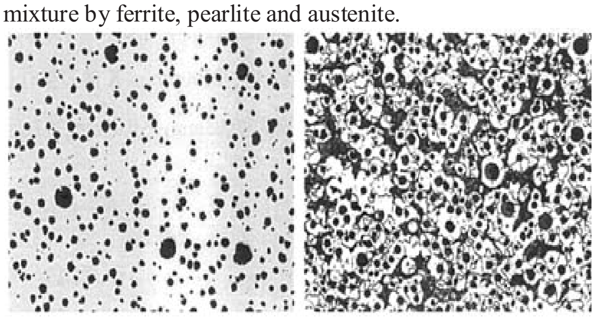
Table 1 chemical composition of ductile iron [*wt*.%]

chemical elements	wt.%
C	3.6
Si	2.85
Mn	≤0.3
P	≤0.1
S	≤0.03
Mg	≥0.02
Re	≥0.02
Al	≤0.05
Ti	≤0.004
Cu	0.6
Mo	0.3
Cr	0.1

Table 2 mechanical properties of ADI

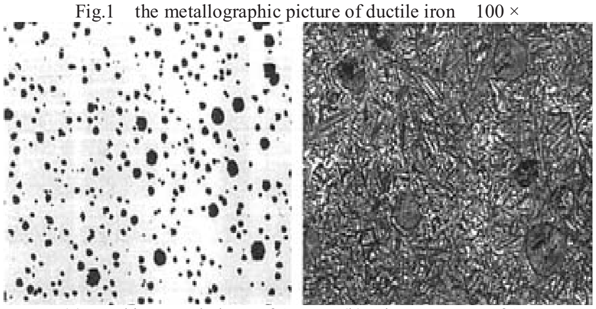
mechanical properties of ADI	
Tensile strength(MPa)	1395
Yield strength(MPa)	900
Elongation(%)	2.75
Compact ductility (J/cm^2)	35
Hardness(HRC)	42

Comparison was done on ductile iron before and after austempering heat treatment by metallographic test. Fig.1 and Fig.2 showed that graphite morphology had hardly any difference, while the microstructure consisting of ferrite needle and stringer-like retained austenite from crude mixture by ferrite, pearlite and austenite.



(a) graphite morphology of cast iron (b) microstructure of cast iron

Fig.1 the metallographic picture of ductile iron 100 ×



(a) graphite morphology of ADI (b) microstructure of ADI

Fig.2 the metallographic picture of austempering ductile iron 100 ×

Experiment Method

This cutting experiment was conducted on a CA6140 lathe. The cutting tool adopted was CC650, a mixed alumina (Al_2O_3)-based ceramic with titanium carbide (TiC). According to ISO standard, the tool was designated as SNGA 120408, square (manufacturer: Sandvik). The cutting tool geometry was listed as follow: rake angle, γ_0=-6°; flank angle, α_o=6°; inclination angle, λ_s=-4°; cutting edge angle, κ_r=75°; minor cutting edge angle, κ_r'=15°; the corner radius r_ε=0.8mm, land width of first face $b_{r1}\times\gamma_{01}$=0.1mm×(-26°). A data acquisition system (Fig.3) made up of Kistler 9257B dynamometer, charge-amplifier, Kistler9403 tool holder and PC was used to measure and acquire the cutting forces.

Multi-factors method was taken in this experiment with cutting parameters selected as below: factor A, cutting speed v_c=81.6, 163.4, 261.4m/min; factor B, feed rate f=0.08, 0.12, 0.16mm/r; factor C, depth of cut a_p=0.05, 0.1, 0.15mm. Measure the cutting forces as much as possible and take the mean value for the purpose of reducing the experimental error. Analysis went on with the help of orthogonal test and the Fuzzy logic tool in Matlab Toolboxes.

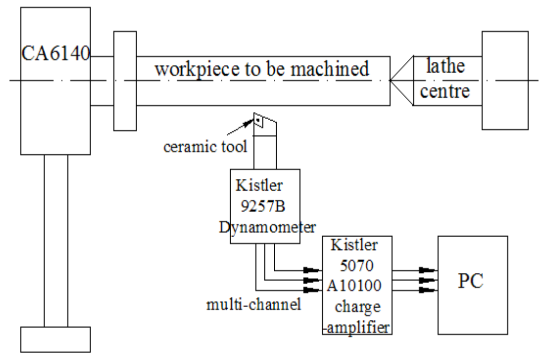


Fig.3 data Acquisition System of cutting forces

3. RESULTS AND DISCUSSION

Empirical formulae of the cutting forces

Cutting parameters also had cooperated effect on cutting forces besides acted separately. An interacted-orthogonal table was used to compute the data measured, with results shown that comparison of *F* value with marginal value revealed the high-significance of cutting speed, feed rate and depth of cut and significance of cross-action of A×B、B×C. The depth of cut had the main influence on cutting force followed by feed rate, cutting speed in turn. The cutting parameters were recommended as below with the consideration of *K* value in such experimental condition, productive efficiency and tool life when cutting ADI under finishing condition: cutting speed, 163.4m/min; feed rate, 0.16mm/r; depth of cut, 0.15mm.

Two types of empirical formulae can be chosen to calculate cutting forces practically, (1) exponential formula; (2) calculating cutting force by unit. Empirical formulae are shown in Table 3, involving three components of forces: main cutting force $F_c(F_z)$; radial thrust force $F_p(F_y)$; axial thrust force $F_f(F_x)$.

Table 3 Empirical formulae of ADI

cutting forces	empirical formulae
main cutting force $F_c(F_z)$	$F_c = 10573.67 a_p^{0.4384} f^{0.4099} v^{-0.3401}$
radial thrust force $F_p(F_y)$	$F_p = 9547.82 a_p^{0.3519} f^{0.2444} v^{-0.2433}$
axial thrust force $F_f(F_x)$	$F_f = 2254.02 a_p^{0.6544} f^{0.1804} v^{-0.6472}$

Table 3 illuminated that cutting forces increased with the growth of depth of cut and feed rate as a result of those positive exponents, decreased with the increase of cutting speed for its negative exponent. The depth of cut had greater influence than feed rate did for depth of cut had bigger exponent, which was in accord with the results of orthogonal test.

Comparing with the results of [7] which cut hardened and tempered steel with cemented carbide tool, it showed the similarity between hardened and tempered steel and ADI that cutting forces increased as depth of cut and feed rate grew with their positive exponents, decreased with increasing of cutting speed owing to the negative exponent.

Influence of cutting speed on cutting force

A single-factor experiment was established to further explore the influence of cutting speed on cutting forces: feed rate f=0.12mm/r; depth of cut a_p=0.2mm; cutting speed v_c=40, 45, 50, 70, 90, 120, 150, 180, 200, 250, 280m/min. Fig.4 is the curve of cutting force F_c against cutting speed v_c with the corresponding chip shape pictures.

Fig.5(a), (b) presented that cutting force nearly linearly increased with growth of depth of cut and feed rate which determined the width and thickness of chip separately. The workpiece (austempering at 350℃ for 1h) consisting of ferrite needle and stringer-like retained austenite had stability in both thermodynamics and dynam. There was no martensite produced when cutting ADI, which had identical views with [9]. Fig.5(c) showed the three-dimensional surface of relationship between feed rate, depth of cut and cutting force as modeled by the fuzzy logic. Cutting force increased at some rate with increasing depth of cut, while at a slower rate with growth of feed rate.

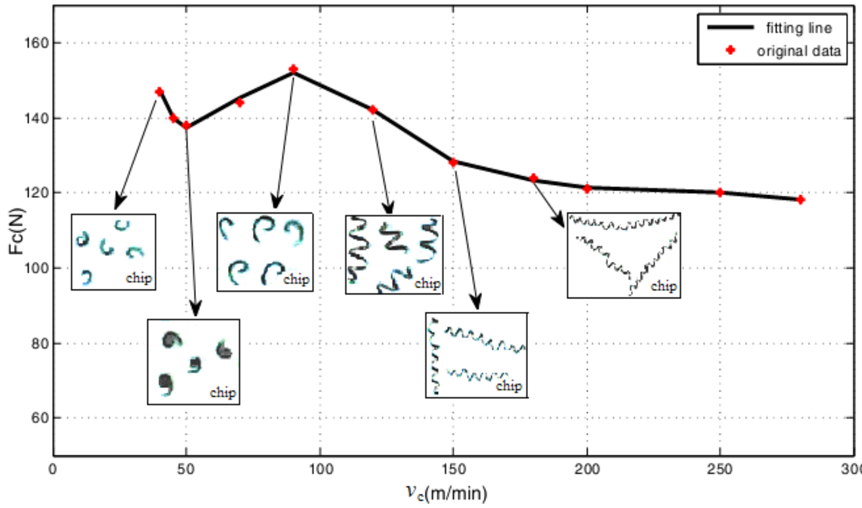


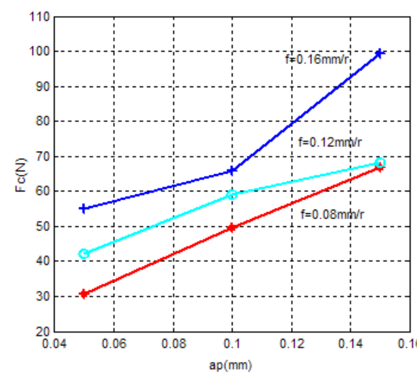
Fig.4 Curve of cutting force Fc against cutting speed vc (f=0.12mm/r, a_p=0.2mm)

Fig.4 indicated that the main cutting force decreased with increasing cutting speed up to a speed of about 50m/min. It was the built up edge (BUE) formed on the cutting edge that makes the tool have bigger real rake angle and lessen the chip deformation. The deforming coefficient would be smaller with growth of cutting speed. The BUE reached the peak till a speed of 50m/min with minimal cutting force about 139N, and the chip shape turned to be C-shaped.

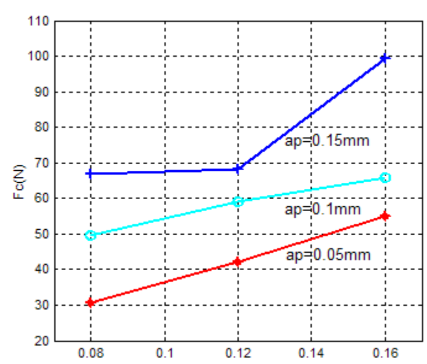
Between speeds of 50 and 90m/min, it increased at some rate for two reasons: (1). The rake angle returned with the fracture of BUE as cutting speed increased, and augmented chip deformation; (2). Interplay between strain hardening and thermal softening brought higher hardness. The main cutting force decreased at a low rate when cutting speed is above 90m/min mainly because of heat in metal cutting. The friction-coefficient between cutting tool and workpiece fell with bigger shear angle Φ and smaller deformation coefficient Λ_h. The chip shape turned to be highly coiled chip from C-shaped chip gradually. Conformity can be concluded after comparing with [8] which milled No. 45 steel with LT55 ceramic tool.

Influence of depth of cut and feed rate on cutting force

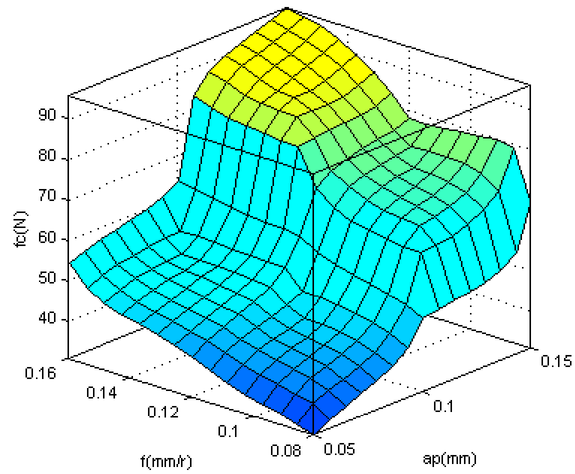
In order to further penetrate the influence of depth of cut and feed rate on cutting force, experiment was conducted: cutting speed v_c=163.4 m/min; depth of cut a_p=0.05, 0.1, 0.2 mm; feed rate f=0.08, 0.12, 0.16mm/r. Fig.5 showed the relationship among depth of cut, feed rate and cutting force F_c.



(a) depth of cut a_p against cutting force F_c



(b) feed rate f against cutting force F_c



(c) Three-dimensional graph describing the relationship among feed rate, depth of cut and cutting force (v_c=163.4m/min)

Fig.5 Influence of depth of cut and feed rate on cutting force

Fig.6 showed the chip shape when cutting ADI with cutting parameters tagged. Chip shape was mainly determined by feed rate and cutting speed (Fig.4), and it turned from highly coiled chip to C-shaped, rewind chips as feed of cut increased, while varied from crack-like and C-shaped chips to coiled chip with the increasing of cutting speed. Depth of cut hardly had any effect on chip.

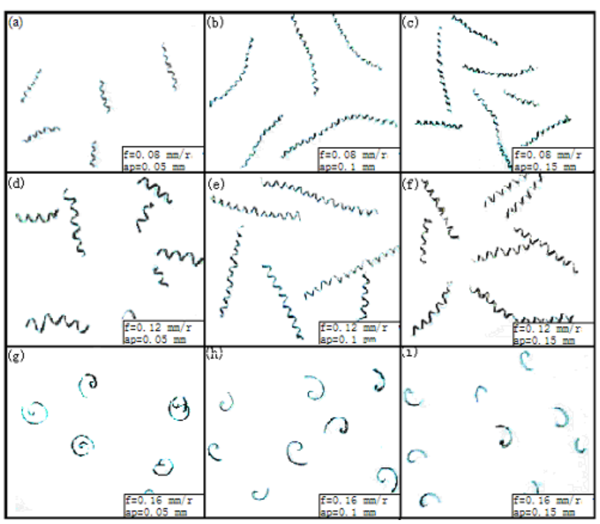


Fig.6 chip shape when cutting ADI (v_c=163.4m/min)

4. CONCLUSIONS

(1) When cutting ADI with CC650 ceramic tool under finishing condition, the depth of cut had the main influence on cutting force followed by the feed of cut, cutting speed in turn with recommended cutting parameters: depth of cutting speed, 163.4m/min; feed rate, 0.16mm/r; depth of cut, 0.15mm.

(2) Cutting force fluctuated by erratic built up edge (BUE) as cutting speed increased, decreased as BUE formed gradually and increased while it fractured, almost keeping constant finally. Cutting force increased nearly linearly with growth of depth of cut and feed rate.

(3) Chip shape was mainly determined by feed rate and cutting speed, and it turned from highly coiled chip to C-shaped, rewind chips with the increasing of feed of cut, while varied from crack-like and C-shaped chips to coiled chip as cutting speed increased.

5. ACKNOWLEDGEMENTS

The authors are grateful to Suzhou Science and Technology Bureau for the support of this research.

6. REFERENCES

[1] Pashby I R, Wallbank J: Ceramic Tool Wear when Machining of Austempered Ductile Iron. Wear, Vol.162-164, No.1, Apr 13, 1993, pp. 22-33.

[2] Goldberg M, Smith G T, Berry J T, Littlefair G: Machinability assessment and surface integrity characteristics of austempered ductile iron (ADI) using ultra-hard cutting tools. 3rd International Machining and Grinding Conference, Cincinnati, Ohio, USA, October 4-7, 1999[C]. 0001.

[3] Cakir, M. Cemal, Bayram, Ali, Isik, Yahya, Salar, Baris: The effects of austempering temperature and time onto the machinability of austempered ductile iron. Materials Science and Engineering, Vol.407, No.1-2, Oct 25, 2005, pp. 147-153.

[4] Seker, Ulvi, Hsirci, Hasan: Evaluation of machinability of austempered ductile irons in terms of cutting forces and surface quality. Journal of Materials Processing Technology, Vol.173, No.3, Apr 20, 2006, pp. 260-268.

[5] K. Katuku, A. Koursaris, I. Sigalas: Wear,cutting forces and chip characteristics when dry turning ASTM Grade 2 austempered ductile iron with PcBN cutting tools under finishing conditions. Journal of Materials Processing Technology, Vol.209, No.5, Mar1, 2009, pp. 2412-2420.

[6] Zhang Hong-tao, Li Hai-bo, Dong Hai, Li Man: Cutting Properties in Cutting Austempered Ductile Iron by PCBN Compact Tools in chinese. Materials for Mechanical Engineering, Vol.32, No.8, Aug 2008, pp. 43-46.

[7] Xu Jin, Ye Bang-yan: Study on High Speed Cutting Hardened Steel with the Coated Insert (CN35) in chinese[D]. South China University of Technology, Guangzhou. Oct 1, 2003.

[8] Ai Xing. High Speed Machining Technology in Chinese [M]: National Defense Industry Press, Bei Jing.2003 .

[9] Deng Hongyun, Wang Chunjing, Zhang Zhou: The Production and Application examples of Austempering Ductile Iron in chinese. Chemical Industry Press, Bei Jing. 2009.

Analysis and Research of Influence of Advanced Ignition Angle on Engine Emission Based on Fuel Quality

Li Jun
School of Mechantronics and Automotive Engineering, Chongqing Jiaotong University
Chongqing, China

and

Zhang Shiyi
School of Marine, Chongqing Jiaotong University
Chongqing, China

and

Yang Lizhong
China Marine Bunker (PetroChina) Co. Ltd QinHuangDao Branch
Hebei, China

ABSTRACT

Because of difference of fuel component and quality in China, this paper discusses and analyzes the results of fuel quality and components and experiment in China. By using the ordinary fuel from HuaZhong and XiNan district, the engine CO_2、HC、NOx emissions are tested and the change of engine emission is studied when adjusting the advanced ignition angle suitably. Engine can operate smoothly and HC and NOx emission can be reduced when the advanced ignition angle is decreased properly against to the number of aromatic hydrocarbon of finished fuel which is higher partially such as in HuaZhong district.

Keywords: Fuel Quality; Vehicle Engine; Engine Emission; Advanced Ignition Angle

1. INTRODUCTION

The vehicle emission is gradually become one of the main air pollution in city. The components and amount of engine emissions are influenced by the fuel quality better or bad directly. Because of the differences of the crude oil components in refinery in China, the finished fuel is different from each other in every district. This paper researchs and studies the vehicle engine emission based on different fuel quality in China. And it is discussed to control and reduce the vehicle emissions by adjusting and changing the some engine parameter against to difference fuel quality and components[1-3,5]. According to investigating, the output of China Petrochemical Corporation and China National Petroleum Corporation is about 90% all of the finished fuel in China[2,3]. And because of difference of the refining technology and epurating process, there are some differences in fuel quality and components in different districts in China. Since analysis of the gasoline and fuel quality from some refinery or gas station in sixteen cities of 7 districts in China, we know that RON(research octane number) is up to the demand basically, but MON(motor octane number) about 20% in $93^{\#}$ is lower. The resistance violent index of fuel is not unsatisfied the Chinese Standard. And the alkene of finished fuel in China is rather higher than America and Japan about 30%~50% commonly and the number of sulfur in China is less and aromatic hydrocarbon and benzene hydrocarbonis also are less than the importing fuel. So it is important to research and control the engine emission based on different fuel quality and component and it has been become a key problem how to adjust some engine performance and structure parameters and control combustion process so as to reducing engine emission.

2. MATHEMATICAL MODELS AND ANALYSIS

According to the characters and process of engine combustion, the И.И.В-Иδе indicates that the consumed percentage x of fuel burning along with time varying is [4]:

$$x = 1 - e^{-f(t)} \tag{1}$$

$$f(t) = n\int^{1}\rho dt \tag{2}$$

$$\rho = k_0 t^{m0} \tag{3}$$

Where: n is proportion coefficient between molecule number of efficient reactive substance and molecule number of being in for reactive substance, ρ is relative density, m is index of fuel quality, k_0 is proportion coefficient.

Then put the equation (3) in (2) and integral transform :

$$f(t) = nk_0\int_0^1 t^m dt = \frac{k}{m+1}t^{m+1} \tag{4}$$

Where: $k = nk_0$. And put equation (4) into (1). Then:

$$x = 1 - \exp\left(-\frac{k}{m+1}t^{m+1}\right) \tag{5}$$

Commonlly the m=1.0-3.0 for gasoline engine[4]. According to experience we know that fuel don't been burned completely. In generally if fuel above 99% is burned, it can been considered as complete burning. So the consumed percentage x of fuel burning along with time varying is:

$$x = 1 - \exp\left[-6.908\left(\frac{t}{t_z}\right)^{m+1}\right] \tag{6}$$

Where: the m is index of fuel quality; the t_z is the presumptive continue time about burning process; the t is burning time; the $\frac{t}{t_z}$ is relative burning time.

However, the relative burning time $\frac{t}{t_z}$ can be substituted with the relative ratio of crank angle $\frac{\varphi_x}{\varphi_z}$.

The φ_x is a crank angle about burning continue time and $\varphi_x = \varphi - \varphi_0$; the φ is a crank angle; the φ_0 is advanced ignition angle; the φ_z is a crank angle about all burning process continue time.

Then:

$$x = 1 - \exp\left[-6.908\left(\frac{\varphi - \varphi_0}{\varphi_z}\right)^{m+1}\right] \tag{7}$$

The quasi-experience equation about combustion velocity $\frac{d_x}{d_\varphi}$ of gasoline engine is:

$$\frac{dx}{d\varphi} = 6.908\frac{m+1}{\varphi_z}\left(\frac{\varphi - \varphi_0}{\varphi_z}\right)^m$$

$$\times \exp\left[-6.908\left(\frac{\varphi - \varphi_0}{\varphi_z}\right)^{m+1}\right] \tag{8}$$

Therefore, we can understand that the combustion velocity of gasoline engine is a function about fuel quality and advanced ignition angle and burning continue time or relative crank angle of burning continue time.

According to above the equations, we can know that the combustion velocity of gasoline engine can result in the burning process directly. The fuel quality can effect on combustion velocity. Then the burning process can influence on engine emission and the number of pollutants[4,5]. So the fuel quality may be able to influence on engine emission. Therefore, we adjust and change some running parameters of gasoline engine suitably and then the burning velocity of engine can be changed and so that the engine emission can be reduced and controlled [6-12].

3. EXPERIMENT

In order to researching the influence of engine emission based on different fuel quality, we have completed the experiments about engine combustion and emission by using the fuel from XiNan district and HuaZhong district in China. Analysis and discussion about influence of fuel quality on gasoline engine emission has been accomplished by changing and adjusting the advanced ignition angle suitably. We used the JL465Q series engine. Its engine displacement is 1.0 cc and compress ratio is 9.5:1. Exhaust analyzer is MEXA-324J. Exhaust testing instrument is BOSCH ETT855 and air-fuel ratio instrument is MEXA-110λ. The fuel is 93# gasoline from XiNan and HuaZhong districts. In the experiment, the engine rotation rate is from 850 to 3500 r/min (rated speed is 2200r/min). The intake temperature of engine is about 25-35℃ steadily. The cooling water temperature of engien is controlled in 85℃ and lubricant oil temperature is controlled in 90℃. The engine load is around 85%-100%.The experiment is completed within two group different parameter conditions against to the number of sulfur; alkene; aromatic hydrocarbon and benzene hydrocarbon of gasoline from the two different districts. We research the engine emission variety based on different fuel quality by adjusting some parameters of engine suitably. And we analyze and discess the change of exhaust number and the method of improving emission.

4. RESULTS AND ANALYSIS

In order to researching the influence of advanced ignition angle on engine emission based on fuel quality, we adjust advanced ignition angle suitably from φ_0=10° to φ_0=15° in these experiments. The exhaust and emission change is shown in Figure1-3. Its number of aromatic hydrocarbon in HuaZhong district fuel is higher than in XiNan although its number of alkene in fuel is near to XiNan's. And it is approximately the highest in China. For the Fig.1, we can know that the CO_2 emission from burning fuel in HuaZhong is higher than in XiNan clearly, because aromatic hydrocarbon contains some benzene hydrocarbon in fuel from HuaZhong district and the number of aromatic hydrocarbon is higher than in XiNan district. When the advanced ignition angle is increased, the delay time of engine fire in combustion process is brought forward and the CO_2 emission can be decreased partly. It is visible greatly in engine operation at 1500-3000r/min to the CO_2 being reduced. Its CO_2 emission number is near to the emission number of Xinan district in which aromatic hydrocarbon is lower. Therefore, the CO_2 emission can be reduced when advanced ignition angle is increased suitably against to some fuel contained higher aromatic hydrocarbon. Shown from Fig.2-3, we can know that the HC is increased partially and NO_x is partly decreased, when advanced ignition angle is augmented. There is a little discrepancy each other in the two districts. But the P_{max} and rate of pressure rise is increased notably in acute burning stage, when advanced ignition angle is augmented and so that the engine works badly and engine operates no-smoothly.

Because of higher octane number of aromatic hydrocarbon, the power consistency of fuel is higher. When the fuel contains higher aromatic hydrocarbon, it results gasoline engine easily in working badly and no-smoothly and in occurring operating vibration and noise acutely. For higher number of aromatic hydrocarbon fuel, when advanced ignition angle φ_0 is increased beyond to some value, the NO_x emission of engine is decreased no-notably and changed slowly. But it becomes great bad and no-smooth prominently to engine working and running. It is shown in Fig.4. Engine working is influenced greatly. Hence, in order to decreasing CO_2 and NO_x emission, we need increase advanced ignition angle suitably, but not exceed to some number. Otherwise, it is not prominent to decreasing of CO_2 and NO_x emission and then the HC emission is increased notably and engine working is bad and no-smooth and engine running vibration and noise is serious fearfully.

5. CONCLUSIONS

Because of difference of fuel component and quality in China, the engine emission is different each other. For higher number of aromatic hydrocarbon of fuel, the CO_2 and NO_x emission can be decreased when advanced ignition angle is increased suitably. But when the angle φ_0 exceeds beyond to some number, it is not marked to reduce CO_2 and NO_x emission. But the HC

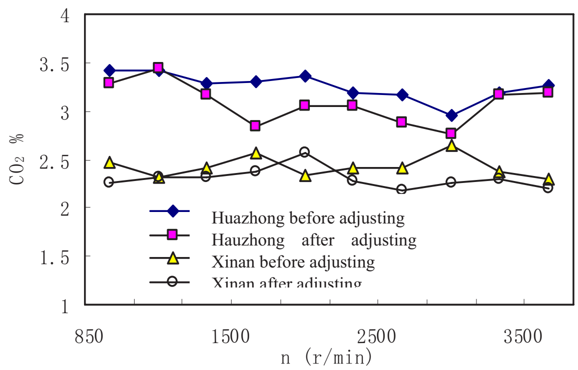


Fig.1 relation of the advanced ignition angle with CO_2 emission

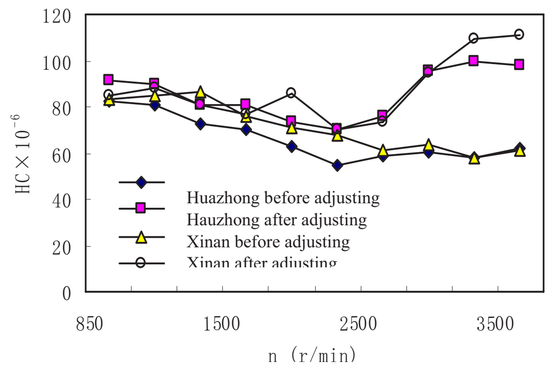


Fig.2 relation of the advanced ignition angle with HC emission

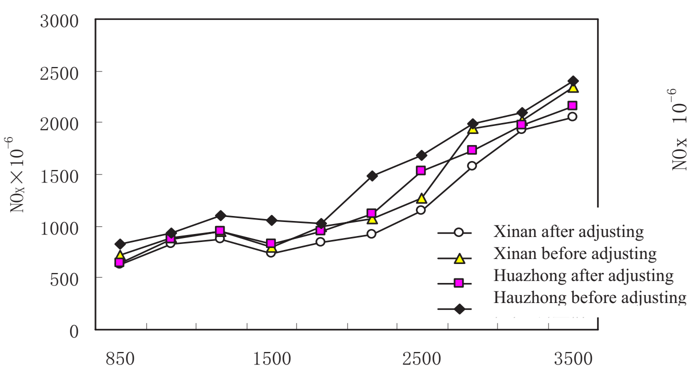
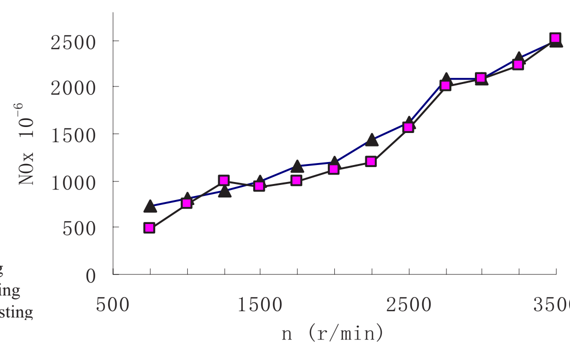


Fig.3 relation of the advanced ignition angle with NO_X emission



▲— $\varphi_0=15°$, □— $\varphi_0=23°$

Fig.4 relation of the advanced ignition angle φ_0 with NO_X emission using the fuel from HuaZhong district

emission is increased observably and engine working is bad and no-smooth. And engine running vibration and noise gradually becomes more fury. So we adjust the advanced ignition angle φ0 suitably and not beyond the permited number.

6. ACKNOWLEDGEMENT

The work was supported by Education Natural Science Foundation Project of Chongqing EDUC(KJ090408), China and by Foundation Project of the Key Laboratory of Chongqing Communication Engineering (2008CQJY002) and by Education and Teaching Reform Project of Chongqing CSTC, China (No.0903070)

7. REFERNCES

[1] Wang Jieqing. Analysis of National Pureness Gasoline Quality and Presenting Actuality. Proceedings of the Tenth the Lubricating Oil and Fuel Committee Conference of Automotive Engineers Society of China[M]. 2002

[2] Zhang Yongguang. Countermeasure of Fuel Components In China Against the Protecting the Circumstance Requirements. Proceedings of

Pureness Fuel and Protecting the Circumstance Proseminar [M]. Beijing:1999

[3] Li Jian. Research on Emission Performance of the Engine Based on Fuel Character. Master Degree Dissertation of Chongqing University[D]，2005

[4] Zhang Zhipei. Principle of Automotive Engine[M]. China Communication Press, Beijing:2003,03

[5] Li Jun, Zhang Shiyi. Investigation and Research in Influence of Excess Air Coefficient on Engine Emission Based on the Fuel Quality[J]. Automotive Engineering, 2009.31(3)：201-204

[6] Li Jian, Qin Datong, Han Weijian. Investigation on Effects of Fuel Properties on Vehicle Emissions [J]. Journal of Chongqing University (Natural Science Edition), 2005,28(7):4~8

[7] Zhang Lijun, Yao Chuanfeng. Research of Influence of Fuel Quality on Particulate and Polycyclic Aromatic Hydrocarbon Emission from Diesel Engines [J]. Chinese Internal Combustion Engine Engineering, 2000(4):35~38

[8] Guo Hejun, Fang Maodong. Research on Effects of Gasoline Qualities on Vehicle Exhaust Emissions [J]. Journal of WuHan University of Technology, 2005,27(9):64~66

[9] Lin Huaiwei, Chen Xiaolong. The influence of Gasoline Quality upon Auto Emissions [J]. World Automobile, 2000(4):18~20

[10] Cheng Yong, Fu Tieqiang. Status Quo and Analysis of Evaporative Emission from Domestic Light-duty Gasoline Vehicle, 2002,24(3):182~186

[11] Kihyung Lee; Changhee Lee. An Experimental Study of the Extent of the Operating Region and Emission Characters of Stratified Combustion Using the Controlled Autoignition Method. Energy & Fuels 2006(20):1862-1869.

[12] C.H.Lee; K. H. Lee. An Experimental Study on the Combustion and Emission Characters of a Stratified Charge Compression Ignitin (SCCI) Engine. Energy & Fuels 2007(21) 1901-190

SWEEP SINE WAVE BASED SHM FOR SHORT COMPOSITE TUBES

Ming LI Gurjiwan SINGH Gurjashan SINGH Alberto GARCIA Ibrahim TANSEL
Mechanical and Materials Engineering, Florida International University
Miami, FL33174, USA

Mustafa DEMETGUL
Technical Education Faculty, Marmara University
Istanbul, Turkey

and

Aylin YENILMEZ
Mechanical Engineering, Istanbul Technical University
Istanbul, Turkey

ABSTRACT

Composite tubes have been widely used for building the structures of the unmanned aerial vehicles (UAV). Monitoring the integrity of their structures would improve the safety and reliability of small UAVs. In this study, sweep sine wave was used for excitation of the carbon fiber tubes of a small four rotor helicopter. The piezoelectric elements attached to the tube acted as actuator and sensor. The actuator created the Lamb waves on the surface of the structure. The experimental data was collected by using a free tube and after a clamp was attached to the different locations of the tube. The S transforms of the sensory signals were used for analysis. Backpropagation type artificial neural networks were used for classification.

Keywords: Structure Health Monitoring, Sweep Sine Wave, S Transformation

1. INTRODUCTION

Composite materials are an excellent choice for building many structural parts of modern aircraft and unmanned aerial vehicles (UAV). Their light weight, high strength and corrosion resistance enable engineers to achieve better performance, reduce maintenance cost and increase the life of air vehicles. Installation of the Structural Health Monitoring (SHM) systems at the composite structures of the unmanned aerial vehicles are expected to improve their reliability while performing various tasks at distant locations.

Most of the SHM systems use active or passive sensors for data collection, and determine the health condition of structures by processing the signals and interpreting the results. Piezoelectric elements are one of the most common actuators/sensors since they are cheap and create vibrations with electricity or vice versa [1]. In this study, the generation of Lamb waves with sweep sine wave through the piezoelectric actuator was proposed to diagnose the condition of a short composite tube which could not be easily diagnosed with single frequency bursts, and to be able to work effectively at low sampling rates.

Lamb waves are generated on the surface with proper excitation. Their propagation changes even with slight defects on the surface. Therefore, many recent SHM methods evaluate the propagation of the Lamb wave by using time-frequency domain methods [2,3,4]. The short time Fourier transformation(STFT), wavelet transformation(WT), and S transformation(ST) are examples of algorithm used with this purpose.

In this study, Lamb waves were generated using sweep sine wave. S transformation was used to extract and encode the features indicating the condition of the structure. Backpropagation type neural network was used for the classification of the cases. The proposed approach was implemented to develop a minimal SHM system for a small four rotor helicopter, DraganFlyer V ti Pro. The validation experiment was carried out on the carbon fiber tubes of the main frame of the helicopter. Not to destroy the tubes, the signals were collected by using free tubes and tubes with a clamp attached to different locations.

The analysis of the experimental data showed that the characteristics of the most significant peaks of the frequency responses varied when the clamps were attached to different locations and tightened at various levels. The results also indicate the proposed method was a convenient approach for detection of the structural and assembly problems of tubes.

2. THEORETICAL BACKGROUND

Fast Fourier Transformation (FFT) is an efficient algorithm for transformation from time to frequency domain. The frequency domain characteristics of the signals can be easily extracted and analyzed after FFT [5]. The frequency-domain representation does not have information about the changes in the time domain. In many applications, such as speech recognition and SHM, the time-frequency representation (TFR) is desirable [6]. The short time Fourier transformation (STFT) [7], wavelet transformation (WT) [8] and S transformation (ST) [9] may be used to study the time-varying characteristics of the spectral information.

The short time Fourier transformation includes time domain information by calculating Fourier transformation of short periods of time defined by a window function. As the window slides along the time axis, the Fourier transforms of the entire time series is obtained piece by piece. Assembly of these pieces represents the characteristics of the signal. The STFT can be written as:

$$\mathrm{STFT}(\tau, f) = \int_{-\infty}^{+\infty} h(t) w(\tau - t) \exp(-i2\pi f t) dt \quad (1)$$

Where $h(t)$ is the time series, w is the window function and τ is the position of the window along the time axis. The optimization of the window width determines the compromise of the frequency and time resolution of the analysis. A wide window gives relatively better frequency resolution, while a narrow window gives better time resolution [3].

Since the fixed window size limits both the frequency and time resolution, WT, ST and other Fourier derived methods have been developed to adapt window width in function of the frequency to achieve optimal resolution. These multiresolution algorithms can achieve good time resolution at high frequencies and good frequency resolution at low frequencies [9].

S transformation was introduced by Stockwell as an extension of continuous wavelet transformation (CWT). It is defined as a "phase correction" of CWT [6],

$$S(\tau, f) = \exp(i2\pi f \tau) W(\tau, d) \quad (2)$$

Where,

$$W(\tau, d) = \int_{-\infty}^{+\infty} h(t) w(t - \tau, d) dt \quad (3)$$

is the CWT of a function h(t). Thus, the S transform is defined as [6]:

$$S(\tau, F) = \int_{-\infty}^{+\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(\tau - t)^2 f^2}{2}} e^{-i2\pi f t} dt \quad (4)$$

S transform is based on a moving and scalable localizing Gausssian window, and it can provide frequency-dependent resolution while at the same time maintain a direct relationship with the Fourier Spectrum [6].

3. EXPERIMENT SETUP

The experiment was carried out on the carbon fiber tubes of the main frame of a small four rotor helicopter, DraganFlyer V ti Pro. Fig.1 shows one of the tubes. The tube is 190mm long; its outer diameter is 5mm and its inner diameter is 3mm. Two identical piezoelectric elements were bonded on the surface of the opposite ends of the composite tubes. The piezoelectric elements have the diameter of 12mm and thickness of 0.6mm. One piezoelectric element was used as an actuator and the other one was used as a sensor. The actuator was excited with a sweep sine wave. The sensor collected the information from the arriving surface waves propagated along the surface of the tube.

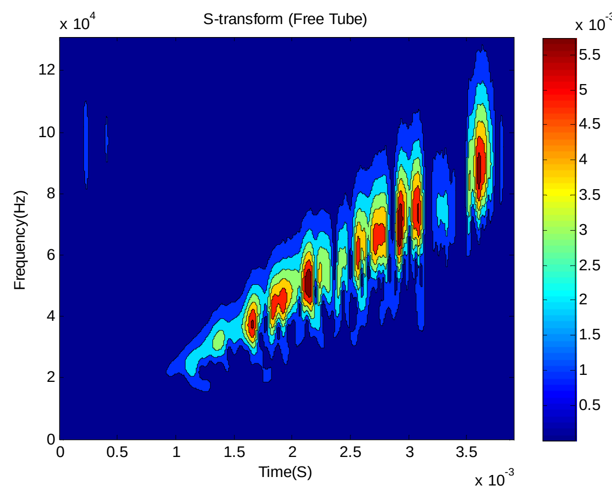
The sweep sine wave was generated by the spectrum analyzer (Stanford Research Systems Model SR780 2 ch Network Analyzer) in Fig.2. The frequency of the generated sweep sine wave signal was between 1Hz and 102.4K Hz. In the experiments, either the tube was left alone or a clamp was attached to study the influence of an obstacle which disturbs the surface waves. The frequency content of the acquired signal at the sensor was analyzed by the same spectrum analyzer. The same experimental procedure was repeated when the clamp was attached 40mm, 80mm and 120mm away from the actuator.

Fig.1 Carbon Fiber Tube with Piezoelectric Elements Attached

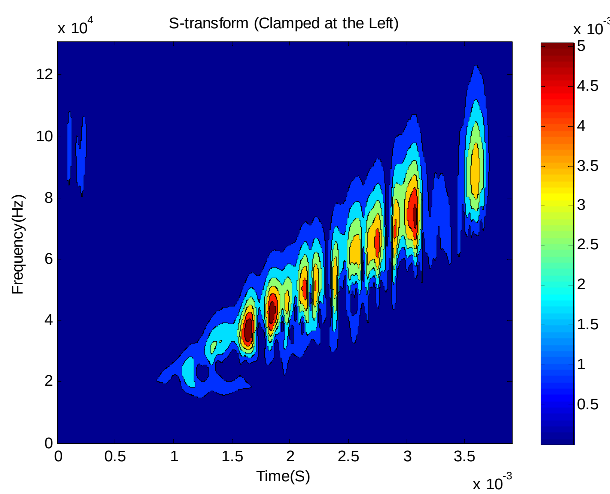
Fig.2 Experiment Setup

4. RESULTS

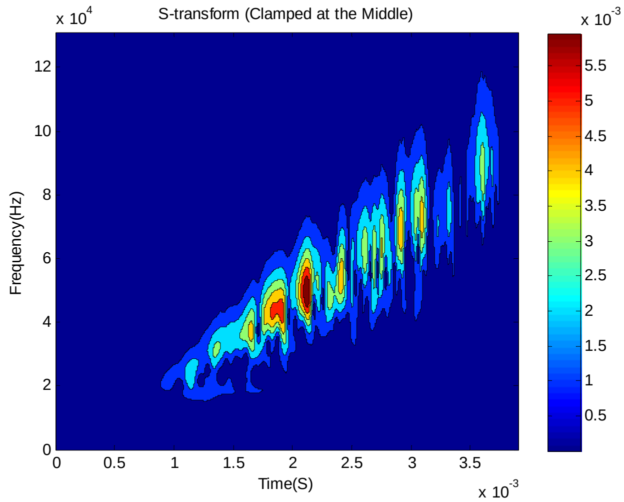
The S transform of the sampled signals at the sensor was calculated. Fig.3 shows the S transforms of the free tube, and the same tube when the clamp was placed at three different locations. In Fig.3, the time-frequency characteristics of the sensory signals of the tube at the above conditions were illustrated by contours. The contours represent the certain high amplitude levels. Fig.3 indicated that the contours of the free tube and the same tube with a clamp had very different characteristics. Free tube had the higher spikes since the clamps obstructed to the propagation of the surface waves and reduced their energies.



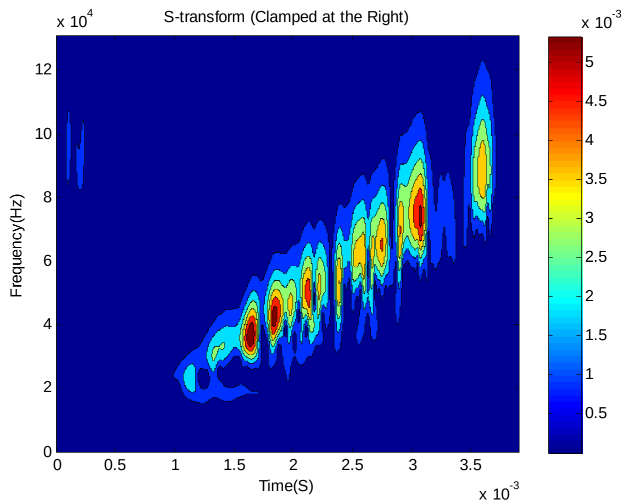
a) Free Tube



b) Clamped at the Left



c) Clamped at the Middle



d) Clamped at the Right

Fig.3 S Transform of the Sensory Signals

Since the excitation signal was a sweep sine wave, the peaks located on a diagonal line correlated the increasing dominant frequency with time. The level of the most significant peaks of the S transform indicated the tube condition. Fig.4 compares the characteristics of the peaks of the free tube and the tube clamped at the right.

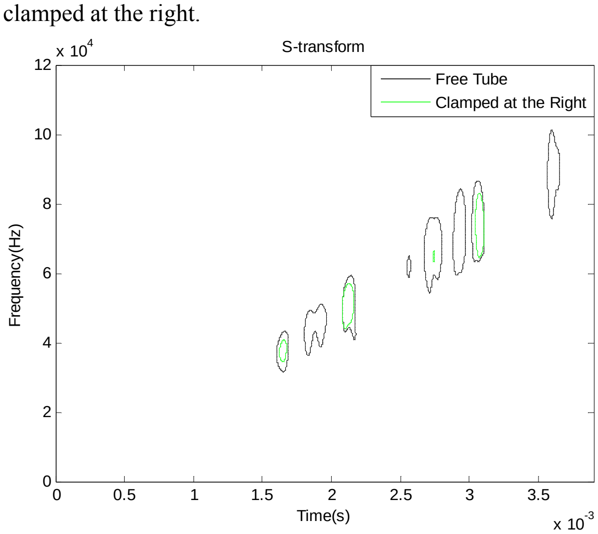


Fig.4 Contour of Peaks

The time, frequency and amplitude of the most significant peaks were isolated and utilized for estimation of the condition of tubes. Fig.5 illustrates the peaks along the time axis and Fig.6 illustrates the peaks along the frequency axis.

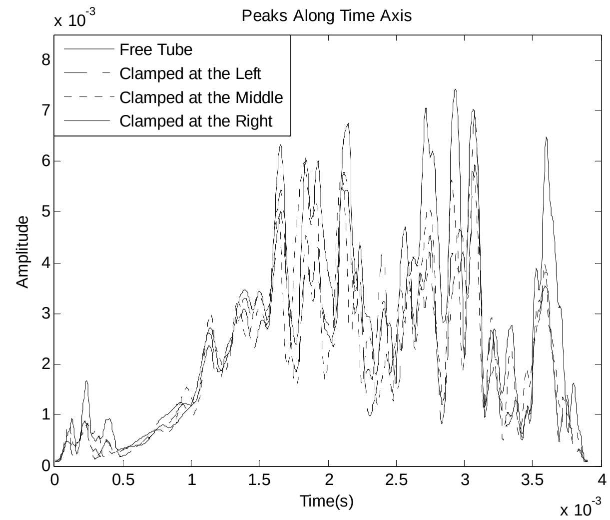


Fig.5 Maximum Amplitude along Time Axis

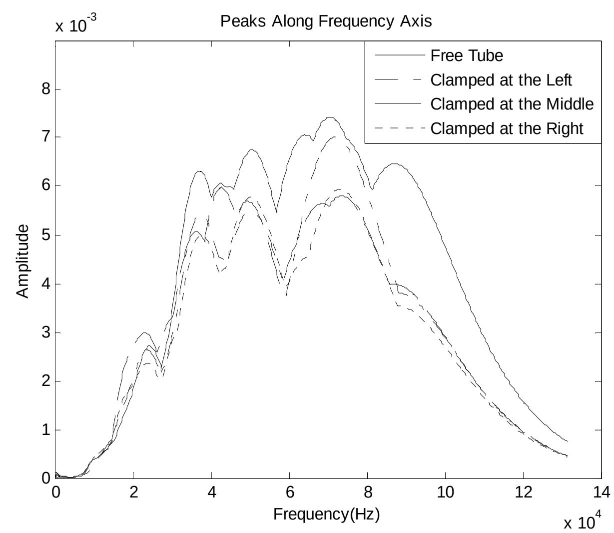


Fig.6 Maximum Amplitude along Frequency Axis

The encoded information of time, frequency and amplitude extracted from peaks of S transforms of the dynamic responses under different conditions was used to train neural networks. The inputs and the outputs of the backpropagation type neural network model are illustrated in Fig.7. The neural network had four inputs, four hidden nodes and one output. The output was an index which represented the condition of the tube. It was a digital value when we wanted to separate the free and clamped tubes. It had an analog value when we estimated the location of the clamp. The inputs were the amplitudes of the frequency response at four critical frequencies identified from the S transform plots. These frequencies were 37, 50, 70 and 87 KHz. The amplitudes of the signal at these four frequencies were presented to the backpropagation type neural network. The experiments were repeated with free tube and the same tube with a clamp at various locations. 16 of these cases were used for training while the other 16 cases were used for testing.

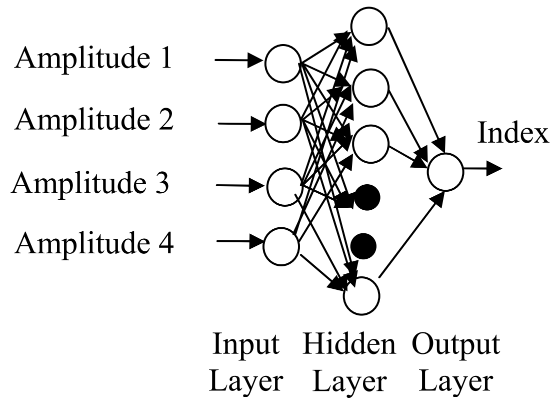


Fig. 7 Backpropagation Neural Network Model

The trainings and tests were carried out with two different purposes. First, the diagnostic method aimed to separate the free and clamped tubes. The index of the training data extracted from the free tube's dynamic response was labeled as 0, and all others were labeled as 1. The neural network converged quickly during the training, and the estimation accuracy of the diagnostic system was 88%. Out of the 16 test cases, only 2 cases were identified wrongly. Secondly, the location of the clamp was estimated by the neural network. The training and test process was the same as the previous one except that the index of the training data was labeled as 2, 3 or 4 according to the location of the clamps (2=left, 3= middle and right=4). After the training process, the neural network estimated health index of the tube with 38% accuracy at the test cases. Out of the 16 test cases, only 6 were classified correctly. Based on the results, the proposed method may be confidently used for determination of the existence of a surface defects or not on a tube. If the clamps used for holding components on a tube loosen up, it would be possible to detect the problem. The confidence level would be much lower when the location of the obstruction is estimated.

5. CONCLUSIONS

In this study, feasibility of using the sweep sine wave was tested to detect defects and evaluate if the clamps were tightly attached to tubes. For evaluation of the defect detection, a clamp was attached to a tube. This approach allowed us to repeat the experiments hundreds of times at different clamp locations and tightness without damaging the tube. The study indicated that sweep sine wave excitations may be used to detect defects disturbing propagation of Lamb waves or attached clamps. The characteristics of S transforms of the propagated waves to the sensor indicated the obstructions which influenced the propagation of surface waves. These obstacles could be either defects or clamps. The encoded parameters of the signal were classified by using the backpropagation type neural networks. The neural network detected the existence of obstacle accurately. However, the location estimation was slightly compromised.

6. ACKNOWLEDGEMENT

The authors would like to thank the Graduate School of Florida International University for providing the Dissertation Year Fellowship and Teaching Assistant Scholarships.

7. REFERENCES

[1] D. P. Garg, M. A. Zikry and G. L. Anderson, "Current and Potential Future Research Activities in Adaptive Structures: an ARO Perspective", Smart Materials and Structures, No.10, 2001, pp. 610-623

[2] Y. Lu, X. Wang, J. Tang and Y. Ding, "Damage Detection Using Piezoelectric Transducers and the Lamb Wave Approach: II. Robust and Quantitative Decision Making", Smart Materials and Structures, No. 7, 2008, pp.1-13.

[3] M. Niethammer, L. J. Jacobs, J. M. Qu and J. Jarzynski, "Time-frequency Representations of Lamb Waves", Acoustical Society of America, Vol. 109, No.5, 2001, pp. 1841-1847.

[4] Kessler S.S., Spearing, S.M. and C. Soutis. "Damage Detection in Composite Materials using Lamb Wave Methods." Proceedings of the American Society for Composites, 9-12 September 2001, Blacksburg, VA.

[5] E. P. Carden and P. Fanning, "Vibration Based Condition Monitoring: A Review", Structure Health Monitoring, Vol.3, No.4, 2004, pp.355-377.

[6] R.G. Stockwell, L. Mansinha, and R. P. Lowe, "Localization of the Complex Spectrum:The S Transorm", IEEE Transactions on Signal Processing, Vol.44, No. 4, 1996, pp.998-1001.

[7] S. Hurlebaus, M. Niethammer, L. J. Jacobs and C. Valle, "Automated Methodology to Locate Notches with Lamb Waves", Acoustical Society of America, Vol. 2, No. 4, 2001, pp. 97-102.

[8] S. Legendre, D. Massicotte, J. Goyette and T. K. Bose, "Wavelet-Transform-Based Method of Analysis for Lamb-Wave Ultrasonic NDE Signals", IEEE Transactions on Instrumentation and Measurement, Vol. 49, No. 3, 2000, pp.524-530.

[9] X. Wang, and I. N. Tansel, "Modeling the Propagation of Lamb Waves using a Genetic Algorithm and S-transformation", Structural Health Monitoring, Vol.6, No.1, 2007, pp. 23-37.

Electromechanical Behavior of CNT Nanocomposites

Yves Ngabonziza*
Math, Engineering and Computer Science Department
LAGCC of the City University of New York
Long Island City, NY 11101
*e-mail: yngabonziza@lagcc.cuny.edu

Jackie Li
Mechanical Engineering Department
The City College of the City University of New York
New York, NY 10031
e-mail: j.li@ccny.cuny.edu

ABSTRACT

Electrical resistance responses of multi-walled carbon nanotubes (MWCNT) reinforced polypropylene (PP) nanocomposites under mechanical tensile loading are studied in this paper. A standard tensile test was conducted while the electrical resistance was measured using 2-probe method. From our previous works on the CNT/PP nanocomposites, the percolation threshold of electrical conductivity is around 3.8 wt% of CNT. The influence of this percolation threshold on the electrical resistance upon mechanical loading was investigated. The results will be discussed and compared.

INTRODUCTION

It is well known that polymers are naturally insulators; combined with CNTs, they become conductive and make them even more attractive due to this additional property on top of their other interesting properties such as light weight, high strength, machinability, and optical properties, among the others. In our previous work [1], we have shown that PP-MWCNT composites produced using injection molding had a percolation threshold of electrical conductivity around 3.8 wt% of CNT.

During the electrical resistance measurement, two techniques can be used. The first, called the two-point probe technique, is an electrical potential method, which is made of two single-terminal electrodes attached to the surface of the conductive structure. A DC *or* AC source current is then applied through the two electrodes and the resulting voltage across the same electrodes is measured. The electrical resistance between these two electrodes is then calculated based on the Ohm's law. The second, called the four-point probe technique, is an electrical impedance method, which uses separate pairs of electrodes for current-sourcing and voltage-sensing. That is, the outer and inner terminals of the electrodes are used as current and voltage contacts, respectively. The key advantage of the four-point probe technique over the traditional two-point probe technique is that by separating current-sourcing and voltage-sensing terminals, the four-point probe technique eliminates the impedance contributions of the wiring and contact resistances. On the other hand, when space is limited, such as the 1-D strip specimens commonly used in laboratory tests, the two-point probe technique can be applied more conveniently than the four-point probe technique.

Several researchers have been interested in the use of piezoresistivity and electrical conductivity for sensing purposes [2-4]. Most of those studies were carried out on carbon fiber composites and proved to be efficient. Due to carbon fiber's inherent nature of piezoresistivity and electrical-conductivity ($\sigma = 5\times10^4$ $S{\cdot}m^{-1}$), this technique has been studied by many researchers for self-sensing of carbon-based composites. Among the others, Chung and her associates [5-15] have conducted extensive research in the area of self-sensing/self-monitoring/self-diagnosing of carbon based system. Pham et al. [16] developed carbon nanotube polymer composite films that can be used as strain sensors with tailored sensitivity. The films were fabricated by either melt processing or solution casting of poly (methyl methacrylate) (PMMA) with MWNT.

However, little work has examined the effect of mechanical loading on electrical conductivity of polymer composites. For the PP nanocomposites used in this study, melt-mixed CNT-PP concentrate was diluted with neat PP in the injection molding process. This letdown of pelletized masterbatches is a very common practice for handling fine particles during injection molding.

This paper contains an investigation of the influence of mechanical loading on the electrical conductivity of the CNT-PP nanocomposites. As commonly known, polymers are naturally insulators; combined with CNTs, they become conductive and make them even more attractive due to this additional property on top of their other interesting properties such as light weight, high strength, machinability, and optical properties, among the others. In addition to that, the CNT-PP nanocomposites would be a good candidate for strain sensing capabilities.

In this paper, the effect of mechanical loading on electrical conductivity will be investigated.

EXPERIMENT

To fabricate the nanocomposites, a CNT-PP masterbatch (concentrate) was dispersed in polypropylene base material using a 55-ton reciprocating screw injection molding machine (Cincinnati Milacron-Fanuc, Model: Robo 55R-22) The polypropylene (BP Amoco's Acclear 8449) was a random copolymer with a melt index of 12g/10min. The MWNT masterbatch was obtained from Hyperion Catalysis (grade: MB 3020-01) and contained about 20 wt% MWNTs.

For mechanical loading, mechanical tensile tests were conducted using an INSTRON universal testing machine; strain were recorded using an extensometer. A mechanical load was applied at a standard rate of 0.01/min. Again for each type of sample, three specimens have been tested and compared for consistency.

Standard 2-Probe method was used to measure the electrical resistance of the nanocomposites. Since the percolation threshold has been evaluated to be around 4% weight CNT, only specimen with CNT weight percentages higher than the percolation threshold were tested. The tested specimens had the CNT weight percentages of 5, 7, 10 and 12.

Once the specimen is ready, electric current is applied to the specimen through an electrical circuit powered by a DC power source; the DC source was used for the tests for its simplicity. Nevertheless, it is worthwhile to note that in real-life practice AC source in 1 kHz is commonly used to avoid inaccuracy caused by polarization**.** The electrical data were then recorded using Labview software.

Figure 1: Specimen in the INSTRON machine with the extensometer and the electrical probes.

RESULTS AND DISCUSSION

During the mechanical loading, the electrical resistance of the system was obtained from the 2-probe measurement as described above; at the same time, the stress-strain curves were obtain form the tensile tests data. Figure 2 shows the stress-strain of the three MWCNT-PP composites with 5 wt% CNT.

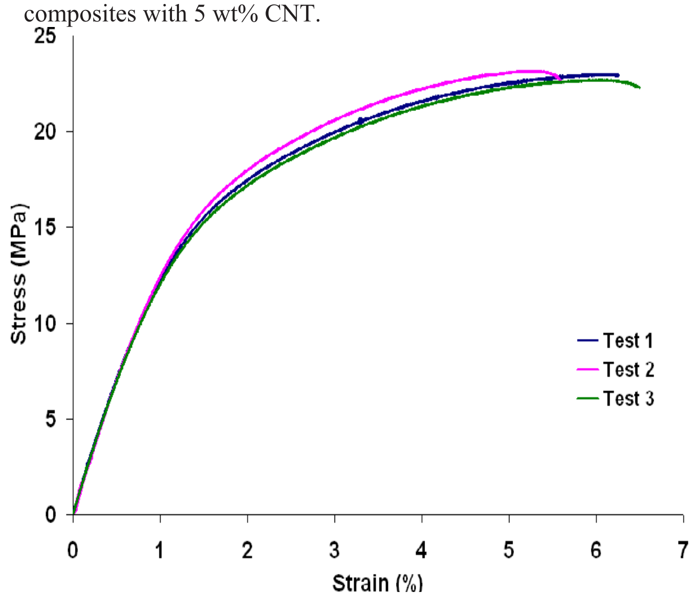


Figure 2: Stress-Strain curves for 5wt %MWCNT-PP composite specimens

The resistance change was also measured using 2-probes method. The change in resistance was not clearly pronounced. This could results to the fact that with 5%CNT, the conductivity of the composite specimen is still at the lowest level to highly sense the electrical resistance change. Figure 3 below shows the change in electrical resistance due to the strain change.

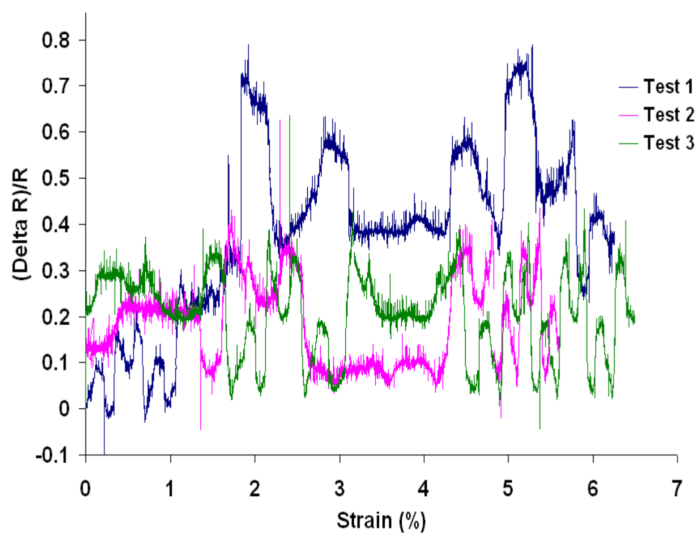


Figure 3: Electrical resistance change due to strain change for 5 wt% MWCNT-PP composite specimens

The same tests that were conducted on the PP-MWCNT composite specimens with 5 wt% CNT were also carried out on the composites specimens with 7% CNT.

Figure 4 shows the stress-strain curves for the composites specimens with 7% CNT.

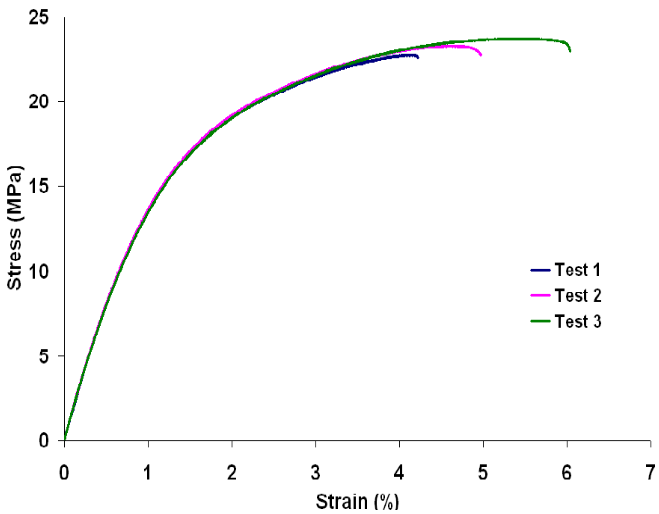


Figure 4: Stress-Strain curves for 7 wt% MWCNT-PP composite specimens

The electrical resistance change was more pronounced in this case even though an obvious pattern was not cleared present. It is illustrated in Figure 5 below.

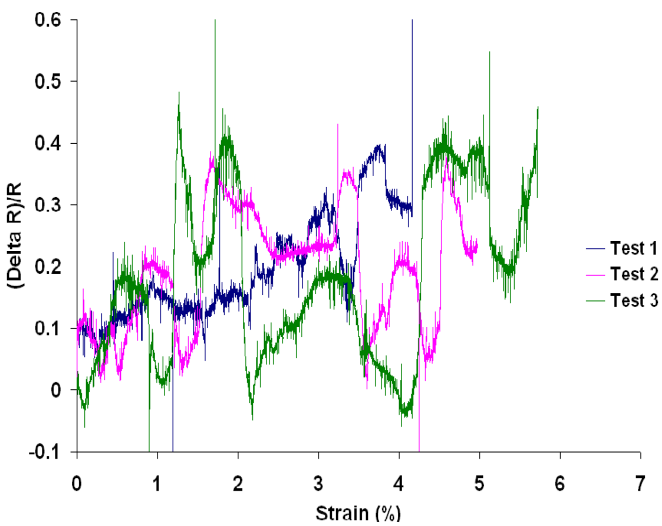


Figure 5: Electrical resistance change due to strain change for 7 wt% MWCNT-PP composite specimens

For the PP nanocomposites with CNT weight percentages of 10% and 12%, the resistance change sensitivity to mechanical loading was more pronounced and clear patterns could be observed. Figures 6 and 7 show the stress-strain and resistance change curves for the composites specimens with 10% CNT respectively.

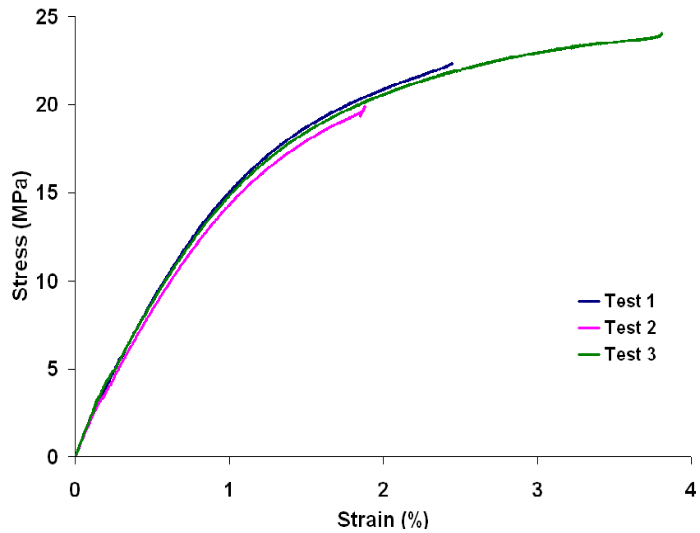


Figure 6: Stress-Strain curves for 10 wt% MWCNT-PP composite specimens

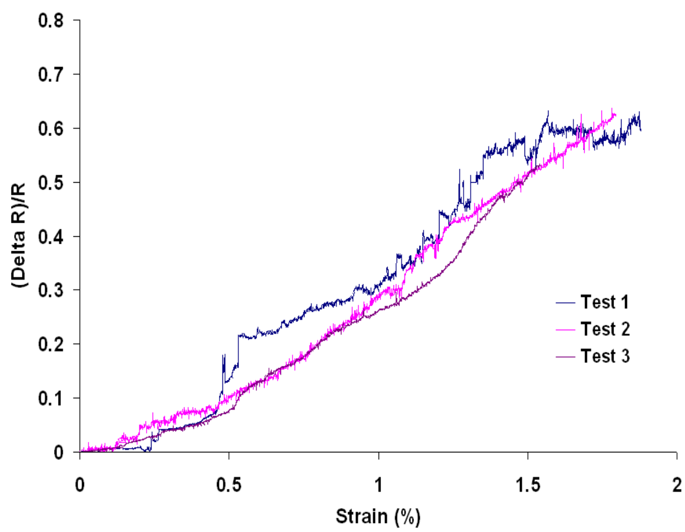


Figure 7: Electrical resistance change due to strain change for 10 wt% MWCNT-PP composite specimens

Form figure 7 above, we can realize the increase in sensitivity to strain change by the electrical electrodes. It can be noted that the change has a clear pattern compared to the CNT percentages of 5% and 7%.

The same increase in sensitivity was also present with PP nanocomposites wit 12 wt% CNT. Figure 8 shows stress-strain curves for the PP-MWCNT composites with 12 wt%.

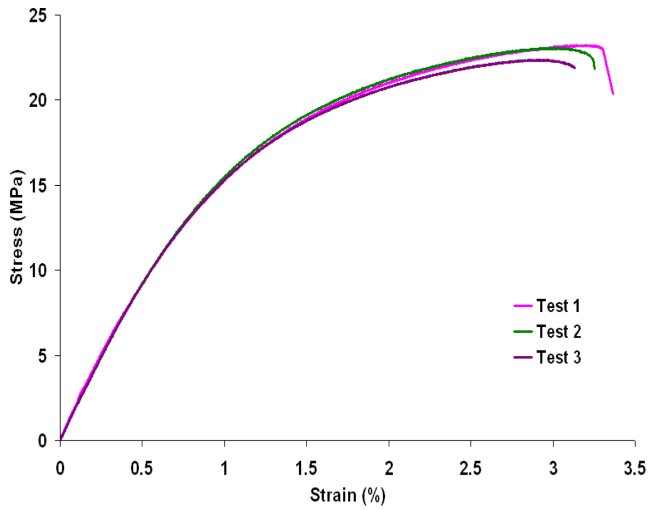


Figure 8: Stress-Strain curves for 12 wt% MWCNT-PP composite specimens

Figure 9 illustrates the resistance change sensitivity to strain change for the PP nanocomposites with 12 wt% CNT. Similarly to the 10 wt% MWCNT-PP composite specimens, the resistance change sensitivity is more pronounced.

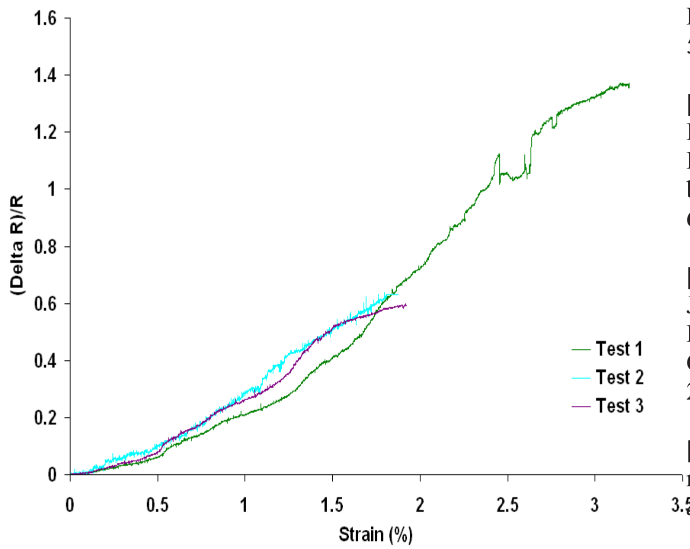


Figure 9: Electrical resistance change due to strain change for 12%MWCNT-PP composite specimens

CONCLUSION

The CNT-PP nanocomposites were produced by injection molding. The composites specimens showed electrical resistance change sensitivity to strain change. The level of sensitivity varies with the CNT wt%; it is less pronounced in composites wit 5 wt% and 7 wt% CNT. The reason might be the above CNT weight contents are close to the percolation threshold of about 3.8 wt%. The sensitivity became more pronounced for nanocomposites with 10 wt% and 12 wt% CNT.

The above results are a promising use of PP-MWCNT composites for the sensing purposes. This additional property will add to the already existing advantageous range of potential applications.

REFERENCES

[1] Ngabonziza, Y., Li, J. and Barry, C.F. 2008. Electrical Conductivity and Elastic Properties of MWCNT-PP Nanocomposites. *Proceedings of the ASME International Mechanical Engineering Congress and Exposition.*

[2] Chung, D.D.L. and Wang, S. 2003. Self-monitoring of Damage and Strain in Carbon Fiber Polymer-Matrix Structural Composites by Electrical Resistance Measurement. *Polymers and Polymer Composites*, **11**(7): 515-525.

[3] Todoroki, A., Tanaka, M. and Shimamura, Y. 2005. Electrical Resistance Change Method for Monitoring Delamination of CFRP laminates: Effect of Spacing between Electrodes. *Composites Science and Technology*, **65**:37-46.

[4] Shen, L., Li, J., Liaw, B.M., Delale, F. and Chung, J.H. 2007. Modeling and Analysis of the Electrical Resistance Measurement of Carbon Fiber Polymer-Matrix Composites. *Composites Science and Technology,* **67**: 2513-2520.

[5] Chung, D.D.L. 1998. Self-monitoring structural materials. Reports: A Review Journal, *Materials Science and Engineering*, **R22**: 57-78.

[6] Wang, S. and Chung, D.D.L. 1997. Self-monitoring of strain and damage by a carbon-carbon composite. *Carbon*, **35**(5): 621-630.

[7] Chung, D.D.L. 2000. Thermal analysis of carbon fiber polymer-matrix composites by electrical resistance measurement. *Thermochimica Acta*, **364**: 121-132.

[8] Mei, Z. and Chung, D.D.L. 2000. Thermal stress-induced thermoplastic composite debonding, studied by contact electrical resistance measurement. *International Journal of Adhesion and Adhesives*, **20**: 135-139.

[9] Wang, S. and Chung, D.D.L. 1999. Apparent negative electrical resistance in carbon fiber composites. *Composites: Part B*, **30**: 579-590.

[10] Wang, S., Mei, Z. and Chung, D.D.L. 2001. Interlaminar damage in carbon fiber polymer-matrix composites, studied by electrical resistance measurement. *International Journal of Adhesion and Adhesives*, **21**: 465-471.

[11] Wang, X. and Chung, D.D.L. 1998. Self-monitoring of fatigue damage and dynamic strain in carbon fiber polymer-matrix composite. *Composites: Part B*, **29B**: 63-73.

[12] Wang, X. and Chung, D.D.L. 1998. Short carbon fiber reinforced epoxy coating as a piezoresistive strain sensor for cement motar. *Sensors and Actuators A*, **71**: 208-212.

[13] Wen, S. and Chung, D.D.L. 1999. Piezoresistivity in continuous carbon fiber cement-matrix composite. *Cement and Concrete Research*, **29**: 445–449.

[14] Chen, P.-W. and Chung, D.D.L. 1996. Concrete as a new strain/stress sensor. *Composites: Part B*, **27B**: 13-23.

[15] Mei, Z. and Chung, D.D.L. 2000. Effects of temperature and stress on the interface between concrete and its carbon fiber epoxy-matrix composite retrofit, studied by electrical resistance measurement. *Cement and Concrete Research*, **30**: 799-802.

[16] Pham,G. T., Park ,Y., Liang, Z., Zhang, C., Wang, B. 2008. Processing and modeling of conductive thermoplastic/carbon nanotube films for strain sensing. *Composites: Part B,* **39**: 209–216.

Application of a Structural Integrity Assessment Software

Marko RAKIN
Faculty of Technology and Metallurgy, University of Belgrade
Belgrade, Serbia
e-mail: marko@tmf.bg.ac.rs

Nenad GUBELJAK
Faculty of Mechanical Engineering, University of Maribor
Maribor, Slovenia

Bojan MEDJO
Faculty of Technology and Metallurgy, University of Belgrade
Belgrade, Serbia

Taško MANESKI
Faculty of Mechanical Engineering, University of Belgrade
Belgrade, Serbia

Aleksandar SEDMAK
Faculty of Mechanical Engineering, University of Belgrade
Belgrade, Serbia

ABSTRACT

In this paper, a brief overview is given of two projects concerning structural integrity assessment, including on-line software based on SINTAP procedure. An example of the software usage is presented, and the obtained results are compared with the experimental data and discussed.

Keywords: Structural Integrity Assessment, Welded Joint, Crack Initiation, Fracture Initiation

INTRODUCTION

Fracture mechanics, having in mind its theoretical and experimental techniques [1, 2], is a good basis for reliable structural integrity assessment procedures. There are several such procedures, and they are already being used for many materials; however, they can still be improved and adjusted, to achieve better efficiency and lower costs of design, construction and maintenance. The approach used for SINTAP (Structural integrity assessment procedure) [3] offers a very good basis for software solution, one of which is developed through MOSTIS (Mobile structural integrity assessment system) project [4], as a segment of a structural integrity assessment system. Four standard assessment levels of SINTAP procedure (0, 1, 2 and 3) are used, ranging from simple but conservative approaches where data availability is limited, to more accurate and complex approaches. A very important issue is assessment of structures comprised of more than one material - welded joints. The results can be presented as crack driving force (CDF) or failure assessment diagrams (FAD). The crack driving force (for example, the applied J-integral) can be plotted as a function of flaw size for different applied loads, or as a function of load for different flaw sizes and compared to the material's resistance to fracture. In a FAD, an assessment is represented by a point or a curve on a diagram and failure judged by the position of the point or curve relative to a failure assessment line. The result of this procedure is information whether the actual or postulated flaw will cause failure of the structure or it is possible to continue with its exploitation. The main principle is that failure occurs if the applied crack driving force exceeds the material's fracture resistance. Another step in the development of assessment procedures was made through FITNET project [5], which included fatigue, corrosion and creep modules, in addition to the fracture procedures included in SINTAP.

BACKGROUND

The purpose of structural integrity assessment is to determine the significance, in terms of fracture and plastic collapse, of flaws present in metallic structures and components. Loading or flaw dimensions can be varied, in order to check the possible increase of loading and/or flaw size that will lead to failure. It is important to note that this approach is not intended to supersede existing methods, but to serve together with them throughout the lifetime of a structure. It can be used for assessing in the design phase - in order to specify the material properties, design stresses, inspection procedures/intervals and acceptance criteria. It can also be used for fitness-for-purpose assessment during the fabrication, with respect to applied fabrication standards. During the operation phase, it can be used to decide whether continued use of a structure or component is safe (i.e. it is safe to continue operation until a repair can be carried out in a controlled manner), despite detected flaws or modified service conditions.

Having in mind that analytical and other methods used for the creation of SINTAP/FITNET procedure (and therefore also the software which is part of MOSTIS system) can give an estimate of the state of the structure, but cannot provide some more detailed information (e.g. stress or strain data at some important location in the structure), a new system - OLMOST (On-line monitoring of structures and fatigue) [6] is currently being developed. It is an integrated hardware and software solution for on-line and on-site measurement of structure state during its service life, in order to prevent failures due to flaws and inappropriate design. Configured as an expert system for on-line monitoring and automatic analysis of measured data, including automatic warning signals to supervisor, it is based on a data bank of materials and stress-strain behavior of components.

Strain state of the structure will be assessed by optical stereometric measurement of surface in critical spots, and/or some other measurement methodologies, depending on the construction type, operating conditions, safety requirements, etc. Having in mind that faults may change the behavior of structure, changes of the structure state can be used to indicate some flaws and anomalies which can not be measured directly. Wherever possible, wireless measurement will be applied; sensors will be wirelessly connected to the processor unit, and the signals will be assembled in mobile computer device. This device can be connected to the internet via GSM mobile network, and then to the server with master program for fault identification. In case of overloading or missing input parameters it will provide different warning signals, depending on the measured data.

Stress-strain behavior of the structure or some of its components will be assessed by numerical - finite element (FE) modeling. Comparison between numerical results and measured strains (or other appropriate quantities) provide relevant information concerning the stress state of the component and loading. The results for critical component in regular service will be used to establish an acceptable loading window. If the deformation state does not fall into this window, the expert system is going to provide decision to safe shut-down or stop the use of structure.

The reliable estimation of flaw size and its position in the component can be made by comparing the measured deformation behavior with the results of the numerical model with assumed size and position of the flaw. Another important usage of this new system is retrieving the loading history and assessment of cumulative damages occurring during the service life of the structure. The main purpose is on-site failure assessment analysis and estimating the remaining lifetime of damaged structure, to improve repair planning and optimize the component life cycle, with a possibility to provide appropriate commands to control the equipment.

APPLIED PROCEDURE

An example for assessment procedure using MOSTIS software is presented on a welded single-edge notched bend (SENB) specimen. The base metal (BM) is high-strength low-alloyed (HSLA) steel NIOMOL 490. Fatigue pre-crack is located in the weld metal (WM), along the axis of symmetry of the joint (one half of the specimen is shown in Fig. 1). Properties of the base metal and weld metal are given in Table 1. It can be seen from these data that the analysed joint is overmatched, having in mind that the mismatch ratio (ratio of the yield strength of the weld metal and the base metal) is larger than 1. As already mentioned, SINTAP/FITNET procedures can take into account this difference of material properties across the welded joint.

The behaviour of the joint under external loading is analysed using level 2 SINTAP procedure, and the results are presented in FAD diagram (Fig. 2). This diagram represents the change of the structure state during the increase of loading (straight line) and the critical state of the structure (failure assessment line). The value K_r, ordinate of the diagram, represents the ratio of the applied stress intensity factor K_I and the critical stress intensity factor K_{Ic}. The abscise L_r is the ratio of the applied loading and the plastic limit load of the structure.

Table 1 Properties of the materials

	WM	BM
E [GPa]	183.8	202.9
$R_{p0.2}$ [MPa]	648	545
R_m [MPa]	744	648

RESULTS AND CONCLUDING REMARKS

It can be seen that the increase of loading from 10 kN (point A) to 70 kN (point C) is shifting the structure toward the critical state, and eventually into the critical state (point B). Each of the points on the straight line corresponds to a specified loading level, and as the load increases (with the increment of 5 kN), the point is moving in the arrow-marked direction. The crack length is kept constant during the calculations (and equal to the initial fatigue pre-crack length), having in mind that the subject of this analysis is crack growth onset. The influence of the joint width on the fracture initiation is analysed in [7], using the local approach to fracture - Gurson-Tvergaard-Needleman (GTN) model.

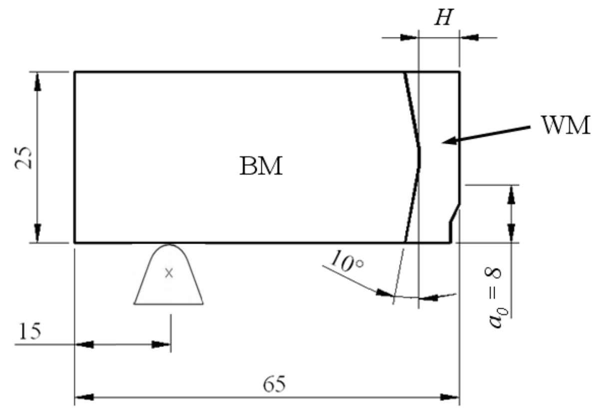


Fig. 1 Dimensions of SENB specimen and welded joint ($2H$ = 6 mm)

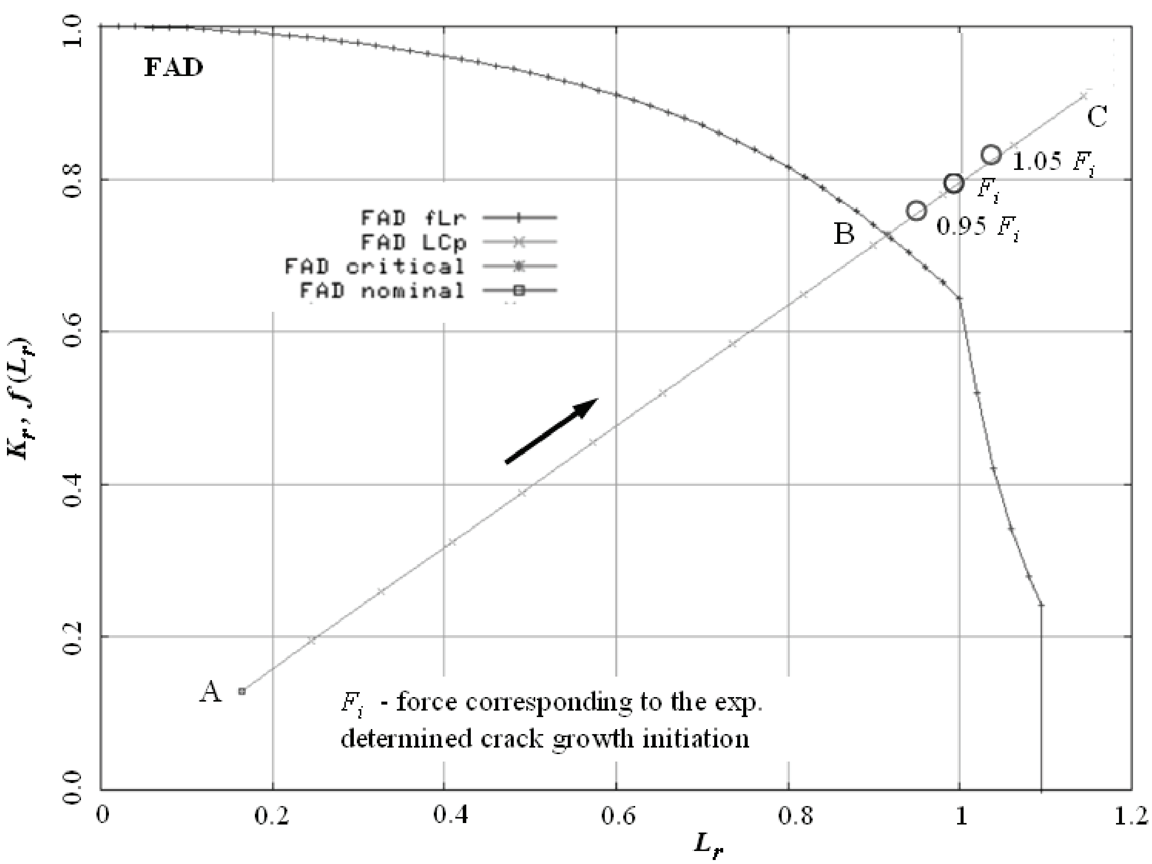


Fig. 2 FAD diagram for welded SENB specimen

The moment of crack growth onset is determined experimentally, and the measured force value that corresponds to it (F_i) is compared with the results obtained using MOSTIS. It turns out that this loading is close to the critical one according to the failure assessment line, because the point that corresponds to F_i at FAD diagram belongs to the critical state of the structure. Therefore, the state of the structure at fracture initiation is correctly predicted, and the assessment using MOSTIS is on the safe side when compared to the experimental investigations, setting a lower loading level as the critical one.

Additionally, the value that corresponds to the experimentally determined crack initiation was varied, in order to check the sensitivity of assessment to the variation of experimental data; the interval $0.95\ F_i < F < 1.05F_i$ was used. It can be seen (Fig. 2) that decrease of this value in amount of 5% brings the appropriate point rather close to the failure assessment line. Having in mind that the crack initiation is difficult to determine exactly, it would be preferable to use a certain safety factor to ensure that the failure assessment will be on the safe-side and will not overestimate the load-carrying capacity of the structure. It can be expected that the highest level of SINTAP procedure would give less conservative results. However, only basic material data are used in the presented example, because they are usually known for most materials in exploitation.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support of Serbian Ministry of Science under the Eureka projects E! 3927 “Mobile structure's integrity system - MOSTIS” and E! 5348 “On line monitoring of structures and fatigue - OLMOST”

REFERENCES

[1] T.L. Anderson, **Fracture mechanics**. London: CRC Press, 1995.

[2] K.H. Schwalbe, **Basic engineering methods of fracture mechanics and fatigue**. Geesthacht: GKSS Research Center, 2001.

[3] SINTAP: Structural Integrity Assessment Procedure, EU-Project BE-1462, Brite Euram Programme, 1999.

[4] MOSTIS: Mobile structure's integrity system, Eureka project E! 3927.

[5] FITNET: European Fitness-for-Service Network, http://www.eurofitnet.org.

[6] OLMOST: On line monitoring of structures and fatigue, Eureka project E! 5348.

[7] M. Rakin, N. Gubeljak, M. Dobrojević, A. Sedmak, “Modelling of ductile fracture initiation in strength mismatched welded joint”. **Engineering Fracture Mechanics**, Vol. 75, 2008, pp. 3499-3510.

Pulsed Atmospheric Pressure Plasma System Applied to PCBs Surface Treatment

Fuhliang WEN, Jhenyuan LIN
Department of Mechanical & Computer-Aided Engineering / Graduate Institute of Automation & Mechatronics, St. John's University, Tamsui, Taipei County 25135, Taiwan

and

Hungjiun WEN, Kuo-Hwa CHANG
Department of Industrial & System Engineering, Chung Yuan Christian University, Chungli, Taoyuan County 32023, Taiwan

ABSTRACT

Since a pulsed atmospheric pressure plasma system (PAPPS) could be applied to surface modification for products with complex geometry in a lack of vacuum environment, it definitely saves a lot of operation cost. The PAPPS has unique features in lower surface charge accumulation and higher processing repeatability due to pulsed parameter adjustment from clever tuning technique in manufacturing recipes, which consisted of an adapted DC power supply, a pulsed high-voltage controller, and a high-frequency pulsed transformer. Based on 2 kW maximum output power and limited working voltage at 20 kV with a bipolar pulsed output mode, through communication interface, the PC remote control system was built for setting the pulsed parameters and for monitoring the variation of electrical power outputs. This low-temperature plasma jet would be employed as the tool for PCB surface cleaning and activating at an atmospheric chamber. Experimental results have demonstrated that the PAPPS is able to offer a stable plasma jet while a tunable DC power source, pulsed parameters and air pressures. After plasma surface treatment, the hydrophilic feature of the printing circuit board (PCB) has been enhanced that meaning of adherent substance ability on PCB surface is improved for component welding and in yield rate.

Keywords: Atmospheric pressure plasma, Pulsed parameter, Surface modification, Plasma jet, PCB treatment

1. INTRODUCTION

In electronic industry, usually printing circuit boards (PCBs) have been adopted as the substrate for mounting ICs, transistors, resistors and other electrical components and wiring connections among each part. PCBs have been developed from a single-layer to multi-layer structure where needs excellent adhesive quality for surface mounting. Therefore, the plasma is used to surface modification to improve the surface function and quality or yield rates for products [1, 2]. Since the atmospheric pressure plasma could be applied to surface modification for products with complex geometry in a lack of vacuum environment, it definitely saves a lot of operation cost. The pulsed atmospheric pressure plasma supply system (PAPPS) has unique features in lower surface charge accumulation and higher processing repeatability due to pulsed parameter adjustment from clever tuning technique in manufacturing recipes [3]. Therefore, the PAPPS was adopted to remove contaminant and to improve the hydrophilic feature on the PCB surface via plasma ion bombardment. Also, the adhesion of metal coating or surface mounting on PCBs would be enhanced and reduce the usage of organic solution for cleaning. There are two major sections described in this paper. First section, the practical design was performed for one set of a pulsed high-voltage power supply involved a high-frequency pulsed transformer which was adapted a DC power supply and a pulsed power controller. Next, under appropriate dry-pressure air and high-pulsed voltage supply, the atmospheric pressure plasma jet was generated via a plasma jetting nozzle. This low-temperature plasma jet would be employed as the tool of surface treatment for PCBs at an atmospheric chamber. Through measured instruments of contact angles and 3D surface profiles, the surface characteristics on PCBs were detected after surface treatment.

2. CONFIGURATION OF PULSED ATMOSPHERIC PRESSURE PLASMA SYSTEM

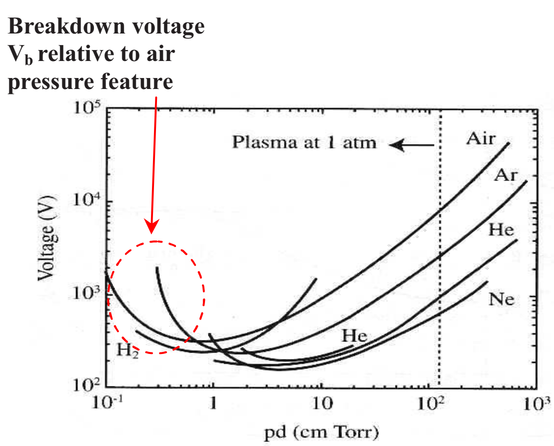


Fig. 1 Expectable breakdown voltage on the cylindrical electrode of the plasma nozzle based on Paschen curve [1].

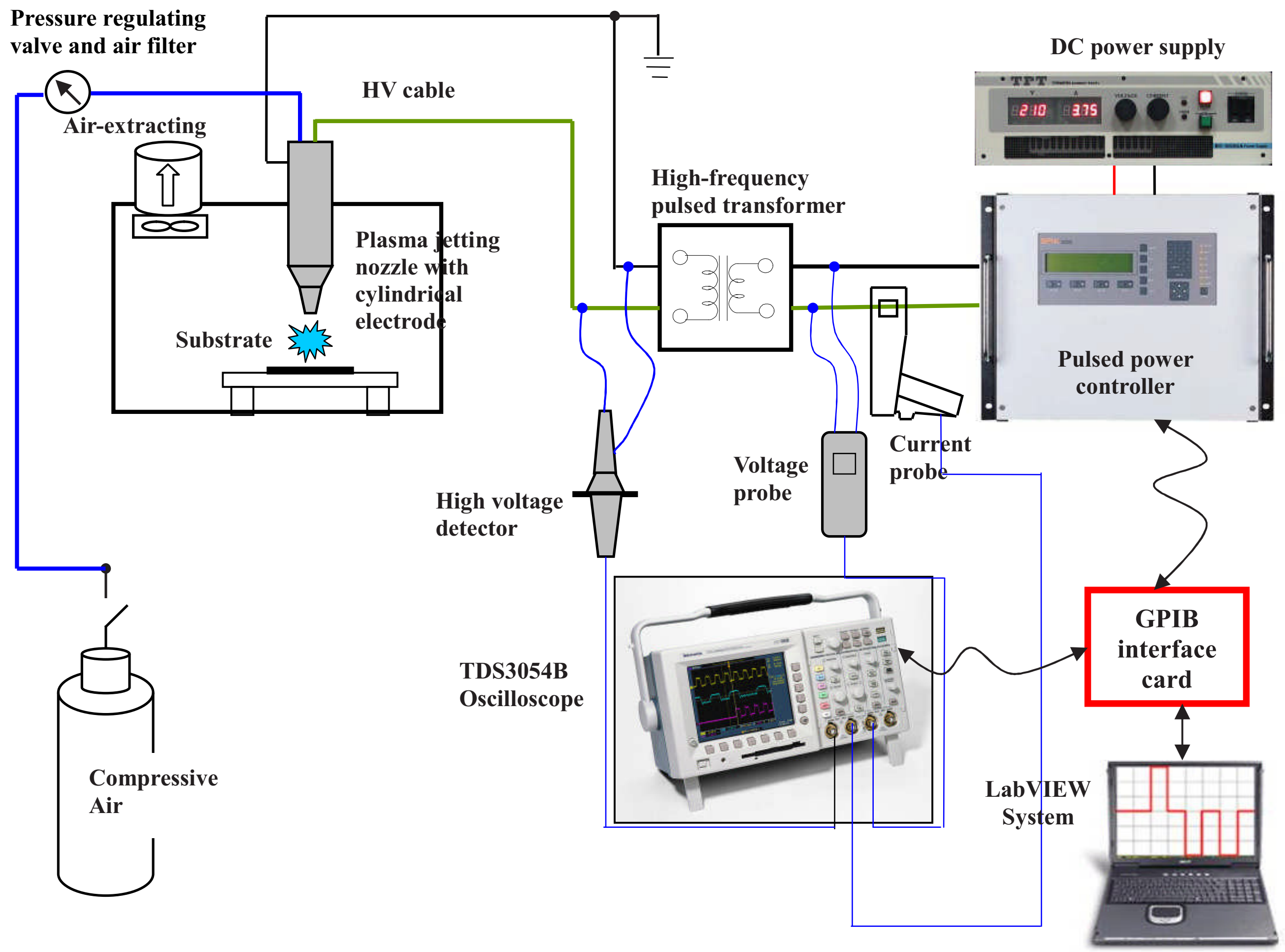


Fig. 2 Pulsed atmospheric pressure plasma configuration and LabVIEW monitoring system

Due to the high breakdown voltage for plasma at one atmospheric pressure, there are intensely collisions among air molecules or ion particles to easily form the electrical arc between electrodes of the plasma jetting nozzle. It is high possibility to damage the electrodes or the treated surface on the substrate. Hence, according to Paschen's law [4, 5], the designate breakdown voltage V_b for the pulsed power supply would be calculated as equation (1),

$$V_b = B\frac{pd}{\ln\left[\dfrac{Apd}{\ln\left(1+\frac{1}{\gamma}\right)}\right]} = \frac{B(p\cdot d)}{\ln[A(p\cdot d)] - \ln[\ln(1+\frac{1}{\gamma})]} \qquad (1)$$

where p is the air pressure, d presents the distance between electrodes, and γ effect displays the average amount in the plasma of secondary electrons released from cathode surface after positive ion struck at each time. In addition, A and B stand for the ionization coefficient of collision, which the effect of constant α depends on the various air feature but it can be obtained from relative experiments. They have the relationship as equation (2) described as,

$$\alpha = pA\exp\left(\frac{B}{\left(\frac{E}{p}\right)}\right) \qquad (2)$$

where E (=V/d) is the electrical field applied on voltage V at the distance d between electrodes. Actually, the breakdown voltage V_b is the function of the multiplication of air pressure p and electrode distance d, i.e., $V_b=f(pd)$. Based on above statement, we always followed Paschen curve, as shown in Fig. 1, to determine the applied voltage.

A pulsed atmospheric pressure plasma system consisted of a pulsed high-voltage power supply, the air piping and supply equipment, and a LabVIEW monitoring platform for electrical powers, as shown in Fig. 2. The pulsed high-voltage power supply provided the proper AC high-voltage field through a DC power supply controlled by pulsed power controller and regulated by a high-frequency pulsed transformer. Its electrical specifications are maximum output power of 2 kW, the operating frequencies within 25 kHz, and limited working voltage at 20 kV with a bipolar pulsed output mode. Through GPIB and RS232 communication interface, the PC remote control system was built for setting the pulsed parameters and for monitoring the variation of electrical power outputs via the

LabVIEW platform, which also consisted of an oscilloscope, a high voltage differentiate probe, a current probe, and a high voltage detector to record the output voltage onto the plasma nozzle electrode after the high-frequency pulsed transformer. In addition, the air supply was a simple piping system which adopted the pump, piping and vessel for the generation of compressive air, plus some adjustable valves and appropriate air filters.

3. PLASMA EXPERIMENTAL DESIGN FOR PCB SURFACE TREATMENT

In recent decade, the pulsed plasma has larger progress in coating technique and surface modification, because of its intensive plasma energy tuned by electrical parameters such as power output, frequency, duty cycle, etc [6]. Also, it has unique feature for the arc-suppressed function. However, there are lots of processing factors involving into the PCB surface modification for adhesive improvement. We adopted that the design of experiment (DOE) method to investigate the correlative model in interdisciplinary knowledge like plasma surface processing. Using the categorized technique for environmental factors and controlled factors, the interactive relationship among these factors could be testified for the better operating parameters. Thus, multiple factor experiments would be identified as the optimal combination for PCB surface modification to enhance hydrophilic or adhesive feature.

3.1 DOE parameter determination

To find the influenced factors for PCB surface modification, we performed DOE method to determine their correlation among factors. Basically, the conventional DOE procedure should be fulfilled as: (1) collecting the possible factors, like processing time, air flow rate and category, plasma jetting distance, frequency, duty cycle, electrical power (including voltage and current), and substrate materials; (2) classified these influenced factors based upon their characteristics and effects on experimental results in percentage; (3) determining reactive levels for factors; (4) setting the level range. Through above procedure, we finally decided the parameters, including plasma treatment time, duty cycle, power, and ignition voltage onto jetting nozzle [7, 8], would be major influenced factors.

3.2 Plasma surface experiment and measurement

According to pre-determined processing parameters, the plasma jetting stream would perform ion bombardment on a PCB of 40 mm × 50 mm for surface modification after solution cleaning with methyl alcohol. Also, all treated pieces would measure the contact angle through pendent drop for surface tension, as well as they were detected by a 3D surface profile measuring instrument. The hydrophilic or hydrophobic feature of the modified PCB surface was identified by contact angles as shown in Fig. 3. After ion bombardment of PAPPS surface treatment, the roughness of the PCB surface has obvious modification, as shown in Fig. 4, where displays the uneven surface.

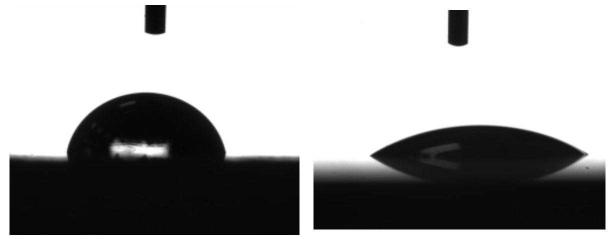


Fig. 3 Measuring contact angle by pendent drop mode: (a) contact angle 99.54° before plasma surface treatment; (b) contact angle 35.55° through plasma surface treatment for 12 seconds.

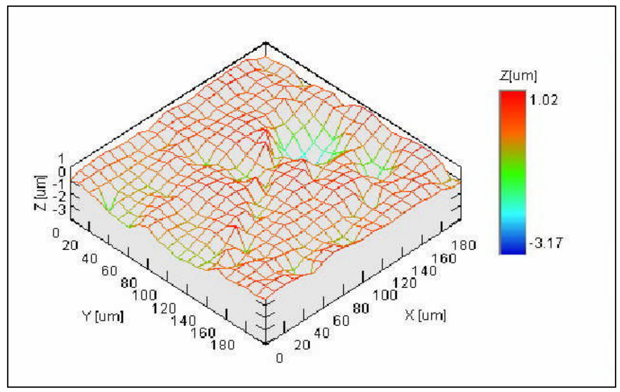


Fig. 4 Surface profile after 12 second pulsed ion bombardment onto the PCB surface with duty cycle 35% near the operating frequency of 11.6 kHz and air pressure of 0.9 kg/cm^2.

4. RESULTS AND DISCUSSION

The pulsed duty cycle is reversely proportional to applied voltage onto the electrodes of the plasma nozzle, as shown in Fig. 5. Since the duty cycle stands for the input energy from the electrical power at each valid period, the larger duty cycle of electrical power input then has more supplied energy to compensate the plasma energy loss when ion bombardment continuously impacts on the PCB surface and causes the ion consumption. In contrast, if less duty cycle needs higher applied voltage onto the electrodes to offer the instant energy stock to sustain the plasma status during the ion bombardment in T_{on} of either 16 µs or 20 µs. However, there are two various slope tendencies where T_{on} of 20 µs has more sensitivity than that of 16 µs pulsed operation parameter. Otherwise, both of them have the better efficiency or called less power consumption within the range of 40%~48% duty cycle, as shown in Fig. 6. Even though, so far, we did not have the obvious proof, perhaps the material properties of the substrate and supplied air categories would be the domain factors which need the further research to demonstrate this guess.

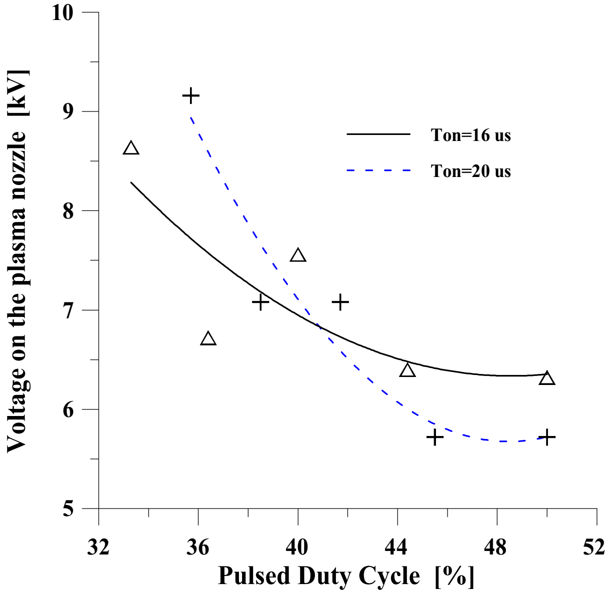


Fig. 5 Pulsed duty cycle with respect to the supplied voltage on the plasma jetting nozzle based on the different T_{on} time.

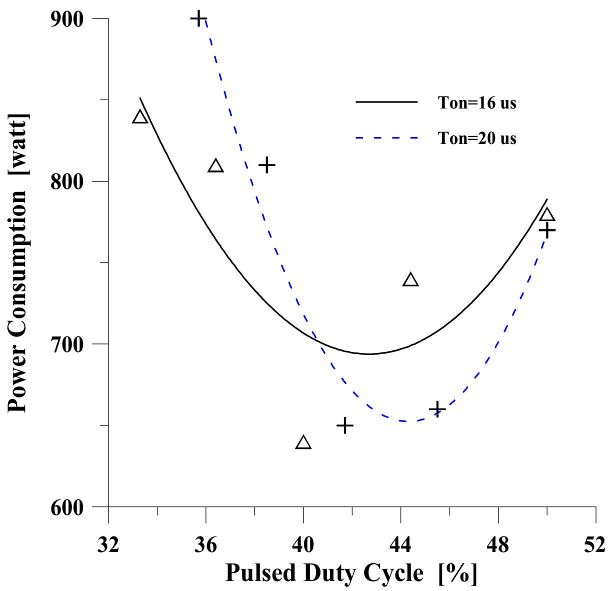


Fig. 6 Pulsed duty cycle with respect to the plasma power consumption based on the different T_{on} time.

Figures 7 and 8 illustrates the treatment time of plasma ion bombardment effected on the contact angle and surface roughness. When the longer duration under plasma ion bombardment, the drop contact angle is reduced to less 50°; that is, the modified surface has more hydrophilic feature. However, at the plasma treatment time between 4 seconds and 10 seconds, there is no obvious improvement in contact angles, except for 12 seconds or above, as shown in Fig. 7. In another point of view, the average roughness curve, Ra, could be

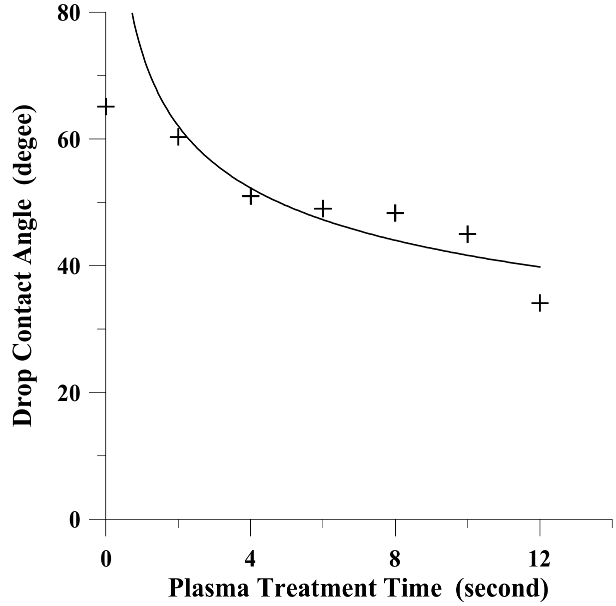


Fig. 7 Effect of plasma treatment time on the contact angle with pendent drop mode.

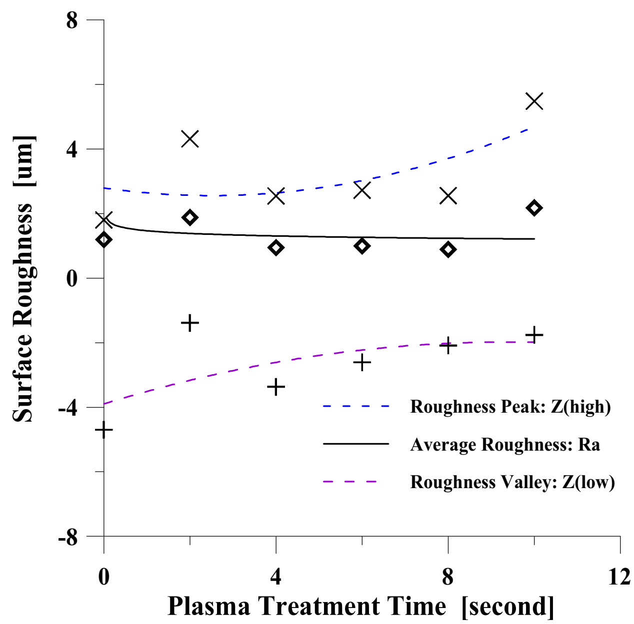


Fig. 8 Effect of plasma treatment time on the surface roughness.

changed effectively by plasma treatment time, as shown in Fig. 8. The plasma ion bombardment is positively affected on the PCB's surface roughness. From this figure, we understands that the improved hydrophile might be the difference in roughness peak Z(high) and roughness valley Z(low), but not in R_a.

5. CONCLUSIONS

Experimental results have demonstrated that the pulsed atmospheric pressure plasma supply system is able to offer a stable plasma stream while a tunable DC power source, pulsed parameters and air pressures. One of optimal parameters includes the operating frequency of 25 kHz and 46％ duty cycle, air pressure in 0.4 kg⁄cm^2, and electrical power of 0.77 kW, resulting in contact angles of drop-shape varied from 65.1° fallen into 34.1° after 12 second plasma surface treatment. Also, the surface average roughness (Ra) was carried from 1.19 μm levitated to 2.17 μm after surface modification in 10 seconds. It is an obvious proof of the PAPPS system useful capability in PCB surface cleaning and activating. After plasma surface treatment, the hydrophilic feature of PCB plates has been enhanced that meaning of adherent substance ability on PCB surface is improved for component welding or mounting.

6. REFERENCES

[1] Andreas Sch¨utze, James Y. Jeong, Steven E. Babayan, Jaeyoung Park, Gary S.Selwyn, and Robert F. Hicks, “The Atmospheric-pressure Plasma Jet: A Review and Comparison to Other Plasma Sources”, **IEEE Transactions on Plasma Science**, Vol. 26, No. 6, 1998.

[2] N St J Braithwaite, ”Introduction to Gas Discharges,” **Plasma Sources Science Technology**., Vol. 9, 2000, pp.517-527.

[3] E. Panousis, F. Clement, J.F. Loiseau, N. Spyrou, B. Held, J. Larrieu, E. Lecoq, and C. Guimon, ”Titanium Alloy Surface Treatment using an Atmospheric Plasma Jet in Nitrogen Pulsed Discharge Conditions,” **Surface and Coatings Technology**, Vol. 201, Issue 16-17, 2007, pp. 7292-7302.

[4] Claire Tendero, Christelle Tixier, Pascal Tristant, Jean Desmaison, and Philippe Leprince, **Spectrochimica Acta,** part B 61, 2006, pp.2-30.

[5] H.E. Wagnera, R. Brandenburga, K.V. Kozlovb, A. Sonnenfeldc, P. Michela, and J.F. Behnke, “The Barrier Discharge: Basic Properties and Applications to Surface Treatment,” **Vacuum**，Vol. 71, 2003, pp.417-436.

[6] G̮unter Mark, **Tutorial: PPST 2003**, June 9th-10th, Tokyo Japan, 2003.

[7] http://www.mie.utoronto.ca/labs/tsl/Research/DCPlasma/DCPlasma.html

[8] Edward V. Barnat and Toh-Ming Lu, **Pulsed and Pulsed Bias Sputtering Principles and Applications,** Kluwer Academic Publishers, 2003.

Understanding the drawbacks of technology backup to promote IT usage among small Mexican firms

Pilar ARROYO
Management and Marketing Science Department, Tecnológico de Monterrey campus Toluca
Toluca, Méx. 50110, México

and

Victoria EROSA
Graduate School of Business Administration, Universidad Autónoma de Tamaulipas campus UAM, Comercio y Administración Victoria.
Cd. Victoria, Tamps. México

ABSTRACT

This work performs a diagnosis about the factors that determine the use of information and communication technologies (ITC) and information systems (IS) among small retailing and service firms located in the downtown area of a large Mexican city. The study uses a multi-case approach to classify firms in terms of their technology infrastructure and the extent of e-business applications. Once firms are classified, their degree of IT usage is related to the following factors: IT innovation promoted by CEO, IT knowledge sharing among employees and technical support provided by consultants and system vendors. Case analysis shows those firms that make full use of their technical infrastructure have CEO/owners with better perceptions about the usefulness of Internet-based technologies, have established collaborative relations with consultants, and employees interested in using IT to facilitate and improve their tasks.

Keywords: Information and communication technologies, small business, knowledge sharing, technical support.

1. INTRODUCTION

Research asserts ITs have the potential to support the competitiveness of small to medium enterprises (SMEs) [5] [8], [9], [17]. However, authors like Nieto and Fernández [13] and Taylor and Murphy [18] present data for European and USA firms that indicate that SMEs are less engaged with digital economy than larger firms. Despite the advantages Internet based technologies represent for the advancement of electronic relations with business partners, the improvement of information flows and customer service, and the reduction in marketing and order-processing costs [9], the current literature evidences that SMEs still have a limited strategic view about e-commerce and e-business initiatives [5]. Then decisions in relation of IT are driven by operational issues such as an immediate perceived benefit or the pressure of suppliers and clients [9].

It has been argued that one of the main reasons of the low IT diffusion among SMEs is their relative disadvantage with respect to larger firms because of their limited capital resources which lead to poor investments in IT infrastructure [10]. However other barriers, even more important, have been identified. Among them figure the beliefs about the inappropriateness of existing technologies to attend SMEs' needs; perceptions about insecurity and privacy (especially about payments online); lack of computing skills and inadequate technical support to end users; inadequate national policies promoting IT adoption among SMEs and a poor vision about the potential contribution of ITs to the firm competitiveness.

Then the research question of this work is stated as: How important with respect to IT infrastructure, are the availability of computing skills, technical and top management support for the effective usage of Information and Communication Technologies (ITC) among Mexican small retailers and service firms? The organization of the paper is as follows. First section discusses literature related to factors affecting IT usage, then the methodology followed to identify and classify firms with different levels of IT use and infrastructure is presented. Third section presents the analysis of the cases and identifies those factors found to be determinant to take advantage of available IT resources. Final section elaborates conclusions, provides directions for future research and discusses the limitations of the study.

2. LITERATURE REVIEW

The perceptions about *usefulness* and *ease of use* proposed by the TAM (Technology Acceptance Model) model to explain the usage of ITs have been shown to be critical factors for IT adoption among SMEs [10], [12], [18]. But according to Ndubisi and Jantan [12], the usage of Information Systems (IS) is driven not only by these perceptions, but also by the users' computing skills and the before and after-sales technical support provided by vendors and designers. The importance of acquiring external computing support and training has been also recognized by other authors like Lin [11] in the context of Taiwan SMEs and by Thong [19] in the case of Singaporean small business.

Caldeira and Ward [1] studied the influence of various factors of the internal context of the small firm on the extent of IS/IT success (measured as usage and management satisfaction with current IS/IT). They conclude that the technical competences of employees, knowledge available and top management attitudes through IS/IT are more relevant to IS/IT usage than financial resources. Those SMEs more concerned with the development of internal technical competences were more successful in adopting and obtaining benefits from IS/IT adoption than firms concentrated on the quality of the acquired software or systems.

The development of technical competences at the organizational level requires of individual abilities, the sharing and application of knowledge to solve or improve work tasks. Usually SMEs are not able to contract IS/IT professionals, however technical skills could be developed in house through more informal training mechanisms such as mentoring and interpersonal interchanges of experiences. The acquisition of external knowledge along with the dissemination and application of such knowledge (identified as absorptive capabilities) to gain competitive advantage is less well recognized among SMEs [16].

Small firms often do not have a systematic approach to develop, share or exploit knowledge. According to Desouza and Awazu [4], at SMEs knowledge transfer from CEO to employees and between employees mainly occurs through socialization. This means knowledge is shared and immediately put into practice via informal mechanisms like face-to-face interactions, observation, tutoring and job rotation to gain experience. To facilitate knowledge sharing is relevant to motivate and stimulate employees to learn from others and clearly communicate the reasons and value of knowledge/experience interchanges. The provision of managerial directions and infrastructure to facilitate knowledge sharing between employees and to learn from customers and suppliers influences the utilization of IS/IT among SMEs [2].

As top managers in SMEs are usually directly involved in most strategic and tactical decisions, they have a strong influence on the adoption of IS/IT [1], [7], [14]. If top managers are convinced on the benefits and potential contribution of IS/IT to the firm's competitiveness, they promote their use and encourage employees to propose new applications to current systems. On the contrary, when top management is not committed to IS/IT or has unrealistic expectations about to what technology can do, diffusion is limited and restricted to specific applications suggested or required by business partners (eg. Online payment imposed by customers).

Since the theoretical models used to identify the internal factors that promote the adoption and use of ITs among SMEs have been mainly validated in developed countries, it is important to gain additional insight about the factors influencing IT usage among Mexican SMEs in order to elaborate suggestions about how to take full advantage of actual IT infrastructure. This work first performs a diagnosis of current IS/ITC applications among local SMEs and then identifies the effect of CEO support, technical backup and employees' knowledge sharing on the exploitation of available technologies.

3. METHODOLOGY

The study was conducted among micro and small firms located in the downtown area of one of the largest and most dynamic cities in the central part of Mexico. The interest in these particular firms is due to the following two reasons:

a) Small firms, and in particular those operating in the service and retailer sectors constitute the primary economic base in the downtown area of large Mexican cities. These firms are usually well established familiar businesses that have been operating for several years in the same geographic area and provide employment to a significant percentage of the local population.
b) Small retailers are subject to a strong pressure for survival due to the allocation of new malls in the periphery of large Mexican cities. Large departmental stores (anchor stores), retailing and service chains in these malls have gained the market previously served by these small downtown firms. These large retailers and service firms not only compete by offering a large variety of merchandise at competitive prices, but also deploy e-business activities to improve customer service and business efficiency.

The case study was the selected research methodology because this qualitative technique enables us to profoundly study a small number of firms from the point of view of the individuals directly working with IS/IT [20]. The method is an open, flexible and convenient approach to generate information useful to explain why small firms are or are not taking full advantage of ICTs.

The on-line revision of the city's yellow-page section allowed the identification of the main commercial activities prevailing in the downtown area. This information was used to define the type of small firms to be included in the study: haberdasheries, office & stationery, hardware, apparel and shoe shops, and beauty parlors. Two firms in each of the identified commercial and service activities were randomly selected from the yellow-page list. The selected firms were personally visited, the objectives of the research were presented to the general manager or owner of each firm and a formal request for participation was presented. All contacted firms accepted to provide the required information, and an appointment for interview was arranged.

The number of business units that compose each firm varied between 2-10 subsidiaries. The number of employees ranged from 2-10 in each business units resulting in a total number of employees that did not exceed 50. Then all participating firms are classified either as micro or small companies according to the official classification in Mexico.

The data collection strategy employed a two-phase approach. The initial phase involved a) in-depth personal interviews with CEO/owners and b) the elaboration of supplementary questions to the technology users to determine how IS and IT are used within the firm. The second phase involved additional interviews with CEO and employees to gather information about the characteristics of the technical services used by the firm, the perceptions of CEO/owner about usefulness of IS/IT, the managerial actions deployed to encourage IS/IT use among employees and the identification of mechanisms used to develop internal technical skills. The mechanisms relevant to this work included: training provided by external consultants hired by the company and the participation of the employee in a knowledge network related to IS/IT usage. All interviews were audio recorded, notes taken and transcripts were prepared. These were the inputs for the classification of firms in terms of the available IS/IT infrastructure and use.

IT use is defined as the application of ITs within the firm to facilitate the organization's operational and strategic activities [6]. The utilization of IT in support of all the activities of business is considered an evolving process [18] that begins with efficient internal and external communication via e-mail, continues with the use of websites and e-commerce activities (order placement and payments online), it advances with the integration of e-business activities and culminates with a transformed organization that uses IS/IT as the base for networking with business partners. Following this evolving e-business process, the interview guide included open questions arranged in the following five sections: 1) General information about the firm, 2) Technologies used for communication and information interchange (telephone, fax, e-mail, Internet, and web pages), 3) Technologies used to maintain and administrate customer relations (Customer databases, e-mail and web pages), and 4) Technologies used to increase process efficiency and control business operations (for example On-line banking, Accounting systems, Inventory Management Systems, Order Processing systems, Financial Planning Systems and Payroll Systems). Respondents were asked to describe how available IS/ITs are deployed, who the users are, and what specific tasks are assisted by these technologies.

4. DISCUSSION OF RESULTS

The number of computers in the selected business unit varied from one to nine. Four of the SMEs did not have Internet connection, all participants have implemented an accounting or financial control system and except by one, all have tried to put in operation an automatic inventory system. Five firms have a Web page but only two are taking advantage of existing ITC for e-purchasing and e-commerce.

Four academic experts (the two authors and another two professors) and two IT professionals ranked the participant firms in terms of the advancement of the IS/IT utilization process. The criteria used to assess the firm's relative IT/IS usage were based on three subjects: use of ITC to sustain communication with business partners and customers, the number of information systems in use and the variety of business process facilitated by IT/IS.

The ordinal data were used as input of a non-metric Multidimensional Scaling (MDS) procedure (a group of techniques for spatial representation of data). The ALSCAL procedure in SPSS resulted in a two-dimensional solution with a stress of 1.4% and $R^2 = 0.993$. The spatial plot of the MDS analysis allowed obtaining a graphical representation of the situation of the participant firms in respect to two critical dimensions (the axis of the spatial plot): 1) IS/IC infrastructure (Vertical) and 2) degree of e-business applications (Horizontal). The position of a firm in respect to others reflects their relative advantage on IS/IT usage but the distances among firms are only qualitative indicators of the firm's situation. The following section gives a detailed description of the situation of the firms classified in each quadrant.

1. Disadvantaged firms. The firms in the south-west quadrant have low ICT infrastructure, limited to telephone, and/or fax. Some of the firms do not have Internet connection (Mercería América, Solo Ajuste and Novensa). Four of the participating firms are located in this quadrant (out of 14). Some of the firms in this quadrant introduced information systems –a financial administrative system and an inventory system- but they were actually not used because they were perceived by CEO and/or users as inadequate or too complicated for the company's needs (Novensa clothes shop and Solo Ajuste clothes). Spreadsheets are not even used to organize financial reports and the Word processing systems are only used to prepare letters and memos, and occasionally to print promotional material (Refaccionaria Jaimes). The firm located most to the west, Mercería América, has only one computer with a MS-DOS system which is used to run a very rudimentary and obsolete inventory and accounting system which has not been updated since its implementation ten years ago. The personnel of this firm do not even perform regular transactions on-line such as tax payments or cash transferences.
2. Limited electronic business. The firms in the north-west quadrant have better ICT infrastructure in respect to firms in the south side of the map, but they under-use the infrastructure. Web pages are mainly used only to post information and not to take customers' orders or receive on-line payments (e-commerce) even though these applications are implemented (Mercería San Jorge). At Stilisimo beauty parlor, part of the available computers is used by the customers to navigate the Internet while they are waiting for service. The inventory system to place purchase orders (e-purchasing) and control inventory is operated only by one employee (Stilisimo) and information updated only when there are stock outs. E-mail is only occasionally used even for communication between business units (Refaccionaria Orsen).
3. Deployment of technical infrastructure. The south-east quadrant depicts firms that have adopted a reduced number of IS/IT (only a few computers and basic information systems) but make full use of them to facilitate tasks and control activities. Escorpio shoe shop uses the available inventory system to keep control of purchasing orders and each salesperson uses it to check products' availability and track new arrivals. The accounting and financial system is used to track sales by product, control cash flows and to identify the line of products with better sales. The Intranet of Capa de Ozono facilitates the replenishment and order processing of the various business units while reducing delivery times. To introduce additional e-business applications, the firms in this quadrant require making additional investments in technology infrastructure and information systems, and complementing their main business model to include customers buying on-line. While Revelación en la moda, an apparel store, has created a customer databases and segment customers (by age and shopping frequency); e-mails are sent to frequent customers to inform them about promotions and season's sales.
4. Competitive. Firms in the north-east quadrant have the best IS/IT infrastructure and applications. There are only two firms clearly located –another one is close to 0-horizontal axis- in this quadrant (14% of participants). One of them (Consorcio papelero RIME) uses Internet and e-mail to communicate with major customers and inform them about new products and order status. This stationery shop is the only one that has fully implemented a selling system to final customers (e-commerce) that contributes to the firm's differentiation and market expansion. The Intranet links the stores of the consortium with a distribution center so each shop can put orders directly to avoid stock-outs and reduce replenishment times. Systems at Leny Jose beauty parlor are used to prepare stylists' schedules and maintain an appointment agenda, to control cash flows via an accounting and financial system using spreadsheets, to administrate the inventory and to keep a record of customer's visits and purchases to assign points in exchange for courtesy services.

The arrangement of firms in the map permits the identification of a high variability among small firms in respect to IS/IT use. Most of the small retailers and service firms require making additional investments on ICTs and advancing e-business applications. Even the more disadvantaged firms have implemented accounting/financial and inventory systems, but they are not fully utilized (Novensa and Solo Ajuste), or need to be updated or restructured according to the actual business needs (Mercería América).

When exploring the reasons for the deficient infrastructure and usage of IS/ITC, perceived complexity of the technology –related to poor technical skills-, inappropriate technical backing and lack of awareness of the benefits related to automation of business operations were cited as the main problems. Cost of technology was mentioned only by the less advanced and smallest firms (Mercería "América", Solo Ajuste and ABC). Factors explaining the differences among quadrants were organized in three categories: CEO's attitudes and support; knowledge interchanges favoring IT/IS usage and the availability of suitable technical support. A detailed discussion follows.

1. Disadvantaged. The CEO/owners of these firms do not perceive the need for updated systems or hardware. They consider technical consultants and software vendors are usually not familiarized with SME's characteristics and therefore, they have implemented non customized systems or no systems at all. The CEO's negative attitude and lack of support inhibits IT/IS usage among employees, in consequence they are uninterested about improving their technical skills. IT/IS are considered too complex and unnecessary to perform job tasks.
2. Limited electronic business. Most of the employees of the firms in this quadrant are familiarized with the use of the available systems and Internet because there are a good number of computers available that they use after work hours (Mercería San Jorge). All CEO/owners complain about standardized commercial IS because they perceive they are designed according to the needs of larger firms and can only be operated by technical experts. The technical backup offered by current technical consultants and software vendors is also perceived inappropriate, and one of the firms (Stilisimo) reported previous bad experiences. Employees are uninformed about how to use available information technologies to improve their work tasks but some consider they could be valuable to develop a differentiation strategy focused on service quality and customer relationship (Stilisimo and ABC stationery).
3. Deployment of technical infrastructure. Two of the firms allocated to this south-east quadrant -Escorpio and Capa de Ozono- have established continuous and collaborative relations with technical consultants to put into practice additional applications. CEO/owners of the firms in this cluster also have positive perceptions about IS/IT usefulness. They have contracted consulting services not just to introduce a standardized system but to design or adjust the system to the particular firm's operations. Employees have found advantages from IT/IS usage (better control of appointments and improved customer service) and are willing to collaborate in order to identify new applications for current IS systems. They rely on mentoring and learning by experience to upgrade their technical capabilities. CEO supports these activities and invites employees to interact with consultants, but he (she) does not interchange knowledge or experiences directly with the employees.
4. Competitive. Interestingly, those firms in this quadrant (better infrastructure and good IT/IS usage) have contracted technical services only during the implementation of information systems. No after sales support is regularly provided except by requested maintenance. Both firms argue they have invested enough in current ICTs and personnel training and their systems are sufficient to cover current needs. However CEO/owners have the most favorable perceptions in respect to ICT and IS usefulness. The owner of RIME considers e-commerce has contributed to the firm's differentiation and market expansion. Employees are encouraged but not instructed, to make a more pro-active use of current technical resources (for example creation of a database to track customer's purchases); they recognize the importance of acquiring technical instruction and like to chat about their experiences with the application software. Employees of the other firm in the quadrant, Leny Jose, judged IS as essential to maintain control on products inventory and they considered sales have increased after they started to keep and use customers' service records. All employees are familiar with the existing applications, they collaborate with consultants to solve problems and help less-experienced users to improve their technical skills.

5. CONCLUSIONS

This study shows that most of the small retailers and service firms located in the downtown area of a large city in the central part of Mexico have a limited technical infrastructure and lack the knowledge to implement e-business applications that would help them to increase the efficiency of business processes and customer relationships. Only few firms have embraced e-business and made full use of available technical resources. Most of the CEO/owners perceive current technical support and information systems are inappropriate to assist SME's needs. These negative perceptions, along with a lack of awareness of management and employees about the potential contribution of IT/IS, are identified as the main reasons for the disadvantaged technical position of most of the participating firms.

Another critical factor influencing IT/IS usage among SMEs is the employees' interest to develop technical skills and share knowledge and experiences with new applications. The support of technical consultants was relevant to customize systems and provide information about alternative applications but less important for the development of internal IS/IT competences. Employees prefer to improve their technical capabilities through more informal instruction mechanisms such as mentoring or apprenticeship training. CEO/owners of the most advanced firms care for the development of internal technical competences and identify opportunities from the introduction of Internet-based technologies (for example e-purchases and customer relation management). However, they do not provide direct instruction to employees or contract technical training services on a continuous basis.

Other studies about adoption and use of information technologies and systems among SMEs have used quantitative methods (surveys and statistical modeling) but we found the qualitative approach more convenient to use with managers and workers of Mexican small business. Respondents were more confident talking about experiences with IT, showing the available resources, expanding and revising their answers, and asking the interviewer about the meaning of some technical terms (for example e-purchasing, Intranet, identification standards). However the use of a reduced number of cases in a particular city limits the generalization of results to all small Mexican firms.

6. REFERENCES

[1] M. M. Caldeira and J. M. Ward, "Using resource-based theory to interpret the successful adoption and use of

information systems and technology in manufacturing small and medium-size enterprises." **European Journal of Information Systems**, Vol. 12, 2003, pp. 127-141.

[2] J. G. Cegarra-Navarro and E. A. Martínez-Conesa, "E-business through knowledge management in Spanish telecommunications companies." **International Journal of Manpower**, Vol. 28, No. 3/4, 2007, pp. 298-314.

[3] P. Dembla, P. Palvia, and B. Krishnan, "Understanding the adoption of web-enabled transaction processing by small businesses," **Journal of Electronic Commerce Research**, Vol. 8, No. 1, 2007, pp. 1-17.

[4] K. C. Desouza and Y. Awazu, "Knowledge management at SMEs: five peculiarities." **Journal of Knowledge Management**, Vol. 10, No. 1, 2006, pp. 32-41.

[5] E. Grandon and J. M. Pearson, "Strategic value and adoption of electronic commerce: an empirical study of Chilean small and medium businesses," **Journal of Global Information Technology Management**, Vol. 6, No. 3, 2003, pp. 22-43.

[6] B. Ives and S. L. Jarvenpaa, "Applications of global information technology: key issues for management," **MIS Quarterly**, Vol. 15, No. 1, 1991, pp. 33-50.

[7] D. A. Johnston and L. Wright, "The E-Business Capability Of Small And Medium Sized Firms In International Supply Chains." **Information Systems and E-business Management**, Vol. 2, No. 2/3, 2004, pp. 223-240.

[8] V. Kula and E. Tatoglu, (2003). "An exploratory study of Internet adoption by SMEs in an emerging market economy." *European Business Review*, 15 (5), pp. 324-334.

[9] R. L. La Rovere, "Diffusion of information technologies and changes in the telecommunications sector: The case of Brazilian small- and medium-sized enterprises." **Information Technology & People,** Vol. 11, No. 3, 1998, pp. 194-204.

[10] J. Lee, "Discriminant analysis of technology adoption behavior: a case of Internet technologies in small business," **Journal of Computer Information Systems**, Vol. 44, No. 4, 2004, pp. 57-66.

[11] M. J. Lin, "The research on the measurement of corporate technology capability." **International Journal of Technology management**, Vol. 13, No. 2, 1997, pp. 133-139.

[12] N. O. Ndubisi and M. Jantan, "Evaluating IS usage in Malaysian small and medium-sized firms using the technology acceptance model." **Logistics Information Management,** Vol. 16, No. 6, 2003, pp. 440-450.

[13] M. J. Nieto and Z. Fernández. "The role of information technology in corporate strategy of small and medium enterprises," **Journal of International Entrepreneurship**, Vol. 3, No. 4, 2006, pp. 251-262.

[14] T. C. Powell and A. Dent-McAllef, "Information technology as a competitive advantage: The role of human, business, and technology resources." **Strategic Management Journal**, Vol. 18, No. 5, 1997, pp. 375-405.

[15] E. Ramsey, P. Ibbotson, J. Bell and B. Gray, "E-opportunities of service sector SMEs: an Irish cross-border study," **Journal of Small Business and Enterprise Development**, Vol. 10, No. 3, 2003, pp. 250-265.

[16] A. Riege, "Three-dozen knowledge-sharing barriers managers must consider." **Journal of Knowledge Management,** Vol. 9, No. 3, 2005, pp. 18-33.

[17] K. Tanabe and C. Watanabe, "Sources of small and medium enterprises excellent business performance in a service oriented economy," **Journal of Services Research**, Vol. 5, No. 1, 2005, pp. 5-21.

[18] M. Taylor and A. Murphy, "SMEs and e-business," **Journal of Small Business and Enterprise Development**, Vol. 11, No. 3, 2004, pp. 280-289.

[19] J. Thong, "Resource constrains and information systems implementation in Singaporean small business," **Omega**, Vol. 29, No. 2, 2001, pp. 143-156.

[20] R. K. Yin, **Case Study Research: Design and Methods**. Thousand Oaks: Sage Publications, 1994.

Portfolio Selection on the Romanian Capital Market in the Era of E-Business

Prof. Stelian STANCU PhD
Department of Economic Cybernetics
Academy of Economic Studies, Bucharest, ROMANIA
Address: str. Dumitru Secareanu, no.5, district 2, Bucharest, Romania

and

PhD Candidate Oana Mădălina PREDESCU
Academy of Economic Studies, Bucharest, ROMANIA
Address: str. Matei Basarab, no.108, district 3, Bucharest, Romania

ABSTRACT

The technological development of the last decades has led to the transformation of the world economies. The capital markets have begun their journey towards electronic trading, more and more world exchanges being transformed along with the adoption of electronic trading systems. The Bucharest Stock Exchange uses an electronic trading system, the benefits of this system being described in the present paper. Essential for a functional and efficient capital market are both the trading system and the capital allocation and portfolio selection decisions. Taking into consideration these aspects, our study focuses also on the theoretical and empirical analysis of the models considered to be the foundation of modern portfolio theory: the Markowitz portfolio selection model and the market model. We conclude that the model developed by Markowitz leads to the best results.

Keywords: Electronic Trading, Capital Market, Return, Risk, Security and Efficient Portfolio.

1. INTRODUCTION

Technology has always represented a critical element in the functioning of capital markets. One of the reasons why technology is considered to be essential in any domain, especially for capital markets is the fact that along with technological developments, information dissemination is done more rapidly, large volumes of information being easily disseminated, efficiency being improved [3]. Like Garbade and Silber [5] have shown that the introduction of domestic telegraph and transatlantic cable speeded the flow of prices on capital markets allowing quicker transaction execution, the technological development has led to the processing of millions of transactions daily on capital markets, the investors having the possibility to trade on different markets using mobile phones.

Technological development has conducted to the *birth* of *Information Age*, the access to information through personal computers or the Internet being much improved. The computational power is now applied to capital markets that have the capacity to react in milliseconds to different events, and large information volumes can be easily stored for their later manipulation [4]. The technological revolution has influenced national capital markets and world economy, technological innovations connecting markets electronically. In the last years the access to information has become essential both for individuals, but also for companies, economies and governments, efficient information manipulation generating economic development and productivity. Capital markets are nowadays revolutionized by *Information Age*, automated trade execution becoming essential for world exchanges, the old floor-trading model being left behind. The digital revolution has permitted capital markets to access, store and process large volumes of information along with automated trading, generating competition on these markets and in the same time stimulating innovation and research.

The Bucharest Stock Exchange uses an electronic trading system that connects all brokerage companies to the central systems of the Exchange. The trading system of the Bucharest Stock Exchange is entitled Horizon [13] and it is an electronic auction system with a central role on the national capital market.

The capital market represents a vital component for an economy, the capital allocation and portfolio selection decisions being essential elements in the optimization process of financial placements.

In the 50's professor Harry Markowitz, winner of the Nobel Prize for Economy (1990), has developed a brilliant approach of investments, recently known as Modern Portfolio Theory. Markowitz's PhD Candidate, William F. Sharpe, also a winner of the Nobel Prize for Economy (1990), has tried to bring modifications to the fundamental portfolio selection model and supplemental information regarding portfolio diversification. The papers of the two researchers represent the foundation of numerous other studies related to the portfolio management topic.

All these aspects have led to the formalization of the present study. Section 2 of this paper focuses on the evolution of the Romanian Capital Market and underlines the characteristics of the actual trading system of the Bucharest Stock Exchange. Taking into consideration the transformations brought by the technological revolution and the global growing inclination to adopt automated trading systems we considered relevant to describe the features of the Bucharest Stock Exchange trading system.

Section 3 focuses on a theoretical and empirical analysis of the models considered to be the core of Modern Portfolio Theory: the Markowitz portfolio selection model [6] and the market model developed by Sharpe [8,9].

We conclude resuming the results of our analysis, being emphasized the fact that the models tested lead to the optimization of financial placements on the Romanian Capital Market [11].

2. THE BUCHAREST STOCK EXCHANGE AND ITS TRADING SYSTEM

The History of the Bucharest Stock Exchange

On the background of a tradition beyond 70 years of commodities-trade exchanges which started to function in the first years of the 18th century according to the model of western exchanges, the history of Romania records the establishment of the first Stock Exchange in 1881 following the adoption of the ″Law on exchanges, exchange and commodities brokers″. Based on a French model, this law stipulated by royal decree the establishment of commodities and stock exchanges. The Bucharest Stock Exchange was inaugurated on December 1st 1882. So, the evolution of the regulatory framework of the capital market was marked by several stages. In 1904 a new law appeared defining more clearly Stock market operations. In 1929 the ″Madgearu Law″ appeared bringing the legislative unification of the Stock Exchange System in Romania and a modern approach of the legislative act. After the 1929-1933 crisis the economic reconstruction put the Bucharest Stock Exchange on an ascending trend, 1939 being the climax year of the inter-war period. The end of World War II left the political framework in Romania unbalanced leading to the end of the capital market and Stock Exchange. The downfall of this extremely dynamic economic sector was caused by the 1948 nationalization process, the establishment of state property leading to the disappearance of specific products like: stocks, corporate bonds, domestic and external T-bills. The 1989 Revolution, which signified an important turning point in our national history, has imperatively imposed, through the resulting reform program, the necessity to recover the capital market and its related institutions, including the Bucharest Stock Exchange [12].

The reconstruction process started in 1992 was a difficult process under every aspect, taking two years to adopt the Law no. 5/1994 regarding securities and stock exchanges. On April 1st 1995, the Bucharest Stock Exchange was reestablished based on a decision of the National Securities Commission as a self-financed public interest institution with juridical personality. This moment represents a turning moment in the development of the Romanian Capital Market, marking the beginning of the stock exchange and altogether the flourishment of the efforts made to create one of the most important institutions of the capital market. The official inauguration took place on June 23rd 1995, the structure of the new stock exchange being based on a modern electronic system capable to cover trading, clearing-settlement and registry functions and sustained by an ensemble of coherent regulations and procedures. The first trading session was held on the 20th of November 1995.

The year 1999, represents a very important year for the Bucharest Stock Exchange because it marks the implementation of a new electronic trading system entitled Horizon.

The Trading System of the Bucharest Stock Exchange

Horizon [13] is a powerful and flexible trading system, with a high degree of accessibility through open and secure communication channels that have the capacity to evolve along with the changing necessities of the business environment. This system, besides accepting distant orders or within the exchange, can make connections with other exchange systems and can record also floor trades. The main characteristics of this trading system are:

- The market can be defined through trading rules and particular control parameters;
- Securities can be traded and settled in different currencies;
- Orders are sent as messages to a high speed trading engine. This engine stores market orders with the purpose to maximize working speed and trade execution. Trade results and statistical messages are recorded in a relational data base needed for historical trades. These results are sent in real time to the system for immediate information discrimination regarding changes emerged on the market.

The trading system is developed through an advanced client/server architecture, which allows shared usage, message management and guaranteed exchange between different components. This architecture also allows trade processing in a configured and distributed environment. Due to the graphical interface, the trading terminal provides current information about the market allows brokers to view reports regarding orders and executed trades and offers facilities for order input and order management. The system offers reports regarding input, modification or deletion of historical orders and trade execution, as well as analysis and control instruments that detect and investigate problems surged during trading activity. For each order and trade, detailed information regarding input and trade execution, as well as each modification or cancellation, are registered in a system journal. Information comprised in the system journal can be used to retrace the situation on the market at a certain moment. All queries regarding market status and regarding orders are available in replay environment. The system can be configured to automatically block trades that surpass predetermined price or volume limits, trading being temporarily suspended for the security in question until the situation is remediated. It is also possible to automatically suspend trading for all the securities comprised in an index, if the respective index varies with a predetermined percent. It can be decided to delay opening, to stop or suspend the trade of certain securities, as well as to cancel orders or trades.

In comparison to the former trading system, Horizon presents the following advantages:

- Friendly graphical interface, which allows the simultaneous opening of a large number of windows and displaying a variety of information;
- Real time automated updating of information regarding orders and trades on user screens, without additional efforts from users;
- Simultaneous functioning of various types of markets, with different characteristics, rules and schedules;
- Work station configuring by user;
- Local alerts set by users that point out when certain parameters have been reached or surpassed;
- Market level alerts for all users set by the Exchange that point out when certain parameters have been reached or surpassed;
- System flexibility that allows the modification of functioning parameters and trading rules through configuration alteration of certain parameters;
- The existence of a large number of orders: limit orders, market orders, without price, hidden orders, cross, hit, take, all or none, minimum fill, minimum block, contingent orders (that activate automatically on the market when the price of the stock reaches a certain value).

The soft of this trading system is built on the trading rules of international exchanges and has a modern technology. Its modular design allows the integration of different modules according to requirements, the system evolving along with the exchange. The system ensures the possibility to build and trade new products, input new trade practices, improve market control and react rapidly to changing regulations.

3. PORTFOLIO SELECTION ON THE ROMANIAN CAPITAL MARKET

The Markowitz Portfolio Selection Model

Harry Markowitz is rightfully regarded to be the founder of Modern Portfolio Theory. In 1952 he published a formal portfolio selection model [1] emphasizing diversification principles and identifying the *efficient set for portfolios* or the *efficient frontier* for risky securities, paving his way to the Nobel Prize for Economy (1990).

Eugene F. Fama refers to the model developed by Markowitz as the *Two-Parameter Portfolio Model* [2]. In this two-parameter model the distributions of returns on any portfolio are normal and can be described from knowledge of portfolio mean and portfolio standard deviation. So, with normal return distributions, a risk-averse investor only considers a portfolio if it has the largest possible expected return being given the standard deviation of return, and if it has the smallest possible standard deviation of return being given its expected return. The portfolio with these two properties is considered to be *efficient*, and the collection of portfolios with these two properties is called *efficient set*. These two portfolio properties can be summarized into a single one which states that for a portfolio to be efficient there must not be another portfolio with the same or higher expected return that has lower standard deviation of return.

So, in a two-parameter world, the portfolio risk is measured by its standard deviation, or its variance of return. A risk-averse investor is averse to dispersion of portfolio return, dispersion that can be completely summarized by variance [2,6] in the case of normally distributed portfolio returns. The risk of a security is determined by the contribution of the security to the variance of the return on the portfolio.

In formal terms, the return, expected return, and variance of return on a portfolio p are:

$$R_p = \sum_{i=1}^{n} x_{ip} R_i \tag{1}$$

$$E(R_p) = \sum_{i=1}^{n} x_{ip} E(R_i) \tag{2}$$

$$\sigma_p^2 = \sum_{i=1}^{n} x_{ip}^2 \sigma^2(R_i) + \sum_{\substack{i=1 \\ j \neq i}}^{n} x_{ip} x_{jp} \sigma_{ij} \tag{3}$$

but $\sigma^2(R_i) = \sigma_{ii}$, transforming Eq. (3) into the following equation:

$$\sigma_p^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ip} x_{jp} \sigma_{ij} \tag{4}$$

where n is the number of securities available for inclusion in the portfolio, x_{ip} is the proportion of portfolio funds invested in security i in portfolio p, R_i represents the return on security i, $E(R_i)$ is the expected return on security i, $\sigma_{ij} = \text{cov}(R_i, R_j)$ is the covariance between returns on securities i and j and $\sigma^2(R_i) = \sigma_{ii}$ is the variance of the return on security i.

Taking into consideration the above equations, there will be n expected return estimates for each of the $i = 1,...,n$ securities, n variance estimates and $\frac{n(n-1)}{2}$ different covariance estimates [1,10].

The contribution of security i to portfolio return $E(R_p)$ is $x_{ip}E(R_i)$, meaning the expected return of security i is weighted with the proportion of portfolio funds invested in security i to obtain portfolio p.

Rewriting Eq. (4) as:

$$\sigma^2(R_p) = \sum_{i=1}^{n} x_{ip} \left(\sum_{j=1}^{n} x_{jp} \sigma_{ij} \right) = \sum_{i=1}^{n} x_{ip} \,\text{cov}(R_i, R_p) \tag{5}$$

the contribution of security i to the risk or variance of the return on p is

$$x_{ip} \left(\sum_{j=1}^{n} x_{jp} \sigma_{ij} \right) = x_{ip} \,\text{cov}(R_i, R_p) \tag{6}$$

So, Eq. (6) can be interpreted as the risk of security i in portfolio p. However, it is more convenient to call the weighted average of covariances

$$\sum_{j=1}^{n} x_{jp} \sigma_{ij} = \text{cov}(R_i, R_p) \tag{7}$$

the risk of security i in p. If Eq. (7) is interpreted as the risk of security i in p, then from Eq. (5) the risk of the portfolio is the weighted average of the risks of individual securities.

For a portfolio to be efficient it must have the property that no other portfolio with the same or higher expected return has lower standard deviation of return. So, if a portfolio is efficient then (a) it has the maximum possible expected return given the variance of its return, and (b) it has the smallest possible variance of return given its expected return. Any portfolio that satisfies condition (b) is called a *Minimum Variance Portfolio* and it is considered to be the solution of the following optimization problem:

$$\underset{\substack{x_{ip} \\ i=1,...,n}}{Min} \ \sigma^2(R_p) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ip} x_{jp} \sigma_{ij} \tag{8a}$$

with the following restrictions:

$$\sum_{i=1}^{n} x_{ip} E(R_i) = E(R_e) \tag{8b}$$

$$\sum_{i=1}^{n} x_{ip} = 1 \tag{8c}$$

$$x_{ip} \geq 0, i = 1,...,n \tag{8d}$$

$E(R_e)$ is a given level of the expected return and Eq. (8d) emphasizes the fact that the weights are nonnegative, meaning that short sales are not allowed. The problem underlined through Eq. (8a) to Eq. (8d) is to choose the proportions $x_{ip}, i = 1,...,n$ invested in individual securities that give a *minimum variance portfolio*, taking into account the restrictions that state the expected return of the portfolio is equal to $E(R_e)$, the sum of the proportions invested in individual securities is 1, and the proportions must be nonnegative.

The contribution of the pair wise covariances between the returns of individual securities in a portfolio to the variance of the portfolio's return does not change in any systematic way as the number of securities in the portfolio is increased. Instead, the contribution of security return variances to the variance of the return on the portfolio decreases as the number of securities in the portfolio increases. So, with the pair wise covariances that have a stable contribution on the variance of portfolio return, and with the contribution of security return variances declining as the number of securities in the portfolio is increased, the concrete result is a decrease in the variance of the portfolio return. In a diversified portfolio, the contribution of an individual security to the variance of the portfolio's return depends mostly on the security's covariances with the returns on other securities than on the security's return [2].

Rubinstein [7] appreciated that Markowitz's research represents the first mathematical formalization of the diversification concept of investments, emphasizing the fact that even though diversification reduces risk, it can not eliminate it completely. So, through diversification risk can be reduced without having any effects on the portfolio expected return. Rubinstein also underlined the fact that one of the main achievements of Markowitz's research was to demonstrate that not the security risk is important for the investor, but the contribution of the security to the variance of the entire portfolio. Thus, a security must be evaluated as part of a whole and not individually.

Testing the Markowitz Portfolio Selection Model on the Romanian Capital Market

In order to test the Markowitz portfolio selection model on the Romanian Capital Market we have chosen a sample of 15 stocks quoted on the first category of the Bucharest Stock Exchange (BSE). We have chosen these stocks because they meet the conditions of financial performance, liquidity and supplemental dimension in comparison to the rest of the stocks quoted on the exchange rate. Thus, we have computed a data base containing daily returns for these stocks for the period 08/01/2002-24/12/2009. The calculation formula used was $R_{it} = \frac{p_{it} - p_{i,t-1}}{p_{i,t-1}}$ (dividends were not included).

Symbol	Company Name	Business Sector
ALR	Alro Slatina	Materials
ATB	Antibiotice Iasi	Pharmaceuticals
AZO	Azomures Tg. Mures	Fertilizers
BRD	BRD Societe Generale	Banks, Insurances and Financial Services
IMP	Impact Developer & Contractor S.A.	Equipments
OIL	Oil Terminal Constanta	Services and Energetic Equipments
OLT	Oltchim Rm.Valcea	Chemistry
SIF1	SIF Banat Crisana	Banks, Insurances and Financial Services
SIF2	SIF Moldova	Banks, Insurances and Financial Services
SIF3	SIF Transilvania	Banks, Insurances and Financial Services
SIF	SIF Muntenia	Banks, Insurances and Financial Services
SIF5	SIF Oltenia	Banks, Insurances and Financial Services
SNP	Petrom	Energetic Resources
TBM	Turbomecanica Bucuresti	Equipments
TLV	Banca Transilvania	Banks, Insurances and Financial Services

Table1. First Category Stocks quoted on the BSE

The optimization problem emphasized through Eq. (8a) to Eq. (8d) was solved with the help of *Microsoft Excel Solver* (the initial portfolio was equally weighted $x_{ip} = 1/15$). For a given expected return of 0.21% the *Minimum Variance Portfolio* was obtained with a risk of 4.19%. The resulting portfolio weights are: 14.02% ALR, 2.79% ATB, 42.17% OIL, 5.12% OLT, 2.08% SNP, 9.76% TBM and 24.06% TLV.

To generate the *efficient frontier* with the help of *Microsoft Excel Solver*, we have changed the expected return in the optimization problem for 23 different values of the expected return comprised between 0.11% - 0.33%.

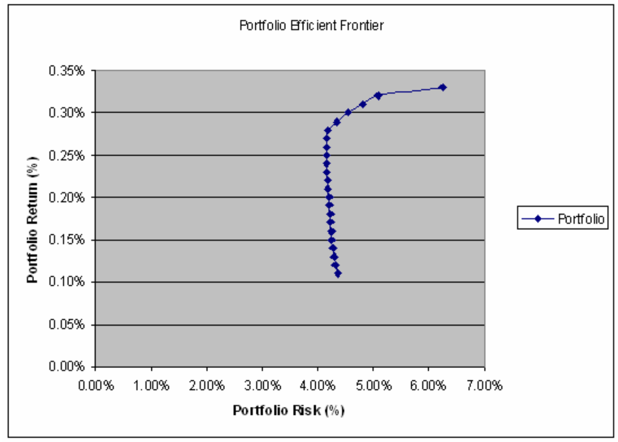


Figure1. Portfolio Efficient Frontier

The Market Model

Markowitz described in an extended footnote the diagonal model or market model and following Markowitz's suggestion this model was to be developed later by William F. Sharpe [7].

The portfolio selection process developed by Markowitz includes three stages: (1) probabilistic estimations of future security performances, (2) analysis of these estimations in order to determine an *efficient set* of portfolios and (3) selecting from that set of portfolios the one that suits best the preferences of the investor. Starting from this process, Sharpe (1963) [8] extended Markowitz's analysis concentrating especially on the second stage – *portfolio analysis*. He considered that the portfolio analysis needed a large number of comparisons and thus the development of a simplified set of hypotheses was necessary in order to reduce the difficulty of this task. Such a set, called the *diagonal model*, highlighted the fact that security returns can be influenced by a single essential factor like the level of a certain index or a factor that is considered to have a significant influence on security returns. Several years later, Sharpe (1967) [9] developed a paper that presented a linear programming algorithm for portfolio selection. In this paper, he emphasized again the importance of the relationship between security returns and a market index and divided portfolio risk into two components. One component includes parts of security risk that can not be eliminated through diversification and the second part includes the risk that can be minimized through diversification.

Fama [2] started from the bivariate normality of securities and portfolio returns and developed a theoretical and empirical study of the market model.

Considering R_i the return on any security $i = 1,...,n$ and R_m the return on the market portfolio of all securities, where each security is given an equal weight in the portfolio, the relationship between R_i and R_m can be expressed as:

$$R_i = \alpha_i + \beta_i R_m + e_i \quad (9)$$

Eq. (9) highlights the fact that there is a linear relationship between R_i and R_m with the coefficients α_i and β_i. This linear relationship is subjected to the disturbance e_i that has a normal distribution with zero mean and constant variance $\sigma^2(e_i)$. This disturbance is independent of the return on the market portfolio R_m.

The variance of the return on security $i = 1,...,n$ can be expressed as:

$$\sigma^2(R_i) = \beta_i^2 \sigma^2(R_m) + \sigma^2(e_i) \quad (10)$$

Eq. (10) divides the variance of the return on security i into two parts: the first part $\beta_i^2\sigma^2(R_m)$ is attributed to the term $\beta_i R_m$ in Eq. (9) and is assumed to be caused by market wide variables, and the second part $\sigma^2(e_i)$ is attributed to the disturbance e_i in Eq. (9) and is assumed to be caused by variables specific to the characteristics of security i. The β_i coefficient is interpreted as the market sensitivity of the return on security i. So, β_i summarizes the sensitivity of R_i to market wide factors, a value of β_i greater than 1 implying a security with both above average market sensitivity and above average risk, and a value of β_i below 1 symbolizing below average market sensitivity and risk in m.

The market model can be used to describe the relationship between the return on any portfolio p and the return on the market portfolio m:

$$R_p = \alpha_p + \beta_p R_m + e_p \tag{11}$$

where:

$$\alpha_p = E(R_p) - \beta_p E(R_m) \tag{12}$$

$$\beta_p = \frac{\mathrm{cov}(R_p, R_m)}{\sigma^2(R_m)} \tag{13}$$

$$\alpha_p = \sum_{i-1}^{n} x_{ip}\alpha_i \tag{14}$$

$$\beta_p = \sum_{i=1}^{n} x_{ip}\beta_i \tag{15}$$

$$e_p = \sum_{i=1}^{n} x_{ip}e_i \tag{16}$$

The variance of the return on the portfolio is given by:

$$\sigma^2(R_p) = \beta_p^2\sigma^2(R_m) + \sigma^2(e_p) \tag{17}$$

The *systematic risk* component [1] of the portfolio variance, the component that depends on market wide factors is $\beta_p^2\sigma^2(R_m)$ and depends on the sensitivity coefficients of the individual securities. This part of the risk depends on portfolio beta and $\sigma^2(R_m)$ and will persist regardless of the extent of portfolio diversification. So, no matter how many securities are included in the portfolio, their common exposure to the market will be reflected in portfolio *systematic risk* that can not be eliminated through diversification.

In contrast, the *nonsystematic component* of the portfolio variance is $\sigma^2(e_p)$ and it is attributed to firm specific components e_i. Because these components are independent and have zero mean, as more and more securities are added in the portfolio the *nonsystematic risk* will become smaller. This risk is considered to be diversifiable.

For the market model the optimization problem is [9,10]:

$$\underset{\substack{x_{ip} \\ i=1,\dots,n}}{Min}\ \sigma^2(R_p) = \beta_p^2\sigma^2(R_m) + \sigma^2(e_p) \tag{18a}$$

having the following restrictions:

$$\sum_{i=1}^{n} x_{ip}\beta_i = \beta_p \tag{18b}$$

$$\sum_{i=1}^{n} x_{ip}E(R_i) = E(R_e) \tag{18c}$$

$$\sum_{i=1}^{n} x_{ip} = 1 \tag{18d}$$

$$x_{ip} \geq 0, i = 1,\dots,n \tag{18e}$$

$E(R_e)$ is a given level for the expected return, and Eq. (18e) highlights the fact that the weights are nonnegative, meaning that short sales are not allowed. The problem underlined through Eq. (18a) to Eq. (18e) is to choose the proportions $x_{ip}, i = 1,\dots,n$ invested in individual securities that lead to a minimized portfolio variance, taking into consideration the restrictions that state the expected return of the portfolio is equal to $E(R_e)$, the sum of the proportions invested in individual securities is 1, and the proportions must be nonnegative.

Testing the Market Model on the Romanian Capital Market

In order to test the market model on the Romanian Capital Market we used the same sample employed for the application of the Markowitz portfolio selection model.

We considered the macroeconomic factor of the model the returns on the reference index of the Bucharest Stock Exchange – BET (Bucharest Exchange Trading).

Using the *Regression* option of *Microsoft Excel* we have estimated the characteristic lines for the stocks comprised in the sample and concluded that BRD, SIF1, SIF2, SIF3, SIF4, SIF5 and SNP have values of the beta coefficient greater than 1, which means that these securities have above average market sensitivity and above average risk, while ALR, ATB, AZO, IMP, OIL, OLT, TBM and TLV have below average market sensitivity and risk.

Comparing the values obtained for the *systematic* and *nonsystematic risk* we have concluded that the securities included in the portfolio are affected in proportion of 0.20% by market wide factors like inflation, variations of the interest rate, recession, that affect all firms simultaneously and only 0.01% of the risk is attributed to firm specific factors, like the success or failure of marketing programs, winning or losing major contracts, as well as other events that take place within companies. Thus, the factors that influence the evolution of the securities included in the portfolio are of systematic nature and can not be eliminated by increasing the number of securities in the portfolio.

The optimization problem emphasized through Eq. (18a) to Eq. (18e) was solved with the help of *Microsoft Excel Solver* (the initial portfolio was equally weighted $x_{ip} = 1/15$). For a given expected return of 0.17%, we obtained the portfolio with a risk of 4.59%. The resulting portfolio weights are: 5.58% ALR, 15.67% ATB, 2.29% BRD, 6.54% IMP, 5.92% OIL, 5.44% OLT, 7.25% SIF1, 4.98% SIF2, 4.24% SIF3, 8.65% SIF4, 5.08% SIF5, 13.28% SNP, 4.79% TBM and 4.95% TLV.

4. CONCLUSIONS

Gorham and Singh [4] considered that the story of any transformation is the story of Schumpeter's *"creative destruction"*, the old systems being destroyed as newer, more transparent and efficient systems take their place. Thus, the old floor-trading model has been replaced in many countries by new electronic trading systems. These trading systems are considered to be more efficient, having the capacity to store and process enormous volumes of information. Electronic trading has led to innovation, competition and collaboration between world exchanges.

In 1999, the Bucharest Stock Exchange has implemented an electronic trading system, the soft of this system being built on the trading rules of international exchanges. Due to its modular design, the system develops along with the exchange, being offered the possibility to develop and trade new products, to introduce new trading practices, to improve market control and to react rapidly to changing regulations.

Also, the capital allocation and portfolio selection decisions represent fundamental elements for a capital market and as a consequence for a national economy. The study conducted in this paper has led to the conclusion that the hypotheses of the tested models lead to the selection of optimal portfolios on the Romanian Capital Market, the best results being obtained in the case of the Markowitz portfolio selection model: for a given expected return of 0.21% the *Minimum Variance Portfolio* has a risk of 4.19%.

5. REFERENCES

[1] Z. Bodie, A.Kane, A.Marcus, **Investments**, The McGraw-Hill Companies, Fifth Edition, 2001, pp. 154-313.

[2] E.F. Fama, **Foundations of Finance**, Basic Books, 1976, pp. 3-132, pp. 212-270.

[3] S.P. Ferris, T.H. McInish, R.A. Wood, "Automated trade execution and trading activity. The case of the Vancouver Stock Exchange", **Journal of International Financial Markets, Institutions and Money**, Vol.1, 1997, pp. 61-72.

[4] M. Gorham, N. Singh, **Electronic Exchanges Transformation from Pits to Bits**, Elsevier Inc., 2009, pp. 1-26.

[5] K.D. Garbade, W.L. Silber, "Technology, Communication and the Performance of Financial Markets: 1840-1975", **The Journal of Finance**, Vol.33, No.3, Papers and Proceedings of the Thirty-Six Annual Meeting American Finance Association, New York City December 28-30, 1977, pp. 819-832.

[6] H. Markowitz, "Portfolio Selection", **The Journal of Finance**, Vol.7, No.1, 1952, pp. 77-91.

[7] M. Rubinstein, "Markowitz's Portfolio Selection: A Fifty- Year Retrospective", **The Journal of Finance**, Vol.57, No.3, 2002, pp. 1041-1045.

[8] W.F. Sharpe, "A Simplified Model for Portfolio Analysis", **Management Science**, Vol. 9, No. 2, 1963, pp. 277-293.

[9] W.F. Sharpe, "A Linear Programming Algorithm for Mutual Fund Portfolio Selection", **Management Science**, Vol. 13, No. 7, 1967, pp. 499-510.

[10] S. Stancu, C. Huidumac, **Teoria Portofoliilor cu aplicatii pe piata financiara**, Ed. Didactica si Pedagogica, 1999, pp. 53-100.

[11] S. Stancu, M. O. Predescu, "The Choice of an Optimal Portfolio on the Romanian Capital Market under Uncertainty and Risk Terms, in the Actual Era of E-Business" The Proceedings of the Ninth International Conference on Informatics in Economy May 2009, Bucharest, Romania, pp. 206-211. Article available at: www.conferenceIE.ase.ro.

[12] ***http://www.bvb.ro;

[13] ***http://www.kmarket.ro;

How Innovative Activity and Competitiveness Relates to Economic Growth of Nations?

Romuald I. ZALEWSKI
Faculty of Commodity Science, Poznań University of Economics, Poland

and

Eulalia SKAWIŃSKA
Chair of Economics, Poznań University of Technology, Poland

ABSTRACT
According to theory, innovative activity gives a chance to increase a competitiveness and economic growth of nation. The purpose of this paper is validation of that assumption using the latest data available for EU countries. Data set of indicators include: global innovation index, (GII), European Summary Innovative Index (SII), Ranking of Competitiveness of Nations (in a form of summary as well as subsidiary data) and set of macro economy data (GDP, labor productivity, export, export of high-tech, R&D expenditure as [as % of GDP] etc as measures of economic growth.
Various regression models: liner, curvilinear, planar or spatial with one or two dependent variables will be calculated and explained. In addition the appropriate 2 D and 3 D-graphs will be used and presented to strengthen verbal arguments and explanation.

The main result of this paper is relationship between innovative activity, competitive ability and growth measured as GDP per capita. Such relationship is shown as fairy good linear span of countries. Only two of them: Luxemburg and Norway due to higher than average growth value are outliers.

The valuable outcome of this paper is classification of nation into groups: highly innovative- highly competitive, highly competitive-non innovative, highly innovative- non competitive and non innovative – non competitive. The last two groups of nations fall in to trap of low competitiveness.

Keywords: incremental innovation, radical innovation, R&D, economic growth, competitiveness, quality of products.

1.Economic growth

GDP, developed in 1930's is the best know measure of macroeconomic activity of nations. Its growth is a key indicator of effectiveness of given economy in short and medium term. It is based on clear and stable over time methodology allowing comparison of countries, regions in time. Unfortunately, GDP does not measure some economical and social phenomena, which importance is growing now, e.g. environmental sustainability or social inclusion. From the other side GDP shows the economic divergence between nations but does not explain why it occurs. Such divergences exist between various counties on different continents. Also in European Union one can find economies in which GDP per capita differs dramatically, from high in Technology Frontier Area –TFA (e.g. Finland, Sweden, Germany) to low in developing ones as Poland, Estonia, Bulgaria. In fact the spread in GDP per capita in Europe is dramatic as a results of political partition after world War II or North-South divergence.

It is commonly accepted by scholars and politicians shat GDP should grow in longer period, despite occurring cycles of economic situation. The economic literature has investigated the drivers of GDP, thus economic growth, for decades. There is a broad dispute and disagreement between scholars concerning the fundamentals of growth; of course there is a substantial agreement in several areas [1,2].

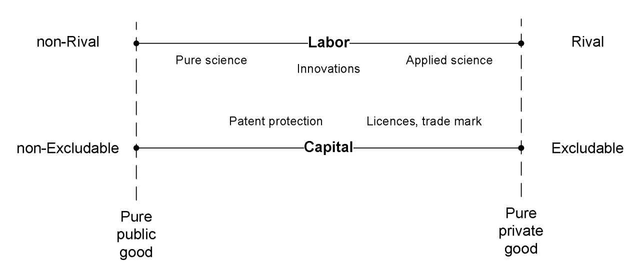
Fundamental components of economic growth in various theoretical models are divided into quantitative and qualitative. The division of production factors in this way reflects divisions of economies into traditional and innovative based on R&D activity, knowledge formulation, education etc. Quantative components are capital K and labor L. At the present among qualitative components we will find technology T, human capital H, institutions IN and economic policy of government P, combined in an equation

$$GDP=F\ (K,L;T,H,\ IN,P)$$

where F means functional relation. Differentiation between 'traditional' or 'quantitative' type of capital and qualitative (T, H, IN, P) could be explained in terms of rivalry and excludability following the paper by Kristian Uppenberg [3]. Fixed capital (K) and labor (L) are "…rival gods, which means that its use by one firm makes it impossible for other firms to use it at the same time. It is also excludable, since an owner of a piece of machinery can prevent other from using it".

By contrast "qualitative" components are typically not rival goods and not necessarily excludable but rather public goods. Non-rivalness means that using for example knowledge by one firm does not diminish the ability of other firm to use the same knowledge. Non-excludability means that one user of knowledge could not prevent other people or organizations from using it once it exists. Pure public goods and pure private goods are opposite points on continuous benchmark scale. In between one can find not pure public goods, partially rival goods or partly excludable goods. The position of some examples is shown in figure 1.

Figure 1. Differentiation between pure and private good



According to Gomułka [4], economic growth in TFA countries rely on R&D activity, knowledge creation, innovations, high quality HR which are various aspects of knowledge economy. At the some time the growth in developing countries to much extent depend on much more passive consumption of technology (diffusion). Thus growth of GDP is a function of capital and improving performance of labor caused by various components of technical (R&D, knowledge, innovation) and social nature (H, IN, P).

Knowledge, as important factor of economic growth was pull out from " the shadow" by Robert Solow [5] and T. Swan [6] and used as starting point for modern empirical analysis of macroeconomic data and theoretical work. New knowledge is the output of investment in the form of spending on R&D, basic research, education etc. The similar relation exist between other investments, made by government and entities consisting of infrastructure, telecommunication, various networks, equipment etc. An investments tends increasing future growth. Unfortunately business investment rate dawn to 20,4 % in the euro area and 20,3 % in EU27 area in the fourth quarter of 2009 [7]. It is obvious (that devoting part of disposable resources to investments helps increasing future output e.g. economic growth. The problem arise for politicians how to divide those recourses between various types of competing investments. For example increasing stock of resources to knowledge creating activity will decrease amount of money in other areas (e.g. current consumption or medical care). For that reason it is difficult for many EU countries to increase expenditures on R&D to Lisbon target (3% GDP), or even EU average value (approx. 2% GDP) relying only on budgetary funds. Fortunately knowledge is not perfectly pure public good as being partly rival and/or partly excludable. In consequence, substantial stock of knowledge and R&D activity is generated by private business funds.

Table 1. Gross Domestic Product expenditure on R&D (GERD) by government sector in 2008 (% of GDP)

No.	Country	% GDP	No.	Country	% GDP
1	JP	15,6	8	Czech R.	41.3
2	Finland	21,8	9	Hungary	41.8
3	Belgium	22.2	10	Estonia	50.0
4	US	27.0	11	Bulgaria	56,7
5	Germany	27.7	12	Poland	59.8
6	EU27	33.5	13	Romanic	70.1
7	France	39.4			

Source: http://epp.eurostat.ec.europa.eu

There is a very big spread of data from 15,6% in Japan to more than 70% in Romania. In developing countries of EU the participatiuon of government in total expenditures on R&D is high (more than 40%). It means that share of business R&D is small, specially in Romania and Poland.

Easterly and Levine [8] argued that Total Factor Productivity (TFP) accounts for majority income differences across countries and not the capital and labor. (for example Taiwan, Ireland). It means that knowledge and innovations are factors increasing economic growth and differentiating between countries. Differentiation results from flow or diffusion (spillover) of knowledge between countries (through the borders). The knowledge spillovers is important over time and depend on many factors. Among them one can mention international trade and foreign direct investment growing dramatically new. FDI disseminate knowledge, new technology, licences etc mostly from richer countries to poorer ones. However reach countries benefit more as they have more and better institutions and policies needed to benefit from knowledge dissemination. From the other side some pure countries are very smart in initiation of innovations, coping technology. Other important factor is human capital, even more important for developing countries [9] than investing in R&D and knowledge formation. Human capital is necessity for absorption of technology and innovation. The well known examples are Ireland and Finland. Years ego politicians decide to develop education at all levels for creation of highly skilled workers in modern branches of science and technology driven industries (ICT, electronics, mobile phones etc.)

There are in TFA different (various) partitions of capital and labor resources into traditional sectors (K and N, respectively) and innovative sectors (M and R, respectively) and innovative sectors (M and R, respectively). The total disposable capital is K+M and labor N+R. If M/K and R/N follow stable growth trajectory however in the period of 'technological revolution' those quotients are much higher. As a consequence growth quotient Y/L increases more dynamically and depends on share of investment in GDP. These quotients are high in TFA countries and promote technology flow towards developing countries. To conclude: the prosperity of nations depend mostly on qualitative components influencing quality of capital and labor and efficiency of their use.

There is no doubt that in many countries of market economy R&D is under-invested, specially in developing countries, national or regional authorities think about wider measures to support R&D of firms directly (by financial means, subsidies, research Programs) and indirectly. The examples are increasing the pool of qualified personnel, searching for talented people, ease immigration of high-skilled people, lowering entry barriers for new-firms, administrative costs, easing credits and venture –capital gains [10]. At the end however the question will be raised: what is the efficiency and effectiveness of public spending on R&D? It is very important issue, because "it is clear that direct and indirect public sector spending on R&D has a positive effect on private R&D spending and on the efficiency of private sector research personal [11]. Empirical research in this area indicated that there is a large potential for increasing efficiency [12]. One can mention the papers by Cincera [13] and Hollanders and Esser [14]. The relationship between inputs and outputs used to calculate Summary Innovation Index for EU countries shows efficiency border and distance of particular nation from this border. One interesting outcome is that some nations spending even less than 1% of GDP on R&D achieve relatively better results than rich due to spillover effects of R&D.

R&D is next factor of great importance for development and productivity growth [15]. The relationship between R&D and economic performance has been demonstrated in various studies. The general outcome is that:

1. return of R&D activity is higher then that of other capital sources.
2. riskness of R&D activity is relatively high,
3. R&D sector is underinvested [16]
4. R&D spending in many countries shows little variability over time (as part of GDP),

R&D activity of nations leads to the development and application of new technology, which in turn yield growth of productivity through: more efficient organization of production, better flexibility, increase of production, lean production, lean innovation etc. One should keep in mind, that radical innovations (product, organizational) will give profit with long lags. Consumers must learn and accept new product its utility, new functions and evaluate quality gains, which took some time. In case of radical organizational innovations, they must interrupt old habits and rules before adopting and learning new ones.

2. Innovative activity of nations (Summary Innovative Index)

> An **innovation** is the implementation of a new or significantly improved product (good or service), or process, a new marketing method, or a new organizational method in business practices, workplace organization or external relations.
>
> **Innovation activities** are all scientific, technological, organizational, financial and commercial steps which actually, or are intended to, lead to the implementation of innovations. Some innovation activities are themselves innovative, others are not novel activities but are necessary for the implementation of innovations. Innovation activities also include R&D that is not directly related to the development of a specific innovation.
>
> (Oslo Manual 2005)
> http://www.ttgv.org.tr/UserFiles/File/OSLO-EN.pdf

The main data source for calculating innovation of countries in Europe is Community Innovation Survey. Aggregated data are disseminated on the Eurostat webpage. The tables cover the basic information of the enterprise, product and process innovation, innovation activity and expenditure, effects of innovation, innovation co-operation, public funding of innovation, source of information for innovation patents, etc.

The European Innovation Scorebord –EIS is a list of countries based on their innovation performance across indicators. The number of indictors increase from more than 20 to 29 in 2009 ranking [17]. Indicators are grouped into five areas: innovative potential, knowledge creation and entrepreneurship (as potential for innovations) as well as applications and property rights as results. Groups of indicators are logically linked to innovative activity and follow the model of process with inputs (potential) and outputs (results) [18]. Innovation performance is calculated as number between 0 and 1. EU27 Member States fall into the following four groups: leaders, followers, moderate innovators and catching-up countries.

3. Competitive ability of nations

The term 'competitiveness' is not identical in meaning. Michael E. Porter in his "Competitive Advantage of Nations" [19] does not define at all what competitiveness is. Some sources explain competitiveness as ability of entity to compete and achieve success.

The "official" definition of OECD of a nation's competitiveness is "the degree to which a country can, under free and fair market conditions, produce goods and services which meet the test of international markets, while simultaneously maintaining and expanding the real incomes of its people over the long term" [20]. Institute of Management Development use the following definition: 'Competitiveness of nations is a field of economic theory, which analyses the facts and policies that shape the ability of a nation to create and maintain an environment that sustains more value creation for its enterprises and more prosperity for its people' (IMD 2003).

Estimation of nations's competitiveness rely on simple or developed measures which reflect the overall results of management and ability of resource's transformation into competitive products and improving the quality of life in long period. Two groups of simple measures are in use. One group describes macroeconomic efficiency through e.g. GDP, rate of inflation or rate of unemployment. Another one deals with international trade through such measures as e.g. share of export in GDP, terms of trade, export structure, share of high-tech products in export, value of export per capita etc.

Unfortunately single measures does not reflect the level of competitive ability of nation. In the literature one can find opinion in favor and against multivariable models. Jeffrey Sachs and his colleagues argued in favor [21]. It is necessary to construct more complex measure composed of variety of simple measures, appropriate weighting factors and calculated according to statistical model. Variety of such holistic models are known in the literature, however two are of great applicability.

World Economic Forum developed 'Growth Competitiveness Index – GCI' on the ground of 12 pillars (institutions, infrastructure, macroeconomic stability, health and education, higher education, efficiency of market (products, labor, finance), adaptation of technology, market volume, management maturity and innovation. Resulting GCI is a rank for given nation.

The similar rank was developed by IMD (Lozanna). The World Competitiveness Yearbook published by IMD focuses on the outcome of the interaction of four competitiveness factors, which generally define a country's national environment. These are:

- Economic Performance (EP),
- Government Efficiency (GE),
- Business Efficiency (BE),
- Infrastructure (I)

On the basis of these four factors and more than 320 criteria (2/3 are hard data and 1/3 are opinion of 4000 managers), the WCY assumes that healthy performance in these dimensions creates a national environment that sustains World Competitiveness Index. The index is published once a year for 60 economies (51 countries and 9 regions). We use in our paper data on overall competitiveness (WCI) of EU countries in 2008 and before the economic crisis. In addition we use data for four mentioned factors (EP, GE, BE and I).

4.**Data and analysis**

Variable set used as a proxy of economic growth, innovativeness and competitiveness of European Union and some additional countries. Most data are from 2008 sources but in some cases we use data from earlier periods. The analysis of time series for innovativeness and competitiveness indicates stability of ranks over last years. All data for growth are from 2008.

Table 2. Variable set used

Variable	Description	Value
Growth		
GDP pc	GDP per capita	in '000
Exp/GDP	Share of export / GDP	USD
Exp pc	Export per capita	% of GDP
R&D G	Share of expenditures of government R&D	in '000 USD
R&D E	Employment in R&D	
EKIS	Employment in knowledge intensive services	in % in % in %

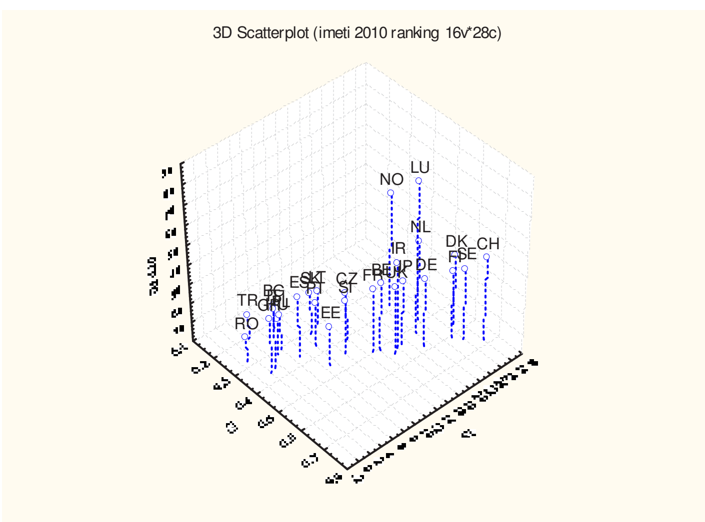
Innovativenes		
SII	Summary Innovative Index	0 - 1
Competitiveness		
CI	Competitiveness index (IMD)	Rank
EC	Economic performance	Rank
GE	Government efficiency	Rank
BE	Business efficiency	Rank
I	Infrastructure	Rank

Growth is approximated by six indicators collected in table 2. Innovative activity for EU and some additional countries is measured by SII as a value from range 0-1. Competitiveness of nations is approximated by a set of variables developed by IMD as a rank. We used general rank CI and four group of ranks for economic performance, government efficiency, business efficiency and infrastructure. All categories: growth, innovative activity and competitiveness rare discussed in sections 1-3 respectively.

For data analysis and presentation we used simple graphical presentation of relations in three dimensional space as previously [23].

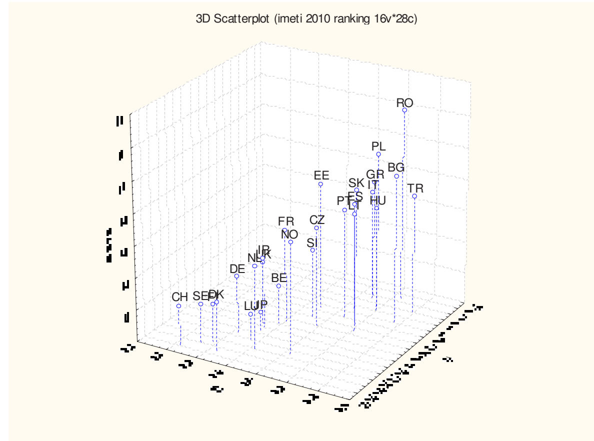
The relationship between innovative activity, competitive ability and growth measured as GDP per capita is a main objective of this paper. Such relationship is show in figure 2 as fairy good span of countries. Only two of them: Luxemburg and Norway due to higher than average growth value are outliers.

Figure 2. Scatterplot of Summary Innovative Index – Competitiveness – GDPpc



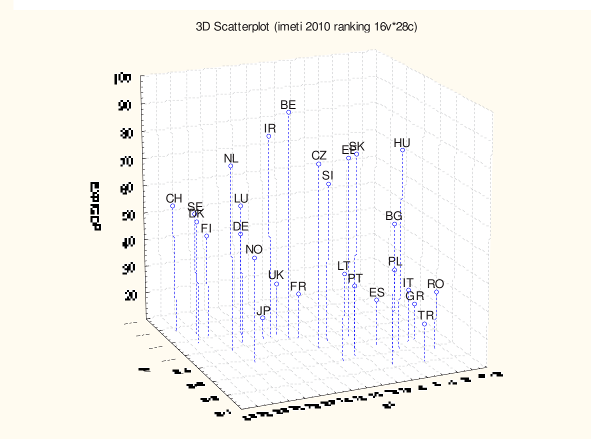
Another interesting plot is between SII-CI and research and development activity financed by government. The plot is presented on figure 3 as fairly good. The high share of government in total cost of R&D (in %) is reflected by high rank, with maximum for Romania. In developed countries the role of government in this type of activity is much weaker. Governments support mainly pure or even 'blue sky' research with a unknown probability of commercial application in future.

Figure 3. Scatterplot of Summary Innovative Index – Competitiveness and R&D G



The spatial distribution of points in the plot: innovative activity- competitive ability and share of export to GDP is shown in figure 4. It give impression of random distribution, however some group of countries could be seen. For example new EU members: Hungary, Czech Republic, Slovakia, Slovenia export a substantial part of their production. Some other countries like Poland, Romania, Latvia reach much worse results. In this same part of plot one can find however Italy, Spain, Portugal and Greece, know as 'PIGS' group, facing a great economic problems now. Remaining countries (highly competitive and highly innovative) differs also against export/GDP share. For example Belgium, Netherlands, Ireland leads, while UK a France are outliers.

Figure 4. Scatterplot of Summary Innovative Index – Competitiveness – Exp/GDP



Proof of such reasoning is given by Principal Component Analysis (PCA) of all data. Two principal components explain more than 77 % of total data variability. The projection of the components on the factor plane is shown in the figure 5A (upper part), while the figure 5B presents projection of countries. Superposition of both will lead to valuable conclusions. It is evident that new member states and some older from the South of Europe realize R&D activity mostly financed by government. Ireland, Belgium, Netherland are very efficient as exporting countries (Exp/GDP), while Greece, Portugal, Romania and even Italy not so much. France and UK are good in economic performance factor of competitiveness (EC). Denmark, Finland and Sweden owe their position to business efficiency (BE), high employment in R&D and in knowledge intensive services (EKIS).

Figure 5A. Projection of the variables on the factor plane

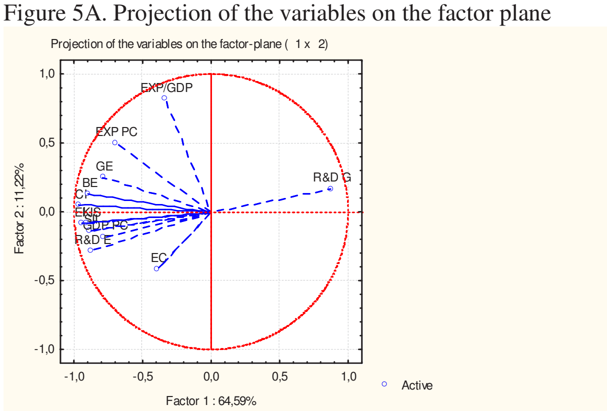
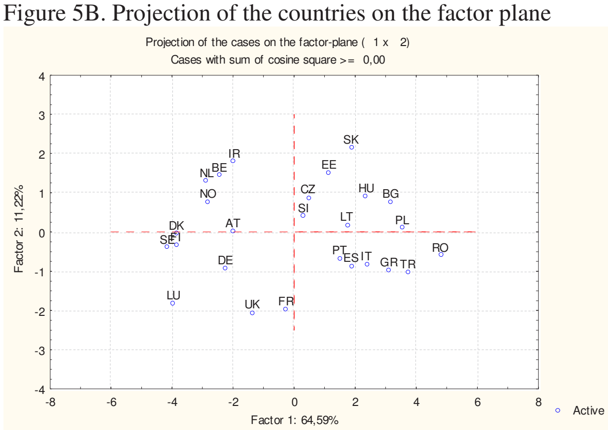


Figure 5B. Projection of the countries on the factor plane



The valuable outcome of this paper is classification of nations into four groups: highly innovative - highly competitive, highly competitive - non innovative, highly innovative - non competitive and non innovative – non competitive. The last group of nations fall into trap.

5.Conclusions

According to theory, innovative activity gives a chance to increase a competitiveness and economic growth of nation. The aim of this paper was validation of that assumption using the latest data available for EU countries. The relationship of Summary Innovative Index – Competitiveness – GDPpc for EU countries is fairly linear with a moderate correlation and support the assumption. The matter is difficult for analysis due to the fact, that are no simple measures for those three dimensions. In fact to various extent all three parameters: SII, CI and GDP overlapped. The further analysis should take this fact into account. Also for better understanding the relations and new trends, the panel of countries must be expanded. Inclusion of Mainland Chaina, India, South Asia and South America economies is necessary.

References

1.Scotchmer S. (2004). **Innovation and Incentives**. MIT Press, Cambridge, Massachusetts, USA.
2.Aghion P., Howitt P., (2009). **The economics of growth**. MIT Press, Cambridge, Massachusetts, USA.
3.Uppenberg K. (2009). Innovation and economic growth, **PIB Papers**, 14(1), 11-35.
4.Gomułka S. (2009). **Mechanizm i źródła wzrostu gospodarczego w świecie**, in: Wzrost gospodarczy w krajach transformacji; konwergencja czy dywergencja, PWE, Warszawa, s. 15-31.
5.Solow R.M. (1956). A contribution to the theory of economic growth. **Quarterly Journal of Economics,** (70), pp. 65-94.
6.Swan T.W. (1956). Economic growth and capital accumulation. **Economic Record**, (32:2), p.334-61.
7. Eurostat newsrelease euroindicators 61/2010, 30 April 2010.
8. Easterly W., Levine R., (2001). It's not factor accumulation: Stylized facts and growth models. **World Bank Economic Review**, (15:2), pp. 221-224.
9.Engelbrecht H.J., (2000). Human capital and international knowledge spillovers in TFP growth of a sample of developing countries: An exploration of alternative approaches. **Applied Economics**, (34), pp.831–841.
10.Mc Morrow K., Röger W. (2009). R&D capital and economic growth: The empirical evidence. **EIB Papers**, (14:1), 95-118.
11.Sveikauskas L. (2007). R&D and productivity growth: A review of the literature. **Bureau of Labour Statistics Working Paper** No. 408.
12.Mandl U., Dierx A., Ilzkovitz F. (2008). The effectiveness and efficiency of public spending. **European Economy Economic Paper** No. 301
13.Cincera M., van Pottelsberghe de la Potterie B., (2001). International R&D spillovers: a survey. **Brussels Economic Review**, (169), pp. 3-32.
14.Hollanders H., Esser F.C. (2007). Measuring innovation efficiency, **INNOMETRIC Paper,** Dec. 2007.
[15]. Los B., Verspagen B. (2007). **Technology spillovers and their impact on productivity**, in Hanusch H., Pyka, A. (eds.), Elgar companion to neo-schumpeterian economics, pp. 574-593, Edward Elgar, Cheltenham, UK.
16. Griffith R., Redding S., van Reenen J. (2004). Mapping the two faces of R&D: productivity growth in a panel of OECD industries. **Review of Economics and Statistics**, (86:4), pp.883-895..
17.Hollanders H. (2009). European Innovation Scoreboard (EIS): Evolution and Lessons Learnt, Innovation Indicators for Latin America Workshop, OECD, 19 March 2009.
18.Skawińska E., Zalewski R.I., **Business Clusters in Regions' Innovation and Competitiveness Development: World – Europe – Poland**, PWE, Warszawa, 2009.
19.Porter M.E. (1990). Competitive Advantage of Nations, McMillan London.
20.Garelli S. http://members.shaw.ca/compilerpress1/Anno%20Garelli%20CN%20Fundamentals.htm
21.Zinnes Z.C., Eilat Y., Sachs J., Benchmarking competitiveness in transition economies, **Economies in Transition**, 9(2), 2001
22.**IMD's World Competitiveness Yearbook**, 2003.
23.Zalewski R.I., Skawińska E., (2009). Impact of technological innovations on economic growth of nations, **Proceedings of 2nd IMETI**, Volume II, pp. 275-280.

Information Technology and its use at TCE-MG

Sebastião CASTRO
Minas Gerais Court of Auditors.
Belo Horizonte, Minas Gerais- 30380-435, Brasil

Armando GRANDIOSO
Minas Gerais Court of Auditors.
Belo Horizonte, Minas Gerais- 30380-435, Brasil

ABSTRACT

This work consists of showing the advantages that Technological Innovation brought to TCE-MG, Minas Gerais Court of Auditors, agency of external control that manages public resources of Minas Gerais State. Monitoring public resources is not an easy task. In order to give a quality service to society and to fulfill with its legal obligation with efficiency, TCE-MG created in its organizational structure, the Information Technology Department (ITD). ITD provides information technology support to the TCE-MG in the form of managing the local network area, maintaining and operating a management information system, providing microcomputer support, and extracting and manipulating automated information for audits. Many of TCE-MG accomplishments such as conducting audit of public agencies accounts with veracity, efficiency and timely as well as working to increase transparency has been improved through ITD support. The ITD at TCE-MG is always looking forward to new technology that could help fulfilling its obligation with ethics, justice, effectiveness, transparency and social commitment.
Keywords: Information Technology, TCE-MG, Audits, SIACE-PCA, Transparency, Government Financial Statements.

INTRODUCTION

Minas Gerais Court of Auditors (TCE-MG) is an agency that has the role of audit the accounts of public agencies as well as a wide range of other public bodies in Minas Gerais state. Its mission is to make this control with efficiency and responsibility. Its goal is to give assurance over many aspects of public revenue and expenditure such as the truth and fairness of public financial statements; the regularity of the expenditure and the fair use of public resource towards society benefit.

The TCE-MG is responsible for analyzing and controlling the revenue and expenditure of 853city councils. In total, TCE-MG has the obligation to monitor and examine 2.292 public agencies, totalizing the amount of 68 billions of reais; approximately 34 billions of dollars.

TCE-MG has been concerned about achieving two goals:

1) To conduct audits of public agencies accounts with veracity, efficiency and timely;
2) To work to increase transparency in public agencies, fighting against corruption and increasing citizens' participation.

TCE-MG X INFORMATION TECHLOGY DEPARTAMENT

In order to improve the accuracy of the information that TCE-MG receives from the public agencies and also to provide society with information and opportunities for participation, TCE-MG has the support from Information Technology Department (ITD) that uses the technology resources to implement TCE-MG external control.

The purpose of ITD is to evaluate the system's internal control design and effectiveness. This includes but is not limited to efficiency and security protocols, development processes, and IT governance or oversight. The goal is to evaluate the organization's ability to protect its information assets and properly dispense information to authorized parties and society. ITD also provides technological tools that allow TCE-MG to receive, audit, analyse, put information available to citizens.

The ITD audit's agenda may be summarized by the following questions:

Will the organization's computer systems be available for business at all times when required? (Availability)

Will the information in the systems be disclosed only to authorized users? (Confidentiality)

Will the information provided by the system always be accurate, reliable, and timely? (Integrity)

Will information held by public agencies be provided to society? ((Transparency)

CONDUCTING AUDITS OF PUBLIC AGENCIES ACCOUNTS WITH VERACITY, EFFICIENCY AND TIMELY

The concept of accountability for use of public resources and government authority is the key to all nations' governing processes. Government officials entrusted with public resources are responsible for carrying out public functions legally, effectively, efficiently, economically, ethically, and equitably[(1)].

According to the article 76 of the Constitution of Minas Gerais, it is clear the obligation that TCE-MG has to control and monitor public resources with effectiveness, transparency and mainly social commitment[(2)].

A study made in 2000 by Marinalva Rita Moreira found out that the frequency of irregularities practiced for the public agencies were large because of many aspects. One of them she described as the necessity of adopting new technology procedures to the effectiveness of TCE-MG work [(3)].

ITD along with TCE-MG staff have developed new technologies that have allowed TCE-MG to gain more agility as to receive all the information as well as to analyze and decided about it. The development of two systems: SGAP and SIACE-PCA [(4)] are examples of ITD contributions.

SGAP : Management System of Accompaniment of Proceeding. (1992): This system is responsible for controlling all the information that arrives at TCE-MG. It also allows all staff involved to have access to each document course whenever necessary.

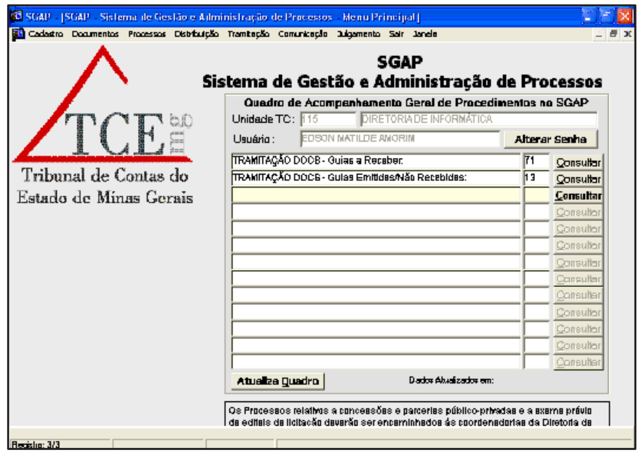


Figure1: accessing SGAP

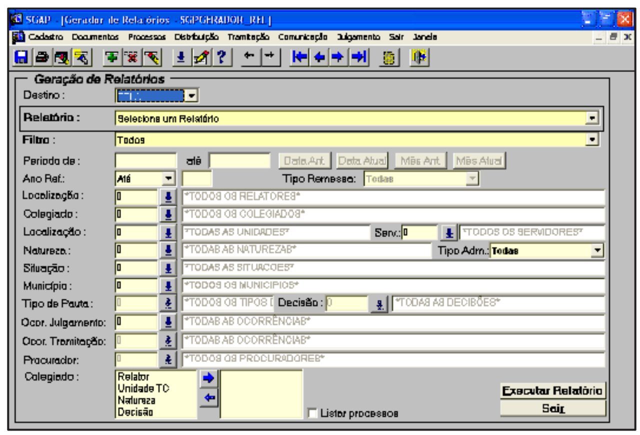


Figure 2: choosing reports

SIACE-PCA :Computerized System that supports External Control (2000): This system receives all the statements regarding to Minas Gerais State, its city councils and others public agencies accounts.

Figure 3: Accessing SIACE-PCA

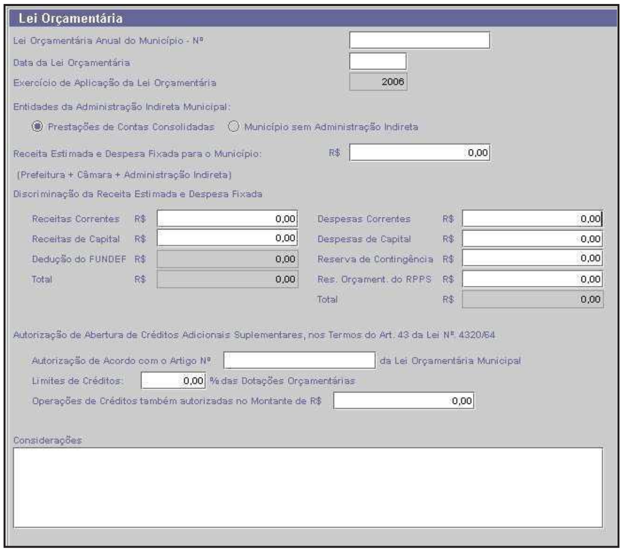


Figure 4: budget information

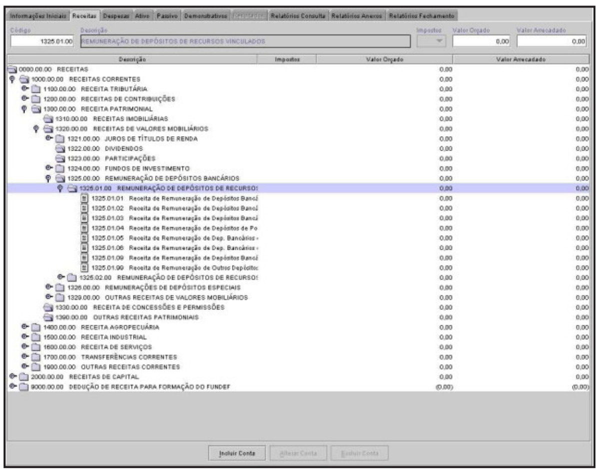


Figure 5: revenue information

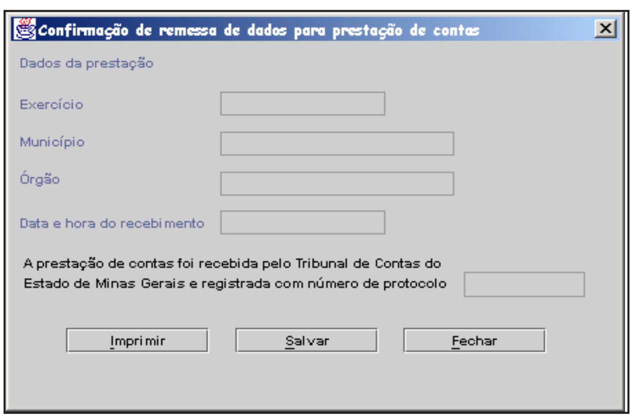


Figure 6: receipt

It is important to mention that all information sent to TCE-MG by public agencies through SIACE-PCA, are important data which are use to planning TCE-MG local examination. Through data processing, TCE-MG receives, audits and analyses the data and, if necessary, checks its veracity personally.

After implanting new programs, redefining the way to analyze the information regarding to accounts from government departments and public agencies, plus computerizing all data, TCE-MG has gained agility and consequently has worked more satisfactorily in society benefit.

The numbers below show the improvements of TCE-MG performance after working along with ID department.

ACCOUNT YEAR	% AUDITED ACCOUNTS AND OPNIONS MADE IN 2010
2008	**96%** (822 city councils)
2002	**46%** (393city councils)

Fonte: SGAP / Total of Minas Gerais city halls: 853

INCRESING TRANSPARENCY

Government that works for its citizens must be open and accountable, use public resources fairly and efficiently, apply the rule of law and provide the public with information and opportunities for participation. Government officials who abuse their power for private gain waste public resources and lose the trust of those they are meant to serve (5).

Many countries have passed legislation that recognizes and protects the right to access information held by public bodies. This is an important component in any effort to fight corruption, but the mere existence of legal instruments is not enough. Habits and cultures on both sides of the information demand and supply relationship must be changed. While public bodies change their attitudes of secrecy to a climate of openness to prevent potentially corrupt situations, citizens also need to capture the spirit of demanding information. . (6)

In May 2000, in order to prevent that mayors and governors spend more than what they obtain collecting through taxes (balance in public accounts) and also to increasing citizens' participation (transparency) Brazilian government approved the

Law of Fiscal Responsibility (Lei de Responsabilidade Fiscal- LRF) [7].

Business intelligence and information management technology can be an enabler for greater government openness and clarity in communication with citizens. Instead of just dumping an agency's raw data on a web site, the information can be organized and tools provided so that external constituents can more easily analyze it. This kind of open access to information will not only increase the public's trust in government but also facilitate non-government research with the goal of identifying, evaluating and prioritizing different solutions to problems challenging public officials[8].

In 27/05/2009, the president of Brazil signed a law named "Capeberibe Law" [9]. According to this law all public agencies have to publish all public accounts (revenue and expenditures) in detailed way in the Internet.

The creation of TCE-MG web page and its constants updates were a major step toward transparency in public accounts. In TCE-MG Web site, society and also public agencies monitored by it have all information regarding to TCE-MG giving services.

Examining with TCE-MG:
Looking forward to work according to Brazilian´s Law and always worrying about transparency issue, in 2004 ITD at TCE-MG launched the first version of Examining with TCE-MG which was reformulated in 2008.
Examining with TCE-MG is an important tool available at TCE-MG web page which allows public to access information regarding to how public agencies are managing public resources. Important information, such as: the revenue and the expenditure in different areas as education and health are easily available to citizens. Examining with TCE-MG also shows if the public agencies are applying the rule of law.

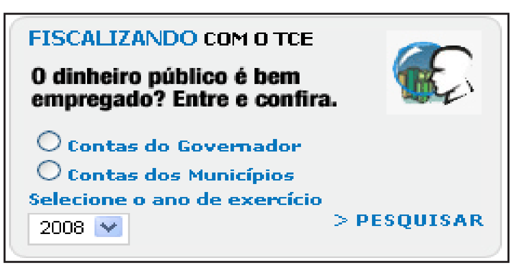


Figure 7: Accessing Examining with TCE-MG

Saúde - Mínimo de 7% para 2000 chegando a 15% em 2004 em diante

Prestação de Contas do Legislativo

Repasse do Executivo	R$ 103.524.315,80
Gasto com Pessoal	R$ 68.668.845,58
Subsídio Anual dos Vereadores	R$ 4.505.250,00
Qtde de Vereadores a partir de 2005 (Resolução 21803 TSE)	41

Fonte: Prestações de Contas Anuais encaminhadas ao TCEMG

Receitas do Município

Receitas Correntes: R$ **4.179.716.767,64**
Receita Tributária: R$ **1.320.680.985,13**

IPTU	R$ 399.305.632,00
Inter-Vivos	R$ 137.919.456,00
ISS	R$ 515.861.376,00
Taxas	R$ 118.575.017,13
Contribuição de Melhoria	R$ 0,00
Demais Receitas Tributárias	R$ 149.019.504,00

Transferências Correntes:R$ **2.414.666.946,06**

Transferência da União	R$ 977.728.965,06
Transferência do Estado	R$ 928.079.778,00
Transferência de Convênios	R$ 55.673.342,18
Demais Transferências Correntes	R$ 453.184.860,82

Demais Receitas Correntes: R$ **397.086.514,67**

Figure 8: information about public resources and its use.

MONTES CLAROS
2008

Prestação de Contas do Executivo

		Indices		
		Apresentados	Apurados	Situação Índices Apurado
Receita Corrente Líquida	R$ 333.225.657,77			
Despesa com Pessoal	R$ 143.169.169,80			
Receita Base de Cálculo	R$ 158.887.034,46			
Despesa com Educação	R$ 40.040.387,27	25,20%	25,07%	
Despesa com Saúde	R$ 29.605.663,66	18,63%	18,62%	
Dívida Consolidada Líquida	R$ 35.011.779,50			

Fonte: Prestações de Contas Anuais encaminhadas ao TC

Figure 9: information about public resources, its use and its conformity with Law and Constitution. (First checking by SIACE-PCA)

Prestação de Contas do Executivo

		Indices		
		Apresentados	Apurados	Situação Índices Apurados
Receita Corrente Líquida	R$ 13.033.257,66			
Despesa com Pessoal	R$ 4.971.664,12			
Receita Base de Cálculo	R$ 9.328.264,62			
Despesa com Educação	R$ 2.363.143,91	25,33%	25,35%	
Despesa com Saúde	R$ 1.774.772,66	19,03%	12,74%	
Dívida Consolidada Líquida	***			

Fonte: Prestações de Contas Anuais encaminhadas ao TCEMG

Figure 10: information about public resources, its use and its conformity with Laws and Constitution. (First checking by SIACE-PCA)

In short, TCE-MG receives information from public agencies through SIACE-PCA. Then SIACE-PCA through itself analyses the data and check if it is according to Brazilian´s Law and Constitution. After this, through audit techniques, TCE-MG analyses and decided about it. It is important to know that TCE-MG put all information available to society at Examining with TCE-MG, updating the information whenever necessary.

Example: Agua Boa - 2007

	INDEX PRESENTED BY MAYOR	INDEX CHECKED BY SIACE	INDEX AFTER TCE-MG AUDIT
EDUCATION	25,86%	22,42%	**23,57%**
HEALTH	16,34%	12,66%	**14,42%**

Legal limits: Education: **25%** - Heath: **15%**

In 2007, the ITD made a research using Google Analytical toll with the purpose of finding out which Web Site area had more access during that year (10). The result of the study is showed below.

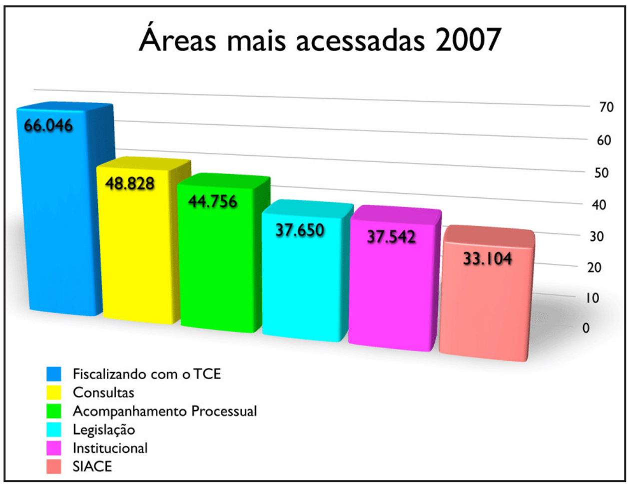


Figure 11: areas accessed

According to the results, it is clear that citizens are more concerned about how public resources are been used. General public are also demanding more information, exercising their right of having access to information held by public bodies. TCE-MG seems to be in the right way when it shows its concerns about transparency issue.

CONCLUSION

The TCE-MG has achieved its goals through Information Technology help. It is impossible to think how all important information that TCE-MG receives could be exchanged, audited, and analyzed with accuracy, veracity and timely without data processing.

SIACE-PCA and SGAP are toll developed by ITD that support TCE-MG achieves its goal. The world nowadays is concerned about transparency in public agency. So, all the benefits achieved by TCE-MG in this field would be harder without technological support.

Analyzing with TCE is an example of TCE-MG concerning about the importance of transparency in public accounts. Citizens and taxpayers deserve to receive all the information regarding to where and how their money has been applied. The population also deserves an efficiency set of data that they can trust from the entity responsible for controlling public resources.

In this context, the role of TCE-MG gains greater importance, given the need to increase transparency and accountability - which are fundamental democratic values - in order to ensure that public funds are soundly managed on behalf of citizens and taxpayers.

Along with ITD, TCE-MG is always looking forward news technology that could help it fulfilling its obligation with ethics, justice, effectiveness, transparency and social commitment.

Nowadays, ITD has updating all the systems used by TCE-MG, installing new programs and revising the current ones. The mainly concern now is to develop a system that allows TCE-MG receives data from public agencies in real time and put this information available to society.

REFERENCES

1)Comptroller General of the United States, **Government Auditing Standards** - July 2007 www.gao.gov

2) Minas Gerais. [Constituição (1989)] **Constituição do Estado de Minas Gerais.** — 13. Ed.— Belo Horizonte: Assembléia Legislativa do Estado de Minas Gerais, 2010.

3) Marinalva, Rita Moreira**. Revista do Tribunal de Contas do Estado de Minas Gerais** - Edição Nº 02 de 2000 - Anos XVIII

4) Convênio TCEMG-DCC/UFMG **SIACE-PCA** : Manual de Instalação e Utilização – Tribunal de Contas do Estado de Minas Gerais .Belo Horizonte, Janeiro /2001

5)Heineman, Ben. **Restoring Trust in Corporate Governance: The Six Essential Tasks of Boards of Directors and Business Leaders**. Policy and Impact Committee of the Committee for Economic Development, January 2010.

6)Riaño,Juanita, Hodess,Robin and Evans, Alastair. **The 2009 Global Corruption Barometer.** www.transparency.org

7)Ministério de Planejamento Orçamento e Gestão.**Lei de Responsabilidade Fiscal –Manual Básico de Treinamento para Municípios**.2ªedição,Brasília2002.

8)Transparency International. **Human Rights and Corruption.** Working paper nº5/2008

9)Diário Oficial da União – **Lei Complementar 131 de 27 de maio de 2009.**

10)Projeto Elaborado pelo Departamento de Tecnologia da Informação e o Departamento de Comunicação Social do TCE-MG; **Novo Portal TCE-MG- 18/11/2008**, Novembro de 2008.

11)Saudagaran, Shahrokh M. **International Accounting: A user Perspective**. Thomson South Western,2003.

System Identification Techniques for Formation Flying Telescope Arrays

Shahin Ghazanshahi
Department of Electrical Engineering, California State University, Fullerton
Fullerton, Ca, 92834, USA

ABSTRACT

Techniques for actively controlling arrays of telescopes flying in formation in space are being studied to control the optical path lengths. Precision control, to a fraction of the wavelength of the incoming light, is necessary to construct ultra-high resolution imagery of stellar disks of nearby stars and for direct imaging of extra-solar planets around these systems. Systems of this type require complex multiple level onboard control techniques to initially phase the telescopes; bringing them in from meters to nanometer and to subsequently hold the telescopes at nanometer tolerances during science observations, with individual spacecraft separations of up to 500 meters. However, the on orbit control system response and some of its parameters are likely to be incompletely known due to differences between ground calibration and the on-orbit environment and also they may drift overtime. Herein we present different system identification techniques potentially feasible for formation flying and optical phasing to determine, both analytically and computationally, the complete in situ closed-loop control system response. The use of system identification techniques to determine the in situ response will enable NASA to further advance the technology of future space-based multiple telescope interferometry.

Keywords: System Identification, Active Optical Control, Formation Flying, Wavefront Sensing.

1. INTRODUCTION

The resolution of stellar disks of nearby stars (Stellar Imager) and direct imaging of extra-solar planets around nearby stars (Planet Imager) will require ultra-high resolution space-based synthetic aperture interferometric imaging systems. Stellar Imager (SI) will resolve the structure of stellar disks of nearby stars and the Terrestrial Planet Finder (TPF) will directly detect and characterize extra-solar planets around nearby stars. Where by the angular resolution is driven by the diameter of the collection aperture, hence the need for large space-borne aperture. Apertures larger than a few 10's of meters can only be realized in space by synthetic aperture interferometric imaging systems. Such large-scale, ultra-high resolution synthetic aperture imaging can, in principle, be accomplished in space by flying an array of phased telescopes each of which comprises a component of a synthetic aperture. Light collected from each of the telescopes are relayed and coherently mixed together to realize imagery with the resolution defined by the diameter of the synthetic aperture as opposed to a single telescope's aperture, albeit with reduced photometric sensitivity. In order to maintain the image quality, the optical path length differences, through the array of telescopes, must be actively controlled to a fraction of the wavelength of the incoming light. A coupled, active optical control system and pointing control system must be used. Techniques for actively controlling such arrays flying in formation are being studied at the NASA Goddard Space Flight Center (GSFC).

GSFC has been studying a multitude of techniques for optimal wavefront sensing and active optimal control, both computationally and through the development of a series of laboratory testbeds for closed-loop control of imaging interferometers. One of these testbeds, Fizeau Interferometry Testbed (FIT), (shown in Figure 1) is being used to explore the requirements for wavefront sensing and for control algorithms for the Stellar Imager and other future Fizeau interferometric systems (e.g., MAXIM and Planet Imager). FIT will be ultimately be used to demonstrate stable closed-loop control of articulated mirrors and the overall system to keep the beams in phase and optimize imaging with thermal and structural perturbations.

Image-based phase retrieval and phase diversity methods [4] are employed on FIT for sensing and solving for the system optical wavefront. One or more images, with varying

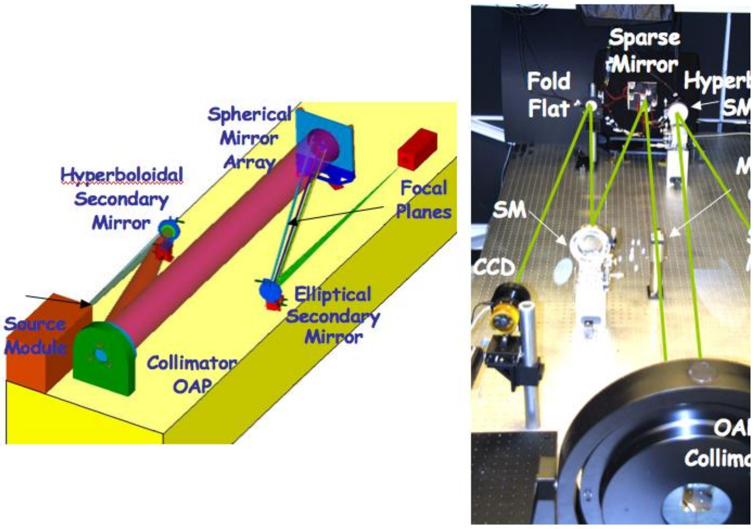


Figure 1 – *Fizeau Interferometry Testbed*

wavelengths and foci, are collected and processed along with knowledge of the system through iterative nonlinear algorithms to recover the optical path differences (wavefront error) between different telescopes comprising

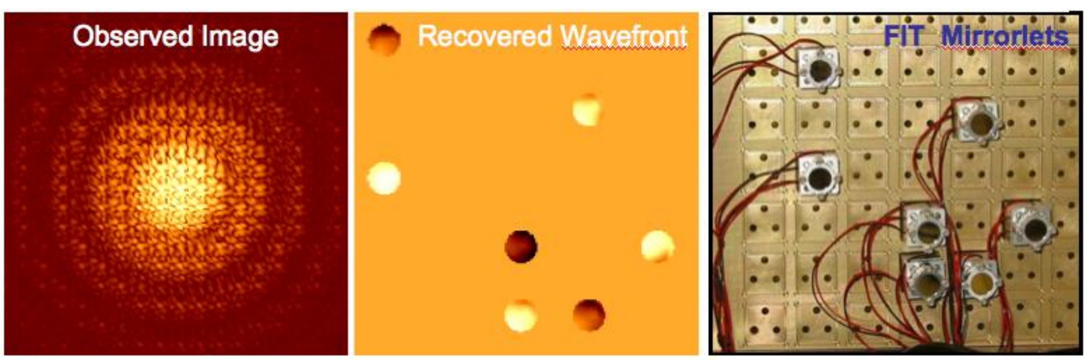


Figure 2 – *Left: Observed FIT Image*
Center: Recovered Wavefront
Right: Photo for FIT Actuated Mirrorlets

the synthetic aperture. Approaches of this type rely on encoding of the wavefront errors within the diffractive spreading of the object to be imaged (left side of Figure 2). The advantage of this approach is that no separate wavefront sensor hardware is required, i.e., the optical system used for science is generally the same system that is used for phase retrieval, thus data are collected with the focal-plane science camera. The resultant wavefront errors are subsequently decomposed into the eigenmodes of the control system (assumed linear) to provide commands for the actuators. The eigenmodes are used to generate linear combinations of voltages that are fed back to the actuators on the articulated mirrors. The mirrors are moved, at some bandwidth, to minimize an estimator of image quality, e.g. the minimum mean squared (rms) wavefront error.

The first phase of FIT contains 7 articulated mirrorlets (right side of Figure 2) each emulating a separate telescope flying in formation, with 3 degrees of freedom

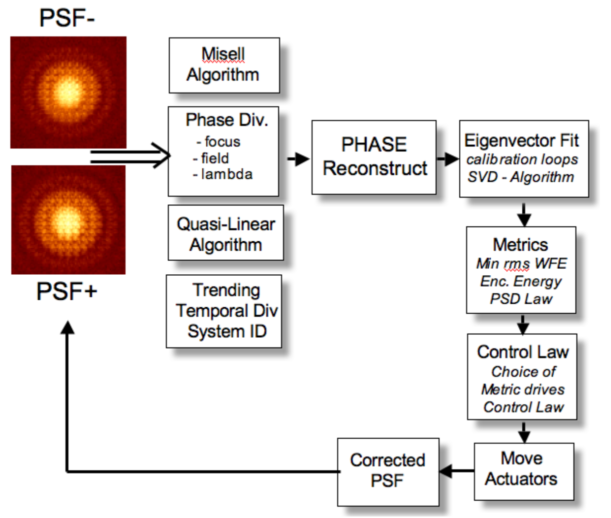


Figure 3 – *Baseline FIT Control Loop*

(tip, tilt, and piston) in a sparse non-redundant distribution which equals to 21 degrees of freedom (DOF) in all. Each mirrorlet comprises a component, i.e. a piece, of the synthetic aperture primary mirror. The baseline control method is sub-optimal due to first indirectly solving for the desired actuator motions via using phase retrieval of images (left side of Figure 2) to solve for the wavefront errors (middle of Figure 2), and secondly by employing a response matrix formalism to decompose the wavefront into the control eigenmodes. This baseline FIT control loop is shown in Figure 3. An alternative approach would be to solve, in a nonlinear fashion, directly for the control eigenmodes from the observed images. However this may be problematic since errors in the initial knowledge of the actuator response, from ground and space calibration, and drift of the actuator response on-orbit will likely change the overall system response. Techniques exist for solving for the actuator response (known as system identification), provided the drift is slow with respect to closed-loop bandwidth of the overall image correction system.

2.0 Optical Control and System Identification

The optical control model used on interferometric systems generally assumes a linear model of the form:

$$W = RV + n \qquad \text{Eq. (1)}$$

where W represents the 2D wavefront error lexigraphically reordered into a 1D vector; V is the vector of actuator voltages; R is the actuator response matrix, and n represents a vector of all noise sources (assumed additive) including readnoise, photon noise, quantization error, flat-fielding errors, and linearity assumptions. The response matrix R represents how the actuated commands affect the wavefront error. The actual implementation of the control is accomplished with a series of matrix multiplications which first map the recovered wavefront to piston, tip and tilt in the optical pupil plane and secondly to piston, tip and tilt of each mirror surface and subsequently to actuator voltages; in this formalism R represents the product of these matrices. Equation (1) can be solved to minimize the root mean squared (rms) value of the residual of wavefront error via pseudo-inverse and/or singular value decomposition techniques.

The maximum likelihood solution, vial linear least–squared solution, of the above model is[5]:

$$V = (R^T C_n^{-1} R)^{-1} R^T C_n^{-1} W \qquad \text{Eq. (2)}$$

where C_n is the noise covariance matrix. Thus the actuator voltages can be calculated, and fed back at each time step, from the phase retrieval recovered wavefront and the assumed known response matrix using matrix techniques via equation (2). Control bandwidth is assumed high enough to adequately correct system disturbances. The response matrix R is estimated during initial control loop calibration of the FIT testbed by applying to each actuator, on a given mirror, a known voltage and measuring the change in the corresponding wavefront. This dataset is collated and singular value decomposition (SVD) used to estimate the matrix R and its eigenmodes. The control actuator voltages are subsequently calculated from the dot product of an associated eigenmode and the recovered wavefront over that mirrorlet.

In principle the above approach would work, however, errors in knowledge of the actuator response matrix R, and time dependent drifts of R, will cause errors in the estimated actuator voltages resulting in errors in the closed-loop corrected wavefront. If the wavefront was not changing dynamically this closed-loop error would be corrected in subsequent time steps, but, generally the wavefront is dynamic resulting in poorer wavefront correction overall. System identification is an algorithmic method whereby the response matrix is periodically updated.

3. RESULTS

Figure 4 shows the wavefront error (radians) versus time for actual laboratory data from the FIT testbed[1]. The cyan curve shows the uncorrected wavefront error which has a mean of approximately 1.2 radians. The red curve shows only tip and tilt correction over each aperture and blue if only piston is corrected.

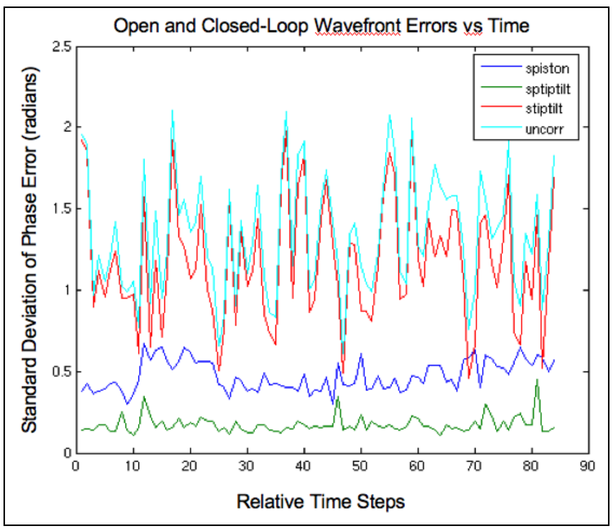


Figure 4 – Open and Closed Loop Wavefront Error

The green curve shows simultaneous correction of piston, tip and tilt. It can be seen that a net reduction in wavefront error of approximately a factor of 10 results. These results were for an ideal response matrix; what would happen with errors in the response matrix?

Figure 5 shows the effect on the closed-loop response with random errors in the response matrix. Plotted is the mean (blue curve) and +/- 1 standard deviation (red dashed curves) of the actuator corrected wavefront error versus the relative error in the response matrix. Relative error in the response matrix is defined with respect to the maximum component of R, i.e. random Gaussian distributed noise was added such that its standard deviation was the fraction plotted on the x-axis of Figure 3. Errors of < 1% in the response matrix cause unacceptable errors in the control of FIT. It is methods designed to compensate for these errors, i.e. system identification methods, that we propose to study.

Examples of such system identification techniques under study include: least-squares method, and observer/Kalman filter identification (OKID) method using time-domain multi input-output data [1,2]. OKID method is a direct Kalman filter gain approach, which can effectively compute the Markov parameters, and identify state-space model and corresponding observer simultaneously using time-domain input-output data.

OKID also can be easily extended to identify closed-loop control systems. Once these approaches are well characterized, and one or more is shown superior, it will be assessed for potential use on formation flying missions such as Stellar Imager.

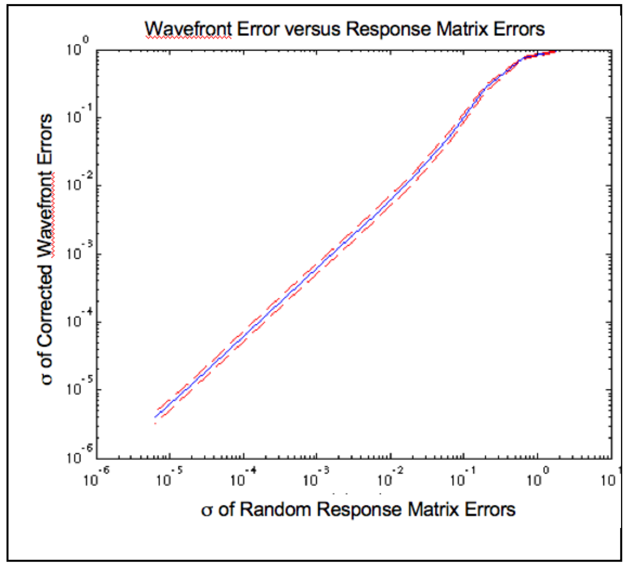


Figure 5- Effect of Errors In Response Matrix

4.0 CONCLUSION

The focus of this project is to develop and evaluate different system identification techniques. The first component of this study will be to identify the set of all algorithms which exist and to identify a limited subset which with potential applications for formation flying optical control. Secondly this subset of techniques will be first studied and characterized both analytically and computationally via Monte-Carlo type simulations to characterize their response in terms of accuracy, precision, dynamic range, computational speed and bandwidth. Third, any approach deemed feasible would be tested on FIT. Pending its performance evaluation in a laboratory environment it would be evaluated and assessed and expanded for space-based formation flying.

5.0 REFERENCES

[1] J.N. Juang, **Applied System Identification**, Prentice-Hall, Englewood Cliffs, N.J., 1994.

[2] J.N. Juang, Phan, M., Horta, L.G. and Longman, R.W., "Identification of Observer/Kalman Filter Markov Parameters: Theory and Experiment", **Journal of Guidance, Control and Dynamics**, Vol. 16, No.2, 1993, pp. 320-329.

[3] R.G. Lyon, K.G., et.al., "Building the Fizeau Interferometer Testbed", **IEEE Aerospace conference**, Big Sky, Montana, March 2004.

[4] R.G. Lyon, J.E. Dorband, and J.M. Hollis, "Hubble Space Telescope Faint Object Camera Calculated Point Spread Functions", **Applied Optics**, 36, No.8, 1997.

[5] K. Ogata, **Discrete-time Control Systems**, Prentice-Hall, Englewood Cliffs, N.J., 1987.

Electrochemical decontamination

Anatolii A. Omel'chuk, Inna N. Yudenkova
V.I. Vernadskii Institute of General and Inorganic Chemistry,
Ukrainian National Academy of Sciences,
Prospect Palladina 32-34, 03680 Kyiv, Ukraine

Yang Li Qiang, Huang Jian Wen
Guangzhou Baiyun Chemical Industry Co., Ltd.
No 1 Yunan Road, Guangzhou Civilian Science & Technology Park, Taihe, 510540 Guangzhou, China

Nikolay A. Maslo
National University of Food Technologies,
vul. Volodymyrska 68, 01601 Kyiv, Ukraine

ABSTRACT

This paper presents results of electrochemical radiological decontamination of fragments of the basic equipment which was used at nuclear power stations and in an experimental nuclear reactor. Equipment made of stainless and low-alloy steels was subjected to radiological decontamination. Mixtures based on orthophosphoric and sulfuric acids ware used as solutions for surface treatment. To protect the products being treated against corrosion and to reduce the working current density and electrolyte temperature, inhibiting compositions based on derivatives of aromatic and aliphatic amines and amides were added to the electrolyte. The optimal compositions of reaction mixtures and conditions of electrochemical radiological decontamination, which ensure the polishing regime, have been determined.

Two modes of radiological decontamination of basic equipment have been developed: mode of stationary tanks and mode of external electrodes.

The possibility to conduct electrochemical surface treatment at very small electrode spacing allows one not only to reduce the specific electrolyte consumption but also to greatly reduce the specific electrical energy consumption and to treat the surface of equipment of large geometrical dimensions without dismounting.

It has been shown that it is possible in principle to carry out electrochemical radiological decontamination of articles made of nonconducting materials.

Keywords: electrochemical radiological decontamination, metal surfaces, stationary regime, external electrodes.

INTRODUCTION

The problem of ensuring a high level of safety of the operating personnel of nuclear reactors and reducing the ecological pressure on the environment is an important applied and scientific problem. According to the technological regulations, the equipment of nuclear power stations and experimental reactors (heat-exchange equipment, pipelines and joints, devices for loading fuel elements, etc) is liable to periodical inspection with the object of detecting damages and preventive repair. During operation, this equipment is contaminated by radionuclides, acquires induced radiation and becomes a source of increased danger for the operating personnel and the environment, therefore it requires radiological decontamination. Besides, as experience shows, emergencies often arise at nuclear power stations. To ensure the safety of the operating personnel and to eliminate the consequences of emergencies, reliable and highly efficient processes and technologies for the radiological decontamination of basic equipment are required.

Under the conditions of nuclear power station, metal surfaces are contaminated mainly by activated corrosion products of structural materials: by the isotopes ^{54}Mn, ^{55}Fe, ^{60}Co, the decay products ^{90}Sr, ^{137}Cs, and in the case of fuel-element failure by ^{238}Pu, ^{239}Pu, ^{240}Pu, ^{241}Am, ^{237}Nb, etc. The radiocontamination of the basic equipment of nuclear power stations and pilot reactor may be arbitrarily divided into in-service contamination and contamination caused by accident. In-service radiocontamination is typical of the equipment of the reactor department and the internal surfaces of the equipment for special water treatment and water purification. Radioactivity caused by accident is most characteristic of reactor, turbine and electrical equipment.

As a rule, the major part of radiocontamination accumulates in metal surface defects (microcracks and microrecesses) of the surface layer, formed by oxide compounds, and can be removed by mechanical, chemical and electrochemical methods [1-3]. The thickness of this layer is estimated to be of the order of 50 μm.

Analysis of the current state of the problem [1-7] shows that among known radiological decontamination methods (chemical, mechanical, electrochemical) only the electrochemical method enables attainment of a high rate of surface layer dissolution and high-quality radiological decontamination at low specific consumption of reagents and electrical energy, ensuring a high corrosion stability of the metal surfaces of the decontaminated equipment.

In spite of the variety of the existing methods for the radiological decontamination of basic equipment, the overwhelming majority of them, including electrochemical methods, call for considerable improvement. For example, it is known that during electrochemical radiological decontamination, the overwhelming majority of metals change into the passive state, owing to which the efficiency of extraction of radionuclides from the surface layer decreases; however, the interrelation between the decontamination conditions (composition of decontaminating solutions and polarization conditions) has not been established. The conditions that allow the transition to stable passivity to be prevented have not been determined.

There is practically no information on the principle of operation of external electrodes, the peculiarities of design approaches to the supply of decontaminating solutions to the decontamination zone, the use of structural materials. There is no information on the interrelation between polarization conditions, decontaminating solution composition and the dissolution rate of the surface layer of various structural materials, the degree of their decontamination, the specific consumption of electrical power and electrolyte. It is known [5-7] that radioactive isotopes can be extracted from disperse nonconductive materials through electrokinetic phenomena, but there is no information on the radiological decontamination of nonconducting massive materials in an electric field. The lack of this information does not make for wide practical application of highly efficient electrochemical processes.

The prospects of using electrochemical processes for radiological decontamination increases the possibility of using not only direct current, but also alternating current, ultrasonic activation. Using an external electrode, one can decontaminate equipment of large geometrical dimensions without dismounting it from installation site, which cannot be done when using stationary tanks.

The lack of information on the conditions of the electrochemical radiological decontamination of different structural materials under different polarization conditions and in decontaminating solutions of different composition necessitated research in this direction.

The results obtained can be used to work out technological regulations for electrochemical decontamination process.

EXPERIMENTAL DETAILS

The dependence of the degree of radiocontamination on electrochemical treatment conditions, the character of radiocontamination and the materials to be decontaminated was determined on the territory of the pilot nuclear reactor of the Institute for Nuclear Research of the Ukrainian National Academy of Sciences. For decontamination, we took fragments of the equipment of the pilot reactor during scheduled preventive repair and specimens from the nuclear power station at Chornobyl.

For electrochemical treatment, current- and voltage-stabilized dc sources were used. The voltammetric investigations were carried out using a PI-50-1 potentiostat with automatic recording of voltammograms. The voltammograms were recorded in a three-electrode cell with respect to a normal hydrogen reference electrode at a potential scan rate of 2 mV s^{-2}.

The radiocontamination level of specimens under investigation before and after radiological decontamination was monitored using equipment of the Institute for Nuclear Research of the Ukrainian NAS by means of special devices (KRB-1, RKB4-1eM, MKS-01R radiation meters, gamma-spectroscopic complex).

The mass loss of specimens under investigation during electrochemical treatment was determined by the gravimetric method.

The sight control of the surface condition of specimens under investigation was performed on a REMMA-101M scanning electron microscope-microanalyzer.

The change in the mechanical properties of the surface was determined by measuring microhardness.

RESULTS AND DISCUSSION

Voltammetric investigations

Electrochemical radiological decontamination is based on the dissolution of metal surface layers by the action of both direct current [1-3] and alternating current [8, 9].

The higher the dissolution rate, the higher the decontamination rate. The investigations carried out showed that the overwhelming majority of metals are passivated under anodic polarization by dc. That is conditions exist under which there is practically no dissolution of the surface layer, and hence no decontamination. A typical trend of polarization curve for carbon steel (St.3) in a solution of sulfuric and phosphoric acids is shown in Fig. 1.

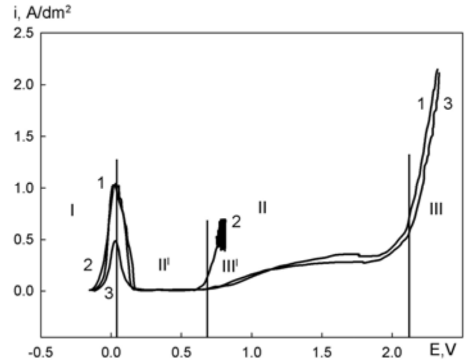


Fig.1. Current-potential curves for carbon steel (St 3) in an acidic decontaminating solution, obtained in the modes: (1) stationary tank at a rate of electrode rotation of 7 rpm; (2) external electrode with rotation; (3) external electrode without rotation. Composition of the solution, wt.%: H_2SO_4, 40; H_3PO_4, 40, H_2O, 20.

The current – potential curve may be arbitrarily divided into three sections. An analysis of the results obtained showed that on the first section (I), dissolution of the surface layer, and hence decontamination, takes place, but the surface of specimens is etched. This decontamination mode is not recommended for articles or equipment of predetermined geometrical dimensions.

There is practically no decontamination at the conditions of anodic polarization, which corresponds to the second section (II, II^I) of the polarization curve. The metal surface is in the passive state. Treatment of the metal surface, which corresponds to the third section (III, III^I), ensures the best decontamination conditions, decontamination being accompanied by polishing in acidic decontaminating solutions based on a phosphoric acid – sulfuric acid mixtures.

The conditions and character of anodic dissolution depend on the nature of metal, the composition of the solution, anodic polarization value. Thus, to optimize the electrochemical radiological decontamination conditions, information on the electrochemical behavior of metals is required.

In the external electrode mode, in contrast to the stationary tank mode, mass exchange between the electrodes is through a porous dielectric material saturated with decontaminating solution. Description of the principle of operation of such a device is given below. The occurrence of electrode processes in the systems where the electrolyte is in the porous material matrix differs from that in ordinary systems, where convective diffusion makes a considerable contribution to mass transfer.

In view of this, current-voltage characteristics for electrodes made of different materials have been investigated under different conditions of mass transfer to the interface. Figure 2 shows schematically electrolytic cells, in which mass transfer to the electrode working surface occurs under the conditions of free (without rotation of the electrode) and forced (with rotation of the electrode, show with an arrow) convection (a) and in a cell where convective transfer is ruled out by a porous dielectric (b).

For the purpose of renewing the electrolyte in the pores of the dielectric material, the surface of the electrode itself was able to rotate.

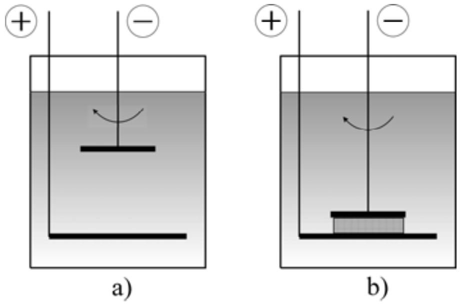


Fig. 2. Arrangement of the electrodes in the case of studying the effect of the conditions of solution delivery to the electrode surface on the shape of current-potential curves: (a) stationary tank mode; (b) external electrode mode. Rate of electrode rotation: 7 rpm.

Such variants of disposition of electrodes simulate decontamination in stationary tank and external electrode modes. Carbon steel (St 3), stainless steel (12Kh18N10T) and copper were investigated as materials that were subjected to electrochemical treatment.

The results obtained (Fig.1) showed that in acidic decontaminating solutions, carbon steel changes to passive state on anodic polarization. The potential of transition to passive state (so-called Flade potential) is 0.0±0.05 V and depends neither on the mode of delivering decontaminating solution to the interface nor on the acid solution composition.

In the stationary tank and external electrode modes in the case of forced solution convection (rate of electrode rotation: 7 rpm), the current density in transition to passive state is about twice higher than without forced solution delivery to the interface.

In the case of forced solution convection in the porous material (translational motion of external electrode relative to the surface being decontaminated), the passivity region decreases in comparison with stationary tank mode, which reduces specific electrical energy consumption.

Moreover, it has been noted that transition from passive state to transpassivation regime in the case of external electrode with its rotation is accompanied by the occurrence of current oscillations. The occurrence of unstable state at the interface contributes to higher rate of the decontamination process.

In comparison with carbon steel, current-potential curves for stainless steel are characterized by a much larger passivity region (0.0 to +(1.4-1.5) V). On current-potential curves there is no clearly defined transpassivation region, therefore it is recommended to carry out radiological decontamination in this case at high current densities. A small shift (of up to 50 mV) towards more positive values on changeover to the external electrode mode with forced solution convection has been noted. In comparison with carbon steel, the current densities corresponding to Flade potential are almost half as high. Typical current-potential curves for stainless steel in acidic decontaminating mixtures of different composition are shown in Fig. 3. The numbering of the curves in the figures corresponds to the conditions of solution mass transfer to the interface given in the captions to Fig. 1.

The results obtained give evidence that only external electrode mode gives rise to unstable states (current oscillations) on the stainless steel surface at relatively low potentials (1.5 – 1.6 V) and hence makes dissolution of surface layers possible at low current densities.

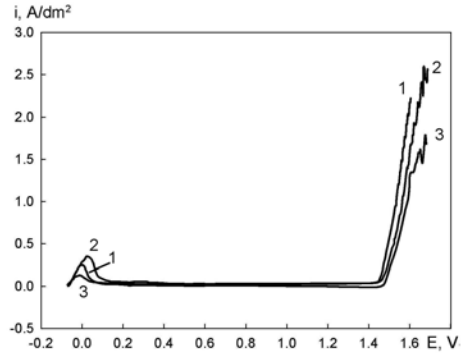


Fig. 3. Current-potential curves for stainless steel (12Kh18N10T) in an acidic decontaminating solution.

Electrochemical decontamination of radioactively contaminated materials

Results of the electrochemical radiological decontamination of different structural materials are given in Table 1. Specimens of the pipeline of the primary coolant circuit (material 12Kh18N10T stainless steel), parts of the heat-exchanger loop: bearing race (40YuT steel), composite: ring with press-fitted bronze bushing (KhN35VTYu steel, BROTs 10-2 bronze), fittings from the reactor room (stainless-clad carbon steel) were decontaminated. Besides, samples of materials removed from the pilot reactor during scheduled repair were decontaminated. These materials were 12Kh18N9T stainless steel, 09G2S steel, beryllium neutron displacer, rubber, polyvinyl chloride, porcelain (Rashek rings), ion-exchange resin.

Metallic specimens were decontaminated in decontaminating solutions of the following composition (wt%): Orthophosphoric acid, 65; Sulfuric acid, 20; Water, 15 and Inhibiting composition (product of polymerization aniline with urotropine and its following condensation), 7.5 vol% with respect to acid mixture.

The nonconducting materials were decontaminated both in an acid solution of the above composition and in a sodium sulfate based solution: Sodium sulfate, 1M; Trilon B, 30 g/L; Seignette salt, 30 g/L.

The basic diagram of the electrolytic cell for the radiological decontamination of nonconducting materials in the field of electrical current is shown in Fig. 4.

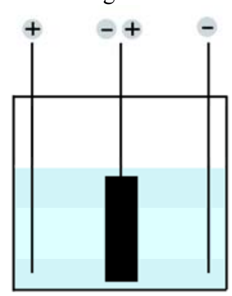


Fig. 4. Basic diagram of an electrolytic cell for the radiological decontamination of nonconducting materials.

The investigation carried out showed that the electrochemical radiological decontamination of surfaces of equipment made of both stainless steel and low-alloy carbon steel ensures a high decontamination level. The exposure rate of radicontamination of the specimens under investigation after decontamination did not exceed the reference level values allowed for rooms where the personnel work periodically.

The investigations carried out showed that the best effect in electrochemical radiological decontamination is achieved in the case of uniform dissolution of the metal surface layer of 20 μm thickness. In some experiments, the rate of the anodic dissolution of the structural materials of which the basic equipment is made was determined (Table 2).

Table 1.

Characteristics of radiological decontamination of various materials used in the basic equipment of nuclear reactors

Material to be decontaminated	Decontamination conditions, modes of radiological decontamination	Radiocontamination (β decays/cm² min)		Character of radiocontamination
		Before decontamination	After decontamination	
12Kh18N10T stainless steel (a part of the water pipeline of the primary coolant circuit)	Acid solution, $i = 35$ A/dm², $\tau = 30$ min.; stationary tank	5000	30	^{137}Cs, ^{134}Cs, ^{60}Co, ^{90}Sr
12Kh18N10T stainless steel (reactor plug)	Acid solution, $i = 7$ A/dm², $\tau = 15$ min.; stationary tank	100	–	^{137}Cs, ^{90}Sr, ^{60}Co, ^{55}Fe
St3 carbon steel (specimens of fittings)	Acid solution, $i = 35$ A/dm², $\tau = 40$ min.; external electrode	$1\cdot10^6$	$3\cdot10^3$	^{137}Cs, ^{90}Sr, ^{60}Co, ^{55}Fe ^{154}Eu
12Kh18N10T stainless steel (inner walls of the "hot" box)	Acid solution, $i = 8$ A/dm², $\tau = 3$ min.; external electrode	200	10	^{137}Cs, ^{134}Cs, ^{60}Co, ^{90}Sr
12Kh18N10T stainless steel (equipment for special water treatment)	Acid solution, $i = 10$ A/dm², $\tau = 4$ min.; external electrode	450	35	^{137}Cs, ^{134}Cs, ^{60}Co, ^{90}Sr, ^{59}Fe
40YuT steel (bearing race)	Acid solution, $i = 20$ A/dm², $\tau = 30$ min.; stationary tank	5000	200	^{137}Cs, ^{90}Sr, ^{60}Co, ^{55}Fe ^{154}Eu
40Kh steel (spacer sleeves)	Acid solution, $i = 5$ A/dm², $\tau = 40$ min.; stationary tank	2300	–	^{137}Cs, ^{134}Cs, ^{60}Co, ^{90}Sr
KhN35VT10 steel, BROTs-10-2 bronze (Ring with press-fitted bronze bushing)	Acid solution, $i = 15$ A/dm², $\tau = 40$ min.; stationary tank	8000	10	^{137}Cs, ^{134}Cs, ^{60}Co, ^{90}Sr
Beryllium neutron displacer	Acid solution $I = 15$A, $U = 20$ B, $\tau = 15$ хв. stationary tank	650	–	^{137}Cs, ^{60}Co, ^{154}Eu, ^{90}Sr
Rubber seal ring of the primary coolant circuit	Neutral solution, $I = 6$ A, $U = 8$ V, $\tau = 35$ min.; stationary tank	450	100	^{137}Cs, ^{60}Co, ^{154}Eu, ^{90}Sr, ^{59}Fe
Porcelain (Rashek rings)	Neutral solution, $I = 6$ A, $U = 8$ V, $\tau = 35$ min.; stationary tank	600	100	^{137}Cs, ^{134}Cs, ^{60}Co, ^{154}Eu
Porcelain (Rashek rings)	Acid solution solution, $I = 15$ A, $U = 15$ V, $\tau = 35$ min.; stationary tank	600	20	^{137}Cs, ^{134}Cs, ^{60}Co, ^{154}Eu
Polyvinyl chloride	Acid solution solution, $I = 15$ A, $U = 26$ V, $\tau = 30$ min.; stationary tank	500	50	^{137}Cs, ^{134}Cs, ^{60}Co

Table 2

Dissolution rate of the metal surface layer in radiological decontamination

Structural materials	Current density, A/dm²	Rate of dissolution of surface layer, μm/min.
12Kh18N10T	35	1,4-1,5
–“”–	15	0,5-0,6
Carbon Steel	40	1,5
–“”–	25	1,4
–“”–	10	0,7
40Kh Steel	10	0,6
–“”–	5	0,5
40YuT steel	30	0,9
–“”–	20	0,7
KhN35VTYu	20	1,2
–“”–	10	0,8
Bronze 10-2	30	1,0
–“”–	15	0,5

The data are given for an acid solution, whose composition was specified above.

The surface layer dissolution rate increases with current density. The largest values are typical of austenitic and carbon steels.

Increasing the temperature and the presence of complex additives in acidic decontaminating solutions also make for higher surface layer dissolution rate. For example, whereas in the case of 12Kh18N10T stainless steel the surface layer dissolution rate is 0.6 μm/min at 25 °C and a current density of 15 A/dm², it is 0.8 μm/min at 60 °C.

The change in the mechanical properties of surfaces before and after electrochemical radiological decontamination was studied by X-ray structural and microindenter analysis and optical microscopy. The electrochemical treatment was carried out in an acid solution at a current density of 15 $A\cdot dm^{-2}$ for 20 min. Kh18N10T austenitic steel and carbon steel (St3) were investigated. As a result of the research carried out by us, it has been found that after electrochemical radiological decontamination, the number of microdefects, microdislocations and other microformations of the second kind on the surface decreases (the parameter $\Delta\alpha/\alpha$ decreases). For instance, whereas before treatment the parameter $\Delta\alpha/\alpha$ for St3 was 1.2×10^{-3}, after treatment it is practically not identifiable. The dislocation density is lower than in the standard.

The research carried out by showed that the electrochemical treatment of Kh18N10T austenitic steel and St3 carbon steel makes for their better mechanical properties. The number of microformations and dislocations decreases noticeably, the structurally stressed state improves.

The observed effect may be attributed to the fact that during electrochemical treatment, the surface layer saturated with stress and dislocation concentrators dissolves. On the metal surface is formed, under anodic polarization, a more homogeneous oxide film, which prevents dislocations form reaching the surface and makes for higher plastic resistance.

There are grounds to believe that the improvement of structural and elastic characteristics after anodic electrochemical treatment is determined by the peculiarities of interface processes and formation thereby of a new surface layer.

External electrode mode

An electrode design has been developed, in which the porous material can be quickly replaced with the aid of a special mount with rope tie wich prevents the contact of the operating personnel with it. This is especially important in replacing radionuclide-saturated material. The principal scheme of the device is shown in Fig.5.

External electrode consists of a telescopic bar 1, on which the handle 2 of a clamping device 5 is fastened, by means of which a porous material 7 is fastened to the operating surface of the electrode 6. By a special tubing 3 decontaminating solution, which is necessary for the electrochemical removal of radionuclides from the surface of the equipment being decontaminated, is conveyed into the porous material 7. Radioisotopes accumulate in this porous material, which is moved over the surface being treated by means of the telescopic bar 1.

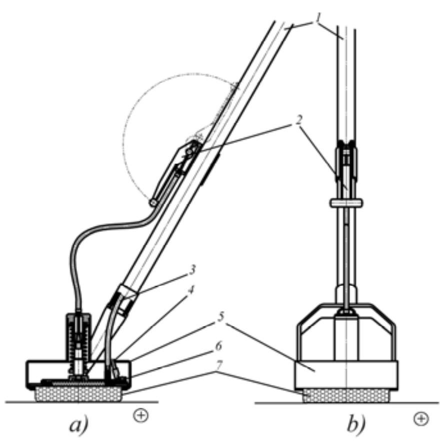


Fig.5. General view of external electrode (its operating part): side view (a) and front view (b)

The following materials are recommended as porous material: polypropylene, silica or basalt multiply cloths. As a result of the investigations carried out it has been found that silica material in the form of multiply cloth shows the highest ability to retain decontaminating solutions. In spite of their higher mechanical stability, polypropylene cloths have a worse ability to retain decontaminating solution. Besides, silica, basalt, carbon cloths and fiber have a much larger sorptive capacity for radionuclides in comparison with polypropylene material. The specific solution consumption in the case of using this material is estimated to be 3-4 g/dm^2 under the action of direct current. Current is supplied to the operating surface by means of a current lead 4.

In the case of decontamination with direct current, the external electrode acts as a cathode, and the surface being decontaminated acts as an anode.

The working surface area of the electrode is 1 dm^2, its mass is 2 kg. The maximum current load is 80 A, which enables 80 A/dm^2 decontamination conditions to be used. The dc power is not over 1.5 kW.

A universal engineering documentation has been worked out, which allows one to organize serial production of external electrodes.

CONCLUSIONS

- It has been found that almost all metals whose isotopes contaminate the surface of the basic equipment (light actinides, cesium, strontium, cobalt, iron, nickel, etc) have a more negative ionization potential than stainless and carbon steel, of which the overwhelming majority of basic equipment is made. An exception is nickel, which has a more positive ionization potential (+ 1.65 V) in sulfate solution than stainless steel. This indicates that solutions based on various inorganic acids and their salts can be used for radiological decontamination. The anodic polarization regions in which radiological decontamination is accompanied by polishing have been established.
- Compositions of solutions for the electrochemical radiological decontamination of various structural materials have been proposed. Orthophosphoric and sulfuric acids, sodium sulfate, nitrite and nitrate form the basis of such solutions.
- To reduce corrosion and to improve the decontaminating and polishing ability, complexing agents and inhibiting compositions are added to solutions.
- The conditions ensuring the predetermined decontamination level of various structural materials have been determined, metal loss during decontamination has been estimated.
- The investigations carried out showed that the electrochemical treatment of Kh18N10T austenitic steel and St 3 carbon steel makes for their better mechanical properties. The number of microdefects and dislocations is greatly reduced, and the structurally stressed state improves.
- It has been found that not only metallic articles, but also articles made of nonconducting materials are amenable to electrochemical radiological decontamination. A design for the radiological decontamination of nonconducting materials in dc field has been proposed.
- It has been shown that the use of the external electrode mode of electrochemical radiological decontamination makes it possible to greatly reduce the consumption of reagents and electrical energy for radiological decontamination.
- Design documents for making experimental equipment for electrochemical radiological decontamination in two modes: in the stationary tank mode and in the external electrode mode have been worked out.
- Experimental models of external electrodes with an operating surface area of 1 dm^2 have been made. A design for the fastening of the porous material, which holds the decontaminating solution, to the electrode has been

developed, which allows one to replace it quickly and conveniently. Designs for the renewal of the decontaminating solution in the external electrode have been found. The designs of external electrodes developed permit one to work with both acidic and neutral salt decontaminating solutions.

- Technological regulations for the radiological decontamination of various materials the stationary tank and external electrode modes have been worked out.

REFERENCES

1. **State of the Art Technology for Decontamination and Dismantling of Nuclear Facilities.** Vienna : International Atomic Energy Agency, 1999. – (Technical reports series, ISSN 0074–1914 ; No. 395).
2. J. H. Lee, Y. K. Lim, H. Y. Yang, S. W. Shin, M. J. Song, **"Application of a Modified Electrochemical System for Surface Decontamination of Radioactive Metal Waste",** Transactions of the 17th International Conference on Structural Mechanics in Reactor Technology (SMiRT 17), Prague, Czech Republic, August 17 –22, 2003.
3. **Decontamination Techniques Used in Decommissioning Activities,** A Report by the NEA Task Group on Decont., http://www.nea.fr/html/rwm/reports/1999/decontec.pdf.
4. P.G. Ogul'nik, " Modern methods of decontamination of reactor equipment", **Atomnaya Tekhnika za Rubezhom,** No 9, 1989, pp. 23-27 (in Russian).
5. B. Kornilovich, N. Mishchuk, K. Abbruzzese, G. Pshinko, R. Klishchenko, "Enhanced electrokinetic remediation of metals-contaminated clay", **Colloids and Surfaces A: Physicochem. Eng. Aspects,** Vol. 265, 2005, pp. 114–123
6. Sibel Pamukcu, "Electro-Chemical Technologies for In-Situ Restoration of Contaminated Subsurface Soils" **Electronic Journal of Geotechnical Engineering (*EJGE*),** paper 1997-04; http://www.ejge.com/1997/Ppr9704/
7. Rainout Lageman, "Electroreclamation", **Environ. Sci. Technol..** Vol. 27, No. 13, 1993, pp. 2648-2650.
8. **USSR invention certificate No 1349347,** C 25F 3/02. A methods for the electrochemical radiological decontamination of stainless steels. G.N. Zakharchuk, L.N.Vasilyeva, N.N.Smirnov, V.Ye.Ulanov, A.N.Usdkin, A.K.Yuryevskii. Published 30 July 1991, Bulletin No 28.
9. **USSR invention certificate No 1262997**, C 25F 3/02. A methods for the electrochemical radiological decontamination of stainless steels. N.N.Ampelogova, R.V.Bolukov, L.I.Vasilyeva, G.A.Zakharchuk, V.V.Morozov, V.Ch.Pentin, G.V.Ryumin, N.A.Smirnov, V.Ye.Ulanov. Published 15 August 1991, Bulletin No 30.)

A New Look at Hydrogen Fueled Supersonic Airliners

Alex Forbes, Anant Patel, Chris Cone, Pierre Valdez, Narayanan Komerath
Daniel Guggenheim School of Aerospace Engineering
Georgia Institute of Technology, Atlanta, GA 30332, USA

Abstract

This paper takes a new look at the prospects for developing supersonic civil airliners, considering global demographics, climate change issues, fuel prices and technological advances. Dramatic changes have occurred in the demographics, economics, and market intensity of the Eastern Hemisphere since the 1990s. Carbon reduction imperatives provide a major incentive to invest in developing hydrogen-fueled airliners. The "point-to-point" air route architecture has proved viable with long range mid-size airliners. With a cruise Mach number of 1.4, a large number of destinations become viable for overland supersonic flight. A conceptual design process is used to estimate cost per seat mile for a range of hydrocarbon and hydrogen fuel costs. An argument based on the ideal shape for minimal wave drag, estimates the drag penalty from using hydrogen. Viable aircraft geometries are shown to exist, that match the theoretical ideal shape, showing that the drag estimate is achievable. Conservative design arguments and market estimates suggest that hydrogen-fueled airliners can achieve seat-mile costs low enough to open a large worldwide market and justify a viable fleet size.

Keywords: Hydrogen supersonic airliner, wave drag, seat mile cost, demographics

1. Introduction

The technical and business cases for liquid hydrogen-fueled supersonic transport airliners (LH2 SST) are re-examined in the light of changes that have occurred in demographics, fuel prices and greenhouse gas reduction imperatives. The paper lays out the cases for the existence of a much larger market than was seen for supersonic airliners in the 1950s through 70s, or in recent studies in the 1990s. It then uses conceptual design to explore the fears regarding the high wave drag penalty of using liquid hydrogen. Finally it projects the cost per seat-distance that can be achieved using hydrogen fueled supersonic airliners, to close the loop on the argument about demand.

The Concorde[1] and the Tupolev 144[2] pioneered supersonic airliner flight in the 1960s, but neither achieved anywhere near the fleet size needed to be viable in the marketplace. Tu-144 regular passenger service across the Soviet Union was cancelled after only 55 flights, citing safety issues. The Concorde was not allowed to fly overland at supersonic speed because of the perceived destructive effects of sonic boom. Of some 200 initial orders only 14 entered commercial service. The oil crises of the 1970s and 80s, the Cold War and US-Europe competition precluded viability of either the Concorde or the American SST concepts. The cost of supersonic travel stayed beyond the means of most travelers, preventing the market from expanding.

The High Speed Civil Transport (HSCT) project in the USA concluded[3] in 1999 that the market did not justify development of SSTs. Experts pointed out that the airlines' business model depends on business/first class travelers to make long-distance routes viable. An SST would take away these high-paying passengers, and thus cut the low-risk profit of the transonic fleet while taking on a huge new risk. This conclusion appeared to be drawn from a market survey that included only US trans-Atlantic and trans-Pacific routes. Current interest in SSTs appears to be limited to business jets.

Hydrogen-fueled airliners were perceived to be impractical for 4 reasons:

- The presumed difficulty in handling liquid hydrogen safely.
- The high wave drag associated with the presumed large volume of liquid hydrogen.
- The high cost of producing and storing hydrogen in sufficient quantities.
- The presumed energy inefficiency and carbon footprint of producing hydrogen starting with fossil-driven power plants.

Against these objections, there are several newer developments that demand a new look at supersonic hydrogen-fueled airliners:

- There may be substantially more demand for supersonic airline travel, than considered before.

- Security and congestion considerations have advanced the point-to-point airline architecture over the hub-and-spoke architecture.
- Point-to-point trips now exceed 17 hours using long-range airliners, showing viable demand despite low payload fractions.
- Reduced time for point-to-point travel would increase trip frequency per aircraft.
- Going to Mach 1.4 may offer enough reduction in travel time to attract a larger market.
- With current technology, using atmospheric winds and density layers, sonic boom will be imperceptible on the ground at up to Mach 1.4.
- The air travel industry's mandate to cut carbon emissions provides a large and unique source of funding, to develop hydrogen-fueled aircraft.
- In the longer term, hydrogen costs should come down, supply being unlimited.

2. Summary of Issues

The problem is distilled to the following questions:

- How have demographics and economic development altered worldwide market projections for supersonic transport? What are viable destinations, and what are the flight times, curfew implications and business implications of supersonic flight between these destinations?
- What is the drag implication of using hydrogen, given the lower fuel weight fraction?
- What is the impact of Global Warming/ Carbon emission reduction initiatives on the prospects for hydrogen-powered flight?
- What are the noise implications of the LH2SST?

3. Growth of World Wide Air Travel

Airline travel has increased by nearly 300% since 1980[4], reaching 4300 billion passenger-kilometers and 160 billion ton-kilometers by 2008. Deregulation of the US airline industry in 1978 increased the number of air travelers[5]. The world has changed drastically since the early 1990s. The Berlin Wall is down, and the European Union integrated. Russia's arctic airspace opened to many new air traffic routes[6]. South Africa is an open and booming economy and provides an intermediate stop for long-distance flights connecting Asia and Middle East economies with South America. African civil air traffic has seen a 5.7% annual growth in the past 15 years and expects a 7% increase in the coming decade. Most dramatic is the rise in the economies of Asia since the early 1980s, and the opening of travel in and to the People's Republic of China. Viable business destinations and international airports abound now in Central and Southern India, with busy air connections throughout India, the Middle East, Sri Lanka, East Asia and Europe. The world economy and job market have become "globalized". Along with this comes the desire of aging parents to visit their children and grandchildren working and living in distant parts of the world. A large new middle class has the desire, means and freedom to see the world, but not necessarily the stamina to survive flights of over 8 to 17 hours. Hence the potential market for supersonic travel may be far greater than that envisaged. Asia and the Pacific are at 28% of the market as of 2006. Based upon their rate of growth compared to the rest of the world they will more than likely gain ground on Europe but will not pass them for at least 40 years, assuming current growth rates[4]. The "broken third leg" of market demand that NASA cited in closing the HSCT project in 1999, is no longer broken when viewed in today's changed realities. The commercial air travel market is also expected to maintain a 4-5% a year increase globally, by conservative estimates, for the next 10 to 15 years. This would result in the market for air travel doubling over this period.

4. Fuel Prices and the Hydrogen Economy

The Hubbert Peak Oil theory[7] holds that fossil fuel prices will rise very sharply as the increase in demand surpasses the increase in supplies[8]. Many experts feel that this may be an imminent event[9], or may occur by 2018[10] or 2030[11]. Currently the airline industry is very reliant on fossil fuels. The industry is under increasing pressure to reduce emissions of carbon dioxide, from its levels of around 300 million tonnes per year[12]. In 2009, the International Air Transport Association (IATA) announced sharp cuts in emissions. In the short term, this can only come from buying carbon credits on the market or funding "clean development" projects around the world, to offset the emissions. Given a nominal price of $20 per ton of CO_2 per year this means buying credits worth over $2B per year, into the indefinite future. Most hydrogen produced today comes from steam reformation of fossil fuel. Shifting to renewable solar or wind sources and improving the efficiency of high-temperature electrolysis in new nuclear reactors will enable hydrogen to be produced at viable costs without generating greenhouse gases[9].

5. Sonic Boom Considerations

Supersonic flight causes a sharp, loud and damaging pressure signature in the shape of an "N" wave on the surface below. However, if the speed of sound at the ground is higher than the aircraft's speed then the boom is not an issue on the ground. This "threshold Mach number" is around 1.20 for many US cities[13]. When atmospheric thermal layers and winds are considered, the flight Mach number can be substantially higher than the threshold without the boom exceeding permissible noise levels[14]. The best flight altitude may thus be substantially lower than those previously considered for supersonic flight.

6. Preliminary Sizing and Performance

A conceptual design study incorporated the general requirements of flying supersonic, the fuel storage issue, and the performance parameters of supersonic cruise. A range of 5,000 statute miles was specified. Supersonic cruise at 45,000 ft was chosen. Following general design guidance[15] validated against Concorde numbers, the aircraft was sized for 200 passengers and 6 crew. An iterative process used the constraints:

- The minimum structure fraction needed to build the aircraft was set at 27%. Composite structures demonstrated with the Boeing 787 allow this.
- Engine technology was assumed at the level of the F-35 Joint Strike Fighter, reputed to have an engine thrust-to-weight ratio over 11.
- Thrust-specific fuel consumption was assumed to be 1.1 per hour, at the level assumed in the NASA HSCT project, at Mach 1.6 cruise.
- The length was limited to 67 m (220 feet).
- The comfort level of modern airline business class seats was assumed.

Figures are presented with British units for the convenience of American readers outside engineering, especially as related to cost metrics.

7. Supersonic Drag Argument

The volume needed to accommodate the payload and fuel, with wings of reasonable thickness, was obtained, for both the Jet-A and LH2 cases. The corresponding Sears Haack shape for minimum wave drag was computed. The drag of an actual airliner can be assumed to be close to this ideal. Once the shape was determined, a sanity check of the layout confirmed that the payload, cockpit and fuel could be accommodated.

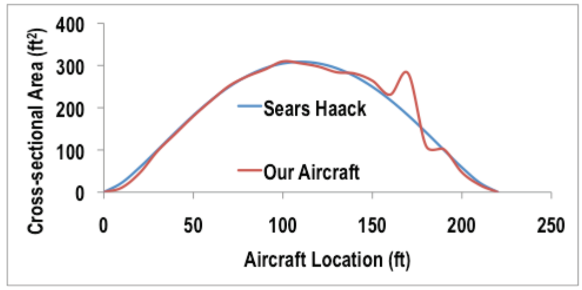


Figure 1: Area distribution of the conventional LH2 configuration, compared to the Sears-Haack minimum wave drag area distribution

Figure 1 shows that a conventional fuselage/ swept wing configuration (shown in Figure 2) can come to within 5% root-mean-square error of the Sears-Haack without much trouble. It is comfortably assumed that actual aircraft designers will be able to smoothen the sharp features.

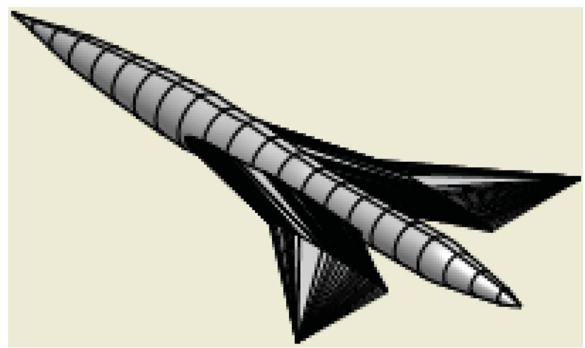


Figure 2: 200-seat Sears-Haack configuration

In supersonic area ruling[16], the area intersected by conical surfaces with the Mach angle (45.6 degrees or higher for the Mach 1.4 cruise case) is used to smooth out discontinuities that would cause shocks. This distribution is shown in Figure 3. It differs by a root mean square error of over 57% from the Sears-Haack, suggesting substantial modification of the wings and redistribution of the fuel into the fuselage.

Some corrections to the above should be considered. The inevitable shock from the nose will cause the relevant Mach number for the fuselage area ruling to be lower than Mach 1.4, thus causing an increase in the Mach cone angle to be used. This would drive the ideal area distribution further towards the Sears-Haack distribution of Figure 1. Nickolic and Jumper (Ref. 16) discuss the issues in comparing the results of different predictions with experimental results, and indicate substantial uncertainties, even in the zero-lift wave drag analysis. Determining the configuration for lowest achievable drag at Mach 1.4 is a matter to be left to more detailed aerodynamic analysis.

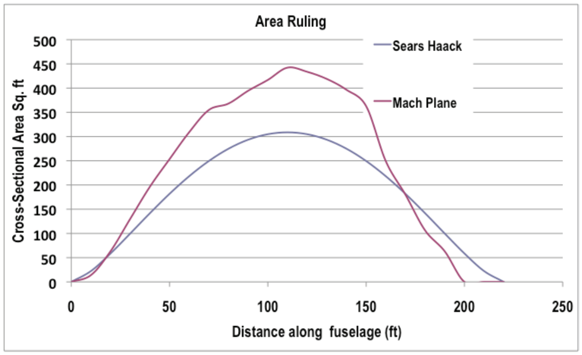


Figure 3: Mach 1.4 conical surface area distribution vs. Sears Haack cross section distribution

The point of the above exercise is to show that a liquid hydrogen-fueled SST can be designed for the 200-passenger, 8000km requirements to conform to the Sears-Haack area distribution. This shape is optimized for transonic conditions, leaving a substantial safety margin because the drag coefficient decreases at supersonic Mach number. This allows us to predict the highest wave drag that should be allowed. Issues and solutions in using liquid hydrogen[17] have been considered elsewhere.

8. Aerodynamics at Supersonic Cruise

Skin friction drag is calculated from the Boeing flat plate correlation for turbulent compressible flow[18]. Reasonable choices of wing loading and spans, give moderate aspect ratio. With the Jet-A SST, to keep the structure weight fraction above 0.27, the payload fraction had to reduced to 9.2%. The range of 8000 km is 20 to 25% greater than that of a Concorde. In contrast, the LH2 SST achieves a payload fraction of 27.5%, even with the structure fraction increased to 30%. Hydrogen generates about 3.8 times as much heat as Jet-A fuel does, even before accounting for the higher thermal efficiency of a hydrogen jet engine due to higher temperatures.

9. Hydrogen Drag Penalty

The choice of a 4.66m (15.3ft) diameter fuselage is conservative, and probably provides substantial volume for hydrogen storage above. However, the additional fuel storage volume for hydrogen beyond that required on the Jet-A craft was found by iteration. The wave drag penalty of including this excess volume brought the total drag coefficient to 0.0394 for the LH2 SST versus 0.027 for the Jet-A SST. Thus the upper bound on the "hydrogen penalty" in drag is a 50% jump in total drag coefficient. However, being substantially lighter, for the same payload and wing loading, the total drag of the hydrogen SST is only 55% of that of the Jet-A. So there is no "hydrogen drag penalty". Other designs were considered, including a Blended Wing-Body and an Oblique Wing. These posed difficult challenges to the Sears-Haack based approach for determining a benchmark calculation. An actual SST design will likely use Blended Wing Body concepts to reduce interference drag and engine noise. Table 2 vindicate the critics of the SST in that a conventional Jet-A fueled 8000km (5000mile) SST is not viable, regardless of noise issues.

10. Seat-Mile Fuel Costs

Airline annual reports circa 2003 indicated that fuel was roughly 20% of total costs (and therefore of averaged cost per ticket). With a sharp increase in jet fuel costs, and cost-cutting in other areas, we assume that fuel costs are now between 30 and 40% of total costs. Below, we estimate only the fuel costs, and the carbon costs attributable to the fuel. Figure 4 considers what happens as the cost of hydrogen fuel varies. This cost is expected to come down with improving technology, infrastructure and market acceptance, because hydrogen supplies are unlimited. It is left as the independent variable.

The use of seat-miles and cost per gallon rather than their metric counterparts in Figure 4 is intended to make it easier for the reader used to these common economic parameters. The seat-mile fuel cost of the LH2 SST is the slanting line. The short horizontal lines mark various levels. The lowest is the seat-mile cost for a long-haul airliner of the Boeing 787 class, with 250 passengers carried for 8000 miles (12,800km), at the current Jet-A price of \$2.237 per US gallon[19] (\$0.59 per litre) as of April 2010. This is 3.12 cents per seat mile, which the LH2 SST can match only at a hydrogen price of \$0.66 per kg.

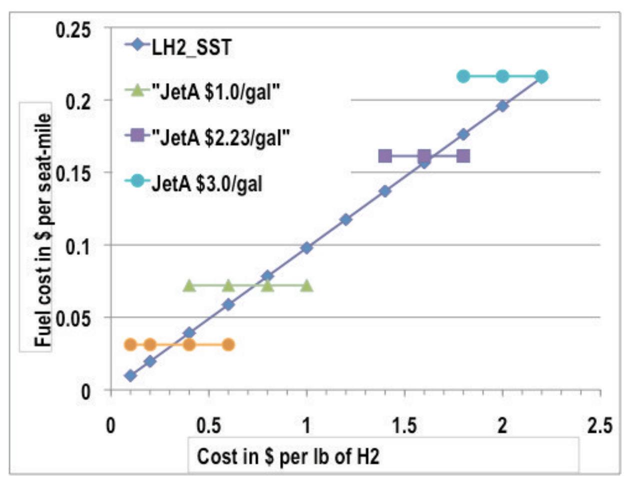


Figure 4: Fuel costs per seat mile.

The next level up is the seat-mile fuel cost of 7.2 cents (4.5 cents per seat-kilometer), of the reference

SST using Jet-A fuel, at the price level of $1/gallon that existed a few years ago. At $1.65 per kg of hydrogen, the LH2 SST would do better. At today's Jet-A price of $2.237 per gallon, the Jet-A SST fuel cost per seat-mile is 16.1 cents, bettered by the LH2 airliner at $3.52 per kg ($1.60 per lb) of hydrogen. The final level shown is for a Jet-A cost of $3 per gallon, where hydrogen can cost $4.84 per kg and still come out better. Today's hydrogen cost from the steam reforming process, liquefied and transported to the point of use, is estimated[20] to be $3.65 per lb. Thus the LH2 SST is already close to being cheaper than the Jet-A airliner at today's prices.

The calculation below uses the example of a transonic long-distance route to arrive at a reasonable comparison of ticket prices. The long route with its unique technology demands, low payload fraction and international issues, is best suited to capture both the true cost to the airline and the effect of marginal fuel costs, compared to the busy US-Europe routes where pricing may depend on many other factors. Assuming that seat-mile fuel cost is 40 percent of total airline cost (the upper bound as indicated above), the seat-mile ticket cost (excluding profit) for an "average" transonic airliner seat on the longest flights comes out to be around 6.24 cents. This works out to about $1000 for a round trip ticket for a 25,600 km (16,000 mile) round trip, a reasonable result given that the Atlanta-Dubai nonstop round trip ticket advance-purchase internet ticket price was around $1100 in December 2009. The LH2 SST at today's hydrogen prices would thus cost about 16.5 cents per seat-mile in fuel, and the round trip ticket would cost the airline $2640, marked up to a $3000 ticket price with economy-class service, but business-class seat room. It is our claim that this ticket price is well within the acceptable range for many who value the comfort and the reduction in flight time.

While it would be great to be able to fly supersonic the entire 12,800km (8000mile) distance non-stop, the paucity of such routes means that aircraft design for this application will probably await the success of the 8000km (5000mile) LH2 SST fleet.

Although the long-term seat-mile cost question is answered in the above, the shorter-term question of development cost remains. Here we could consider the carbon cost. At $20 per ton of CO_2, the transonic airliner adds a carbon cost of $0.00267 per seat-mile. A fleet of 500 LH2 200-seat airliners operating three 8000 km flights per week would save $208 million per year. Looking ahead a decade, over $2B of carbon savings can be reasonably projected, as a source of development funding for the SST.

Table 2: Parameters and results of the 3 conceptual designs compared

Concept	Jet-A SST	LH2 SST	Transonic Jet-A
Range, km	8000	8000	12800
Passengers	200	200	250
Cargo, tons	10	10	10
Payload fraction	9.2%	29%	22%
Gross weight, Metric tons	358	114	175
Wing Loading, N/m^2	4978	4978	5505
Aspect Ratio	6.24	9.33	6.02
CL	0.25	0.25	0.74
Engine T/W	11	11	11
L/D	9.16	5.4	15.27
Fuel Fraction	61%	37.7%	48%
Structure Fraction	27%	30.6%	27.2%

11. Conclusions

This paper argues for a new look at hydrogen-fueled supersonic airliners. Dramatic changes in demographics, globalization of trade markets and employment, and the maturing of expatriate worker communities, and the opening of the Communist Bloc nations and South Africa, all imply large and significant changes in the market for supersonic transport. A technical approach using the Sears-Haack body for minimum transonic wave drag is used to obtain a conservative comparison of the performance achievable using hydrocarbon (Jet-A_ and hydrogen-fueled supersonic airliners. Five main points are shown in this paper:

1. Hydrocarbon-fueled SSTs are not likely to be viable for an 8000 km range needed to reach an adequate number of busy non-stop destinations.
2. The aerodynamics of LH2 SSTs can be designed to be quite effective for 5000-mile range.
3. The "hydrogen drag penalty" of carrying a large quantity of liquid hydrogen for intercontinental flights, is non-existent, as these aircraft will have much lower drag than comparable Jet-A SSTs.
4. At today's costs of Jet-A and hydrogen, the LH2 SST is already more cost-effective than the Jet-A SST when carbon costs are included.
5. At today's costs of Jet-A and hydrogen, the viable ticket price on LH2 SSTs will be about 3 times that of advance-purchase transonic long-distance tickets. With mass-production efficiencies reducing hydrogen costs, it is realistic to expect LH2 SST ticket prices to come down to the level of today's transonic airliner tickets.
6. The carbon savings of a fleet of 500 LH2 SSTs would provide over $2B in a decade, as a justification of investment in LH2 SSTs.

12. Acknowledgments

The authors gratefully acknowledge the guidance and historical background provided by Dr. B. Kulfan at the Boeing Company.

13. References

[1] Heppenheimer, T.A., **The Concorde Supersonic Transport**. U.S. Centennial of Flight Commission, http://www.centennialofflight.gov/essay/Aerospace/Concorde/Aero53.htm Viewed 12/30/2009.

[2] Anon., **Tuploev Tu_144**. Wikipedia, http://en.wikipedia.org/wiki/Tupolev_Tu-144. Viewed 12/30/2009

[3] Wilhite, A.W., Shaw, R.J., **An Overview of NASA's High Speed Research Program**. ICAS 2000 Congress, Paper 112. http://icas-proceedings.net/ICAS2000/PAPERS/ICA0112.PDF

[4] Air Trasport Association, **World Air Travel and World Air Freight Carried, 1950-2008**. Data cited by Rodrigue, J-P. at Hofstra People. http://people.hofstra.edu/geotrans/eng/ch3en/conc3en/evolairtransport.html. Seen 04 Mar. 2010.

[5] Anon at PR Newswire, **Global Airline Industry Program**. MIT. Web. 04 Mar. 2010. http://web.mit.edu/airlines/analysis/analysis_airline_industry.html.

[6] Anon, **Northwest Airlines to Become First Carrier to Operate New Route Through Russian Airspace**. HighBeam Research: Online Press Releases. Web. Viewed 10 Mar. 2010. http://www.highbeam.com/doc/1G1-55395491.html

[7]Hubbert, Marion King. **Nuclear Energy and Fossil Fuels**. Report, 1956. Presented before the Spring Meeting of the Southern District, American Petroleum Institute, Plaza Hotel, San Antonio, Texas, March 7-8-9, 1956.

[8]Hirsch, Robert L. **The Inevitable Peaking of World Oil**. The Atlantic Council, Oct. 2005. Web. 14 Feb. 2010. http://www.acus.org/docs/051007 Hirsch_World_Oil_Production.pdf

[9]Anon, **The Hydrogen Economy**. Washinton D.C.: National Academies Press. 2004. Web. 22 Feb. 2010. http://www.nap.edu/openbook.php?record_id=10922&page=97. Web. 10 April 2010.

[10] Anon, **World Oil Production Close to Peak**. Online Posting. *Science Daily*. 7 Apr. 2007. <http://www.sciencedaily.com/releases/2007/03/070330100802.htm>. Viewed 14 Feb. 2010.

[11] Wood, J. H., Long, G.R., Morehouse. D.F., **Long Term World Oil Supply**. Energy Information Administration. 18 Aug. 2004. Web. 12 Feb. 2010. <http://www.eia.doe.gov/pub/oil_gas/petroleum/feature_articles/2004/worldoilsupply/oilsupply04.html>

[12] Jungmeyer, L. **EU Adds Aviation In Its Carbon Plans, Airlines Criticize Decision**. Online Posting. Environmental Leader Daily. 27 Oct. 2008. <http://www.environmentalleader.com/2008/10/27/eu-adds-aviation-in-its-carbon-plans-airlines-criticize-decision/>. Viewed Nov. 2009.

[13] Kulfan, R. M., **Application of Favorable Aerodynamic Interference to Supersonic Airplane Design**, SAE 901988, Oct. 1990

[14] Seebass, Richard and Agrow, Brian. **Sonic Boom Minimization Revisited**. AIAA 98-2956, 1998

[15] Roskam, J., **Airplane Design, Part I, II, V:** Preliminary Sizing of Airplanes, Preliminary Configuration Design and Integration of the Propulsion System, Component Weight Estimation. Roskam Aviation and Engineering, Ottawa, KS 1985

[16] Nikolic, V., Jumper, V.J., **Zero-lift Wave Drag Calculation Using Supersonic Area Rule And Its Modifications**. AIAA 2004-217, January 2004.

[17] Department of Energy, Feb. 2010 <http://www1.eere.energy.gov/hydrogenandfuelcells/storage/hydrogen_storage.html#liquid

[18] Kulfan, B.M., **Reynolds Numbers Considerations for Supersonic Flight**. AIAA 2002-2839, June 2002.

[19] Anon, **Jet Fuel Price Monitor**. International Air Transport Association, viewed April 13, 2010. http://www.iata.org/whatwedo/economics/fuel_monitor/Pages/index.aspx.

[20] Anon, **What are some approximate price numbers for hydrogen?** Chemical Process Business, CheResources.com, Chemical Engineering tools and Information, http://www.cheresources.com/questions/chemical_process_business-375.html. Viewed April 13, 2010.

THE (FUTURE) CUSTOMER VALUE IN THE FOCUS
An axiomatic design method combined with a Delphi approach to improve the success rate of new strategies, products or services

Vittorio M. FRANZELLIN
Faculty of Science and Technology, Free University of Bolzano
Bolzano, I-39100, Italy

Dominik T. MATT
Faculty of Science and Technology, Free University of Bolzano
Bolzano, I-39100, Italy

Erwin RAUCH
Faculty of Science and Technology, Free University of Bolzano
Bolzano, I-39100, Italy

ABSTRACT

The success rate of new strategies, products and services can only be increased by a consequent (pre-) orientation to the future markets and respectively the stakeholder's and the customer's benefit.
This paper presents a methodology which helps predicting new market trends and identifying systematically the customer needs and shows how to define on this basis a consistent vision and strategy for the company (from the high-level targets to the lower-level strategies).
The approach is based on the combination of different methodologies like the Delphi techniques and the Axiomatic Design aims to the systematic identification of share-/stakeholders and customer's benefits and requirements.
The Axiomatic Design is the cornerstone of the proposed methodology and is applied in different stages of the approach: first, to set up specific surveying / interviewing guidelines (that integrate the Delphi applications), and second to drill down the different strategic options and to set up new strategies and product or service concepts with the highest probability of success.
An example taken from the durable goods industry helps to illustrate the successful application of this approach (case study).

Keywords: Axiomatic design, innovation, product and service design, corporate strategy, strategy development, customer value

1. INTRODUCTION

Competitive pressure in volume markets has become stronger because of increasing economical and technical emancipation of the so called low labor cost countries [1]. Due to increasing customer expectations and growing international competition, companies are forced to move faster, to offer a huge product variety and at the same time to reduce the product life cycles.
The physiological (strategic) reaction of many small and middle sized enterprises (SME) to this development is the retreat into niche markets. However, lower production volumes, increasing complexity of the product range and shorter product life cycles render difficult the amortization of the investments, especially for SMEs. Flops on the market are hard to cope with and often even endanger a company's survival. [2].
The success rate of new strategies, products and services can only be increased by a consequent (pre-) orientation to the future markets and the customer benefit since the benefit drives the buying decision!
A company will only be able to define proper business strategies and to offer products or services with a better cost/benefit relation if the company is close enough to the market and to the customer and if it is capable to recognize and predict future markets trends and opportunities at the right time: not only the technological benchmark or the comparison with the competition are the "key factors" of success, but the more than ever the orientation at the (future) customers' needs.
This fact is proved by the fact that companies exploit ideas coming to a high percentage from external sources: almost 75% of the ideas come from market research and marketing as well as from other external sources, e.g. retail sector and suppliers [3]. It is also rather difficult to develop technologically oriented ideas to such an extent that it results in a tangibly noticeable customer benefit [2]. The past showed that in just 20% of the cases, this target could be achieved [4]. This indicates that the orientation to the customer and the really relevant issues (for the customer) is a basic condition for success in strategy, product or service development.
First of all, there has to be differentiated between the terms „customer benefit" and „customer requirement". Customer benefit can be defined as the quantitatively measurable and/or qualitatively perceived customer profit when purchasing a product or utilizing a service. Customer needs will be created if a supplier of such a product or such a service can communicate this benefit adequately [2]. The targeted search of ideas for successful products or services, and the corresponding strategies has to begin therefore with the identification of these customer benefits, and the related "benefit-driver". The "benefit-drivers" are influencing-variables related with the middle- to long term market prediction. (e.g. at a macro-level: political, demographical, economical, cultural, social trends and/or scenarios; at a micro-level the benefits-drivers are typically related with technology, design, materials, product- or service functions, ergonomics..).
On the other side customer requirements are particular characteristics and specifications of a good or service indicated

by a customer. They are usually identified by surveys or by market analyses [5]. The customer requirements are thus directly related to the respective customer benefit, which possibly could also be created by another (competitive or substitutive) product or service [2]. They can be also derived systematically from an identifiable customer benefit. The list of customer requirements (briefing) determines the direction of the product development onto which the functions and materials of the product are being developed.

The prediction of new trends and future markets, the identification and communication of the customer benefit is often very difficult especially in the area of durable goods.

Regarding the pursuit of customer benefits, the automobile industry, for example, discussed at length this problematic some years ago. A number of methods were developed for this topic, such as the QFD (Quality Function Deployment) or the Conjoint Analysis [6, 7].

All these methods can be meaningfully applied where customer needs regarding an already existing product- and/or service concept have to be evaluated [8]. However, these methods are insufficient or even not deployable for novel products, services or product/service combinations: after all, to a large degree, it is about ascertaining what the customer would like, but what does not yet exist in this form. How should he or she then be able to verbalize such a requirement? Even if the alternatives are obvious or known (which is an important condition for the functioning of the Conjoint Analysis, which tries to evaluate preferred samples by a comparison in pairs), the comparison in a panel or a group of people can lead to mistakes, as the Arrow's Impossibility Theorem shows [9].

On the other side, the systematic definition of a coherent customer and market-oriented strategy-framework (including e.g. commercial-, industrial-, financial- and corporate guidelines) is hard to achieve but it is crucial for a company to have clear stated goals and visions and to perform strategy and tasks accordingly.

Using a well-structured / oriented strategy, a company can achieve its goals faster than if the work was carried out in a less structured manner. According to Hacker (1998), there is a well-established idea that the success of an enterprise depends not only on its strategy planning, but also on how this strategy is implemented, monitored, and adjusted. Sull (1999) stated that many good companies fail due to an inability to take the appropriate action, while Nordlund (1996) found out that too many strategies lack action plans to fulfill their high-level goals. Usually, the company grows based on a specific strategy but fails monitoring and predicting environmental changes and adjusting strategy.

There are very few tools for customizing and designing a strategy to a company-specific and detailed level [10].

Among the approaches suggested in the management literature regarding to what should be implemented in a strategy, the most important ones are: Porter's five market forces for evaluating market attractiveness [11] , the "generic strategies" [11], the value chain analysis and its nine activities for increasing customer value (Porter, 1985); Ansoff's "growth strategies" (Ansoff, 1957, 1965); Andrew's "SWOT Analysis" and his strategy emerging from the alignment between "environmental opportunity" and "corporate capability" (Kennth R. Andrews 1965); the Prahald's and Hamel's "core competence focusing" (also focus on specific factor(s) that provides customer benefits, is(are) not easy to imitate, can be leveraged to many products and markets) for improving competitive advantages [12], 1990); Ghemawat's way to achieve competitive advantage using positioning analysis, level of flexibility and sustainability by commitment to sticky factors (i.e. untradeable, specialized, and durable) [13]; Erickson's and Shorey's performance pursuit by defining a strategy for high-level stakeholders and then trim it to fit important business processes, organization and resources [14]; Kotler's headquarters driven business units strategies [15]; Kerin's "aided matrix models" to suggest generic strategies depending upon the level of market growth rate and relative market share, or market attractiveness vs. business strength (Kerin et al., 1990); the Peters's and Waterman's 7-S framework and the related co-dependency between strategy successful implementation and company's structure, system, know-how, staff and values (7-S framework from Peters and Waterman Jr., 1982); et al.

The strategy design and (solving-) processes suggested in literature are to some extent contradictory, and focus mostly on high-level decisions [10]. Some researchers praise the concept of strategic business units (e.g. Hax and Majluf, 1996), whereas others prefer the principle of core competences as an alternative (organizational) way [12]. Many of the mentioned techniques also stress that a strategy has to be applied at all levels of the company, but few of them suggest how this should be realized.

Tools and frameworks for strategy-designing to fit companies at all levels are less frequent in management literature.

The first challenge therefore lies in the identification of a potential area of benefit, in which benefits can be derived and be translated into (product- or service-) demands. The second one is to initiate and to manage this innovation process into a coherent decisional framework that consider and define the strategy, according to the top-down approach of Hayes and Wheelwright (1984), at the corporate, business and functional level (marketing, sales, manufacturing, R&D, finance, et cetera).

How to reduce / manage complexity by SME's all-round strategy review?

How to recognize in a very early stage potential risks and / or "tactical circular reference" (e.g.. by the product / brand positioning, by the sales channels, by investments policies…).

How to formulate lower-level strategies respecting the high level goals, and also guarantee the link between tasks and stated goals and vision?

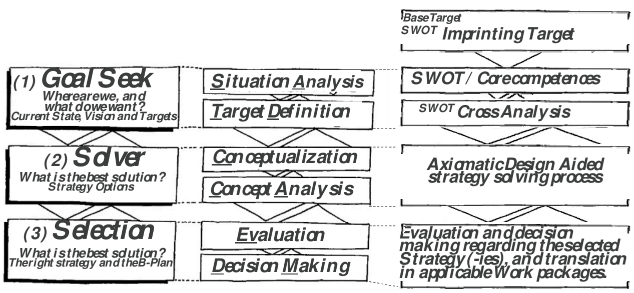
In this paper we suggest the use of Axiomatic Design (in the following referred to as "AD") as a tool for tackling the problems stated above. The first results of this research field, only focused on the product or service AD-aided development, were published in 2008 as CVF I Customer Value in the Focus Methodology [2]. The challenge of this new research chapter is now to define a structural framework for the market prediction, and the consequent definition of customer and market oriented corporate, business and functional strategies (Customer Value in the Focus II, in the following referred to as "CVF II").

2. METHODS, IDEAS AND SOLUTIONS

The developed and as CVF II (Customer Value in the Focus) proposed methodology is a collection of guidelines studied and tested for the prediction of future markets and customer needs, and for the systematic design of new company goals and strategies (from the corporate strategy level to the product or service ones).

Build up on a base Züst approach [19], the purposed framework foresees, according to Figures 1 and 2, three main work steps: (1) the "Goal Seek" (Where are we and what do we want?), consisting in a "Situation Analysis" based on a upgraded SWOT methodology, and the "Target Definition" (Company's Vision and Targets); (2) the "Solver" (What is the best solution?) consisting in two AD-Aided tasks: "Conceptualization" and "Concept Analysis" aims to strategy options design; and the (3) "Selection" (What is the best solution?), consisting in "Evaluation" and "Decision Making", also a quantitative and qualitative statement about the designed strategy options and the structuring of consequent implementation work packages.

Figure 1. CVF II (Customer Value in the Focus) Framework



This framework (CVF II) combines a set of base techniques and methodologies in an innovative way: many of them are known and reviewed in the literature, such as the SWOT Analysis (Andrews 1965 et al.), or the Delphi forecasting techniques (Dalkey & Helmer, 1962, Hill & Fowles, 1975; Linstone & Turoff, 1975; Lock, 1987; Parente´ & Anderson-Parente´, 1987; Stewart, 1987; Rowe, Wright & Bolger, 1991, Gordon 1994 et al.), dedicated to the "Situation Analysis" and other (in this context) rather diffused methodologies as AD (Axiomatic Design – Suh, 1990).

Figure 2. CVF II Framework – extended view

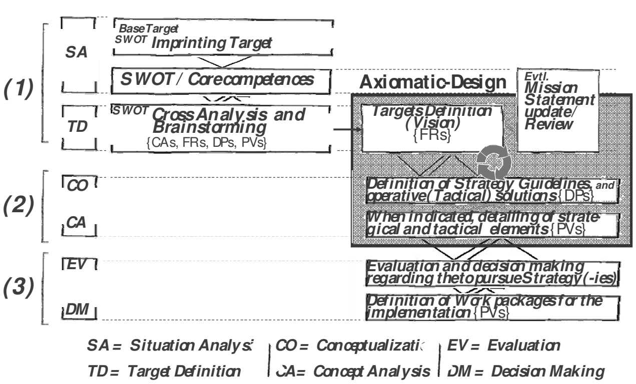
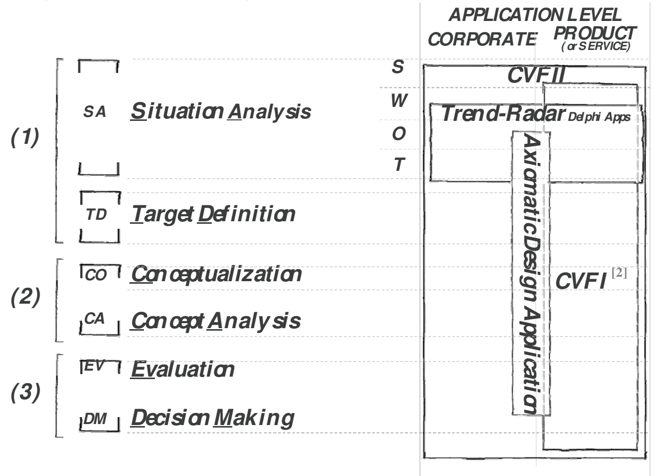


Figure 3. CVF II - Tooling "Bill of Material"



Especially the first two steps of the CVF II methodology require specific explanation and examples:

Step 1.1: (1) Goal Seek > (1) SA Situation Analysis

According to Martin and Kar [16] leading to the statement of business mission and goals, a careful analysis of customer needs and expectations represents one of the critical success factors of the strategy development. According to Certo (1993), strategy development begins with environmental analysis. In this analysis, all external and internal factors affecting the business unit are considered in a medium to long term perspective. Socio-cultural aspects, workers skills, governmental laws, and environmental considerations need to be analyzed accurately.

The Situation Analysis consists in a SWOT Analysis, upgraded with Delphi forecasting techniques, and applied for the identification and the prediction of: (a) future trends and customer "benefit drivers" related with market, competitive environment, legal – social – cultural – economical and technological issues; (b) customer needs, and the potential of new product or service concept or development directions [2].

According to Dalkey and Helmer (1962), the Delphi technique is as a procedure to "obtain the most reliable consensus of opinion of a group of experts… by a series of intensive questionnaires interspersed with controlled opinion feedback". In particular, the structure of the technique is intended to allow access to the positive attributes of interacting groups (knowledge from a variety of sources, creative synthesis, etc.), while pre-empting their negative aspects (attributable to social, personal and political conflicts, etc.). From a practical perspective, the method allows input from a larger number of experts geographically dispersed.

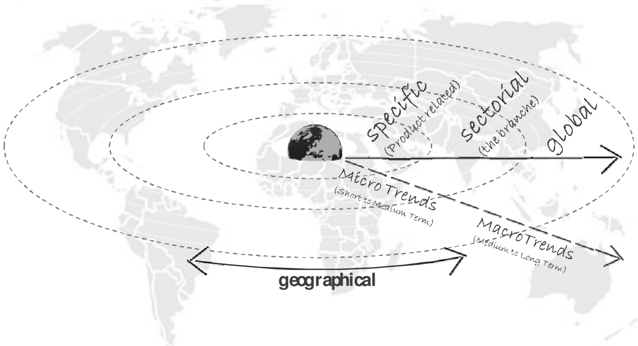
According to Rowe and Wright [17], Delphi is not a procedure intended to challenge statistical or model based procedures, against which human judgment is generally shown to be inferior: it is intended in judgment and forecasting situations in which pure model-based statistical methods are not feasible because of the lack of appropriate historical / economic / technical data, and this where some form of human judgmental input is necessary (e.g. Wright, Lawrence & Collopy, 1996).

Four key features may be regarded as necessary for defining a procedure as a "Delphi". These are: anonymity, iteration, controlled feedback, and the statistical aggregation of group response. Anonymity is achieved through the use of questionnaires and / or anonymous interviews. The iteration of the questionnaire over a number of rounds, the individuals are

given the opportunity to change their opinions and judgment without fear of losing face. Between each questionnaire iteration, controlled feedback is provided through which the group members are informed of the opinions of their anonymous colleagues.
An examination of recent literature, for example, reveals how widespread the use of Delphi is: applications in very different areas such as health care industry (Hudak, Brooke, Finstuen & Riley, 1993), marketing (Lunsford & Fussel, 1993), education (Olshfski & Joseph, 1991), information systems (Neiderman, Brancheau & Wetherbe, 1991), and transportation and engineering (Saito & Sinha, 1991).

The results of this first work-step (1) “Situation Analysis” consist in a traditional SWOT Analysis report, combined with a “Trend-Radar” based on the Delphi-forecasting results (see Figure 4) and a core-competence profile edited according to Prahald and Hamel [12], namely on the results of specific customer surveying processes (due to the direct relation with customer benefit and competitive advantages position).

Figure 4. Trend Radar reporting framework



Step 1.2: (1) Goal Seek > (2) TD Target Definition
The above described Situation Analysis that includes an analysis of internal and external strategic factors (Step 1.1) affecting the organizational performance, inputs from the stakeholders and the market needs guide now iteratively the definition of the corporate goals (Step 1.2) and strategies (Step 2.1) using AD.
According to Suh (1990), Nordlund (1996), Engelhardt (1998), Lobo, Cochrand and Lima (2000) [18], Martin and Kar (2002) [16], Schnetzler, Sennheiser and Schönsleben (2006) et al. Axiomatic Design can be used as a tool for the design of non engineering design object, such as technology strategies, business plans, and organizations.
AD differentiates four so called Design Domains: the Customer Domain describes the so called customer-benefit attributes (CAs; customer attributes), the Function Domain deducts from there the functional demands (FRs; functional requirements), the Design Domain provides Design Parameters (DPs) for the implementation of the FRs, whose transformation into processes shall be secured by the Process Variables (PVs) in the Process Domain [20]. The essential core of the Theory of AD is represented by two axioms, the Independence Axiom (1st axiom) and the Information Axiom (2nd axiom), which represent a necessary and sufficient condition for a “good” design of a product or a system. For this purpose, FRs and DPs are mathematically shown as vectors {FR} and {DP}.

The Design Matrix describes the relation between the two vectors:

$$\{FR\} = [DM]\,\{DP\} \tag{1}$$

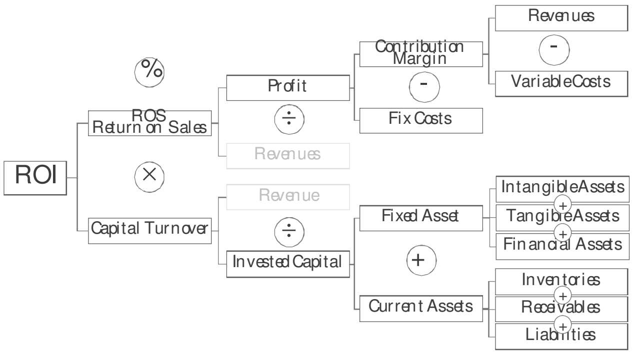
where

$$DM_{ij} = \frac{\partial FR_i}{\partial DP_j} \tag{2}$$

The first axiom demands the independence of the functional requirements (FRs). A potentially good design exists if exactly one Design Parameter (DP) can be found to fulfill the allocated FR without influencing the other FRs. To fulfill the Independence Axiom, the Design Matrix must be either a diagonal or a triangle matrix. In the case of a diagonal matrix, it is called an uncoupled design. This represents the ideal case, as every FR can be fulfilled with exactly one DP without being in any interrelation whatsoever to other FRs. In triangle matrices there is a so called decoupled design. These functions can only be fulfilled independently from each other by adhering to a certain sequence. All other cases represent a (badly) coupled design [20].
According and expanding Engelhardt and Nordlung [10], by adapting AD to the strategic design, the following terminological parallelism (respectively in engineering and business applications) can be assumed: Customer Attributes (CAs) = Customer / Shareholders needs, Functional Requirements (FRs) = Goals (∑ Goals = Mission and Vision); Design Parameters (DPs) = Strategies; Process Variables (PVs) = Activities.
The next step in strategy development is setting the directions that will guide the enterprise: the mission and the strategic objectives (FRs).
On this base, the logical and mathematical framework of the Axiomatic Design Theory helps to explore and to structure systematically all the goals (FRs) and the strategic options (DPs) (e.g. at commercial, product and service level) and to identify, at a very early stage, possible couplings (inconsistencies) between the different targets and solutions.
The passage between the function, the design and the process domain can very systematically be developed by the two axioms and the underlying methodology. In contrast, Nam P. Suh does not present an uniform methodological approach for the identification and translation of the customer attributes (CAs Customer or Share/ Stakeholders Needs). The analysis of the many examples which are meant to prove the validity of the axioms, does, however, show a logical pattern at the identification of the customer benefit attributes: from a purely economical point of view, the benefit is connected with a measurable value generation. The latter consists of a different perception of the value term for each user or group of users / shareholders.
The shareholder of a (profit-oriented) company measures the benefit of a strategy drawn upon to the extent to which this contributes to the increase in the value of the goodwill of the company (at least in medium and long term). According to the respective point of view, the goodwill can be illustrated by the Economic Value Added (EVA) or by the Return on Investment (ROI, see Figure 5).
Both methods illustrated are based on analytically-mathematically connected systems of key figures, which help to drill the benefit “goodwill” down into the smallest units.
The individual “leaves” of the respective key figure trees represent hereby ideal approaches for “benefit creating” strategies, products or services as partial benefits of the achievement of an increase in goodwill. For this purpose it is

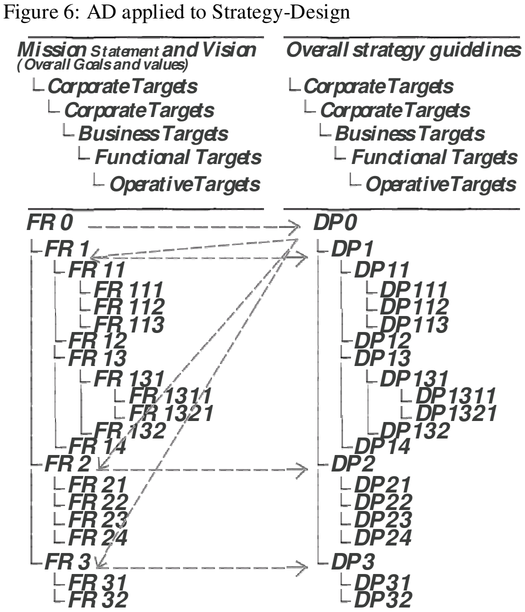
sufficient to calculate the target "earnings before tax" as the difference between contribution and fixed costs.

Figure 5: The Shareholder benefit logic (CAs): The ROI tree



On the top level therefore, the "father" functional requirement FR-0 can be directly attributed to the shareholder benefit "profit maximization" and be defined as follows:
FR-0 (Goal 0) Sustainable high profitability
On this level, FR-0 can be directly attributed to a design parameter DP-0, which is illustrated as follows:
DP-0 Define a Corporate, business and functional strategy for a high and sustainable profitability
The first design round explores now according to the ROI's mapped CAs (Customer, respectively Shareholders needs), the first level of company goals (FRs), which can be displayed as follows:
FR-1 (Goal 1) Increase Sales Volumes
FR-2 (Goal 2) Optimize the costs
FR-3 (Goal 3) Minimize the capital employed
By defining the Target structure on this top level other factors can also be considered: e.g. overall company / shareholders values, general purpose, existence reason, social-cultural principles.. . In this case the tree structure can be opportunely adapted.
As shown in Figure 2 the Target Definition process (Step 1.2) and the Solving Process (Step 2) are iteratively combined (AD "Zig-Zagging" process).

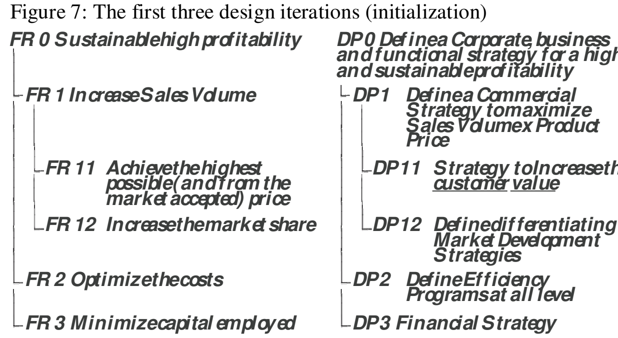
Figure 6: AD applied to Strategy-Design



The significance of the requirement for a targeted strategy design is, however, still too weak on the top specification level. In line with the methodology of the Axiomatic Design, the FR – respectively DP- tree is to be extended onto lower levels by the so called "zigzagging", as explained in the following.

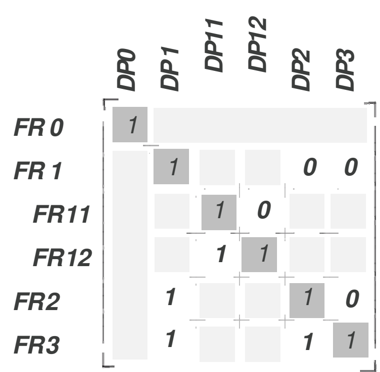
Step 2: (2) Solver > (1) Conceptualization
Strategy development based on AD starts also with setting high-level goals, and then corresponding strategies are defined to achieve these goals.
The strategic design process progresses from a system level to levels of more detail. High-level goals and the corresponding strategies are decomposed into more detailed sub-goals and strategies in terms of a design hierarchy. The decisions at higher levels affect the statement of the goals at lower levels (see Figure 6).
During strategy development, a strategic design process includes high-level decisions that make up the corporate level strategy and goes on to levels with increasing details. The business and functional level strategies are formed in the lower levels of decomposition.
At each level of the strategy development, there exist a set of goals. Before a certain goal is decomposed, the corresponding strategies must be determined. Once a business goal can be satisfied by a corresponding strategy, that goal can be decomposed into a set o sub-goals, and the zigzagging process is repeated iteratively. This process of mapping and zigzagging must continue until the design is completed. The first three iterations are shown in the following Figure 7.

Figure 7: The first three design iterations (initialization)



The information generated during mapping is captured in a strategic design matrix, which shows the relationships between each goal and strategy (Figure 8).

Figure 8: The first three design iterations – the design matrix



This matrix shows a decoupled design: these goals and strategies can only be fulfilled independently from each other

by adhering to a certain sequence: in our example the "cost optimization" can be applied just after "volume increasing", and after "market and product development"; "Financial Strategy" as last.
The mapping process between the domains can also be expressed mathematically in terms of the characteristic vectors that define the design goals and the design solutions.
Design Matrices are set up for all the goal-strategy relations at the different levels in each branch of the goal-strategy tree. Knowledge for configuring the Design Matrices comes from the Delphi interviews (see Step 1.1), cross functional groups, existing process descriptions etc.
The strategy has to be consistent all the way from high-level company targets and visions (FRs), down to the operative tasks carried out by the employees (DPs). The company's personality, culture, philosophy, organization and areas of business provide company-specific needs. Those needs have to be considered when designing (customizing) a strategy. In addition, the results of the SWOT simple and cross analysis (Stages 1.1) as a cluster of CAs, FRs, DPs and PVs, and the defined goals must also set priorities. After the first initializing-rounds, the next detailing must be effected in the light of the individual market-sector and application. This is now demonstrated below with a practical example.

3. ILLUSTRATIVE EXAMPLE

The company concerned is a medium sized, worldwide operating durable goods producer offering products and solutions of the highest quality. The company is market leader in their sector and wants to review the own middle to long run targets and strategies (for the next 10 to 15 years).

Step (1)

The Situation Analysis, realized through one-rounded Delphi campaign (25 customer and sector's experts worldwide involved – see e.g. a std. Delphi report Figure 9), combined with traditional internal analysis helps to identify trends and future scenarios, regarding the competitive situation, the customer needs, the products and technology development (see a Trend-Radar excerpt - Figure 10).

Figure 9: Delphi response standard report

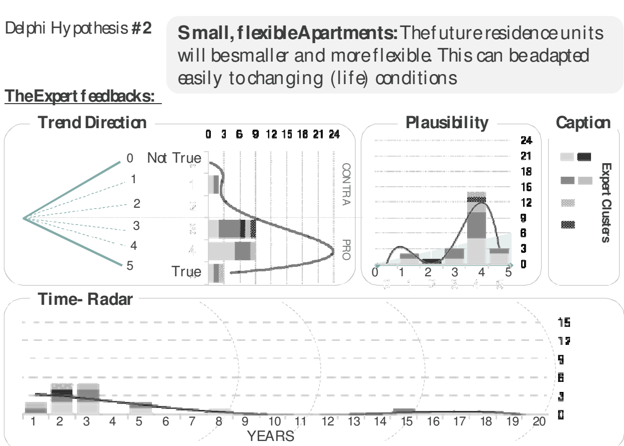
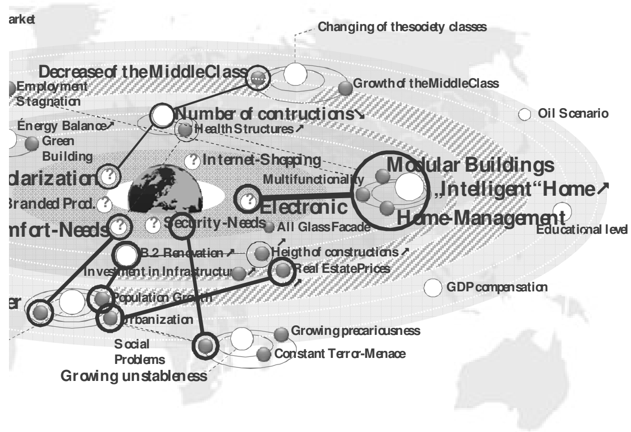


Figure 9: Trend Radar report sample excerpt



All this inputs and information can be capitalized for the following strategy design using e.g. a SWOT-Cross analysis, also crossing and studying the different combinations between identified Strengths and Opportunities ("attack"), Strengths and Threats ("prevention"), Weaknesses and Threats ("defense"), Weaknesses and Opportunities ("improvement").

Step (2)

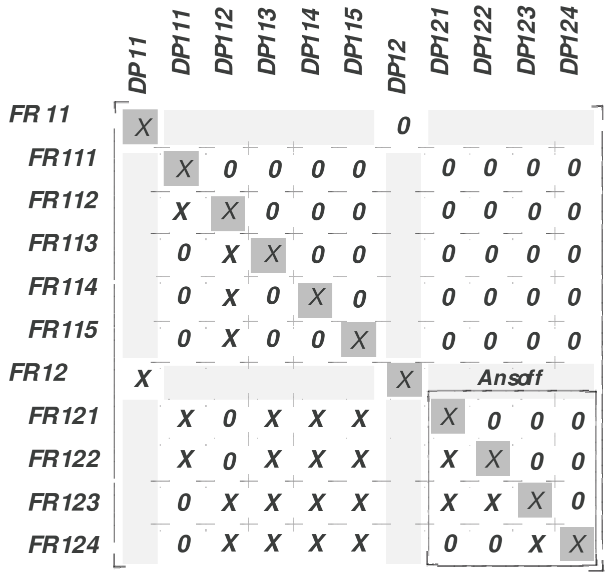
In the Solving process, all the "random" CAs, FRs, DPs and PVs identified above (external, internal analysis, future trends / customer need predictions…), can be distilled together with the company management experience and know-how, structured and fulfilled using AD multi-level decomposition methodology.
According to the first three design iteration described above (see Figure 8), as illustrative example, in this paper, only the decomposition of the FR 1 Increase the Revenues > FR 11 Achieve the highest possible price, and FR 12 Increase the market shares > is shown.

Figure 10: Illustrative-Example Decomposition

FR 11 Achieve the highest possible (and from the market accepted) price	*DP 11 Strategy to Increase the customer value*
FR 111 Constant good and right quality at different price levels	*DP 111 Product Quality Strategy*
FR 112 Higher (customer oriented) product attractiveness	*DP 112 Develop transnational and customer oriented products strategies*
FR 113 Competitive delivery times	*DP 113 Develop customer oriented logistic solutions*
FR 114 Binding and high delivery performance	*DP 114 Guaranteed delivery performance*
FR 115 Political correct company image (environmental, locations…)	*DP 115 Update and communicate the company image*
FR 12 Increase the market share	*DP 12 Define differentiating Market Development Strategies*
FR 121 Market Penetration	*DP 121 Develop present customer and acquire new customer from competitors in current markets*
FR 122 Market Development	*DP 122 Expansion in new geographical markets with current products*
FR 123 Product Development	*DP 123 Customer-segment oriented product development without cannibalization of existing products*
FR 124 Diversification	*DP 124 Expansion in new geographical markets through new products or services*

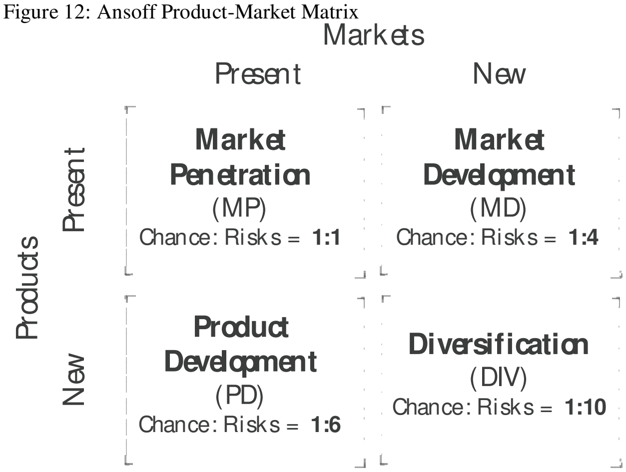
The corresponding related design matrix (Figure 11) shows a decoupled design, also a potentially good strategy design subordinated to the displayed implementation sequence.

Figure 11: Illustrative-Example Design Matrix



According with Ansoff (1957, 1965) and Becker (2001), the FR12 Increase the market share can be achieved only through a well defined strategic roadmap. Considering in fact the four sub-FRs, directly derived from the Ansoff's Product-Market Matrix (see Figure 12), FR121 Market Penetration (MP), FR122 Market Development (MD), FR123 Product Development (PD), FR124 Diversification (DIV -Horizonatal, Vertical or Lateral), AD helps with the evidences of the Design Matrix to demonstrate the better strategic way to go.

Figure 12: Ansoff Product-Market Matrix



Analyzing in detail the independency between the different FRs and the related DPs (see Figures 13 and 14):

DP121(MP) affect FR122(MD)

The implementation of a Market Penetration strategy can influence the Market Development Strategy, e.g. through the choice of marketing activities and promotion, which should be discussed and coordinated between MP and MD.

DP121(MP) affect FR123(PD)

DP122(MD) affect FR123(PD)

MP influences the Product Development, e.g. an old product could be "reanimated" through a specific marketing campaign.

DP123(PD) affect FR124(DIV)

PD influences the Diversification, e.g. the improvement of current product interfaces, can simplify an horizontal diversification.

Figure 13: Becker's Product-Market Matrix Strategic-Roadmaps demonstration through Axiomatic Design (Part I)

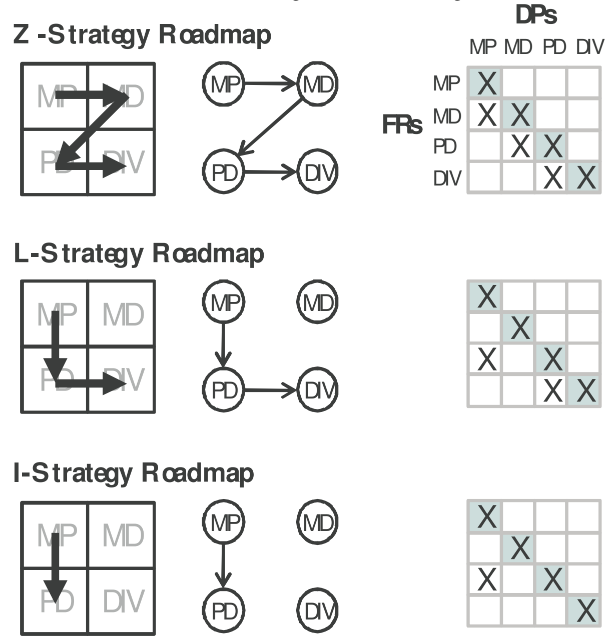
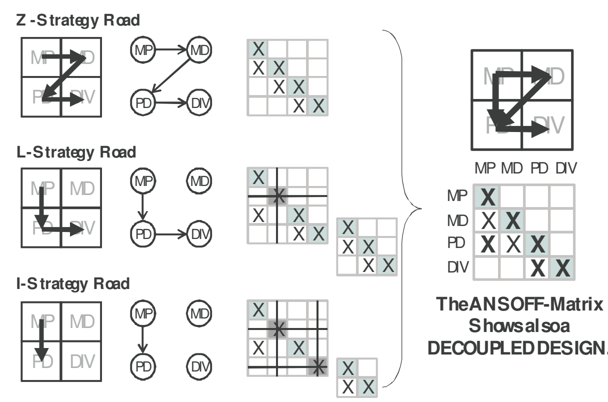


Figure 14: Becker's Product-Market Matrix Strategic-Roadmaps demonstration through Axiomatic Design (Part II)



This AD-Aided design process enables also the design of a total strategy with no contradictory goals. Feedback from company experts, strategic team-members and other stakeholders provides new knowledge and enables further trimming and redesign of the strategic output.

4. RESULTS AND CONCLUSIONS

Solving the equation systems expressed in the Design Matrices creates a process sequence that minimized unproductive iteration and reduces rework, thereby speeding up implementation. The final result can be produced in form of a process /or a flow chart that displays the different relations (time / cause-effect) between certain activities and others in time.

Since Axiomatic Design is a top-down design method, it is very suitable for consistently transferring high-level company goals and visions down to specific projects. This breakdown of abstract goals into more concrete ones improves company efficiency (Robbins, 1994). It also allows the firm to adopt more rapidly a new strategy by defining how low-level goals and strategies as well as tangible activities are related to overall strategic vision, which improves employee participation and communication. In the case of a business strategy one might find that the product strategy and the marketing strategy (strategic e.g. product segmentation: quality image vs. low cost performance) are fully coupled: in this cases the decoupling requires specific statement and strategic decision, in order to avoid dangerous “short-circuits”. According to Engelhardt and Nordlung [10] other examples of the effects of a tight coupling between what the company aims for and what it actually does could be: (1) less resources needed for achieving the goals, (2) selection of proper technologies for chosen markets, (3) hiring procedures focused on getting the employees needed for planned tasks, (4) motivated employees that know the reason for what they are doing, etc.

AD helps the designer also to set up design equations that express the relationships between goals, strategies and activities. The framework provided by AD gains simplicity by having a one-to-one mapping between activities, strategies, and goals.

AD proved to be a very useful tool for designing and customizing a technology strategy for the company. The approach also helped to identify strategies and activities that from unnecessary coupling are candidates for redesign.

In this paper, a methodology based on the Theory of Axiomatic Design combined with Delphi techniques has been demonstrated, which supports companies in their systematic design of new strategies, products or services. By means of consequent derivation of FRs and first DP hypotheses for the (future) strategies, products or services from a clearly defined share-, stakeholder or customer benefit description, a framework for guided expert surveys is being developed. The results of the survey show a clear picture of the future market and customer needs. On this basis, targeted and detailed developments of innovative products, respectively services can be commenced. The successful application of the methodology has been demonstrated in an example from the durable goods industry. Future research will be focussed on the validation and improvement of the approach by means of its application in different areas.

5. REFERENCES

[1] E. Kirner, O. Som, C. Dreher, V. Wiesenmaier, **Innovation in SME – Der ganzheitliche Innovationsansatz und die Bedeutung von Innovationsroutinen für den Innovationsprozess,** Karlsruhe, Project Report InnoSME, 2006.

[2] D. T. Matt, V. M. Franzellin, **The customer value in the focus** – an axiomatic design based method to improve the success rate of new product or service, IAMOT International Association for Management of Technology, Dubai, 2008.

[3] K. Kashani, J. Miller, T. Clayton, **A Virtuous Cycle: Innovation, Customer Value and Communication. Research Evidence from Today’s Brand-Builders**. Brand Innovation Study PIMS/IMD, 1999.

[4] R. Rüttimann, **Wodurch sich gute Innovationen auszeichnen.** Innovation Management, Vol. 1, 2007, pp. 12-17.

[5] D. Spath, D. T. Matt, S. Riedmiller, **The Communication Circle -A model for the support of a market-driven development and production of innovative products**. Proceedings of the CIRP Design Seminar, Berlin, 1998 pp. 27-35.

[6] D. Spath, D. T. Matt,D. Geisinger, **Marktorientierte Neuproduktentwicklung,** ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb, Vol. 93, No. 5, 1998, pp. 200-202.

[7] Gaul, W and Baier, D, **Marktforschung & Marketing Management (in German)**, München: Oldenbourg, 1994.

[8] R. Züst, D. T. Matt, D. Geisinger, R. Winkler, **A Chance of Paradigm for Planning Life-Cycle-Optimised and Innovative Products.** Proceedings. Int. Conference on Engineering Design ICED, Munich, Vol. 99, 1999, pp. 1081-1084.

[9] Hazelrigg, GA **Theoretical Foundations of Engineering Design**, Massachusetts Institute of Technology, 1998.

[10] F. Engelhardt, M. Nordlund, **Strategic Planning based on axiomatic design**, Proceedings of ICAD 2000 First International Conference on Axiomatic Design, Cambridge, MA, 2000.

[11] M. Porter, **Competitive Strategy, Techniques for Analyzing Industries and Competitors,** New York, The Free Press, 1980.

[12] C. K. Prahalad, G. Hamel, **The Core Competence of the Corporation,** Cambridge, Harvard Business Review, May-June 1990, pp. 79-91.

[13] P. Ghemawat, **Commitment: the dynamic of strategy,** The Free Press, New York: New York.

[14] T. J. Erickson, C. E. Shorey, **Business Strategy: New Thinking for the ’90,** Prism, Fourth Quarter, 1992, pp. 19-35.

[15] P. Kotler, **Marketing Management: Analysis, Plannning, Implementation and Control,** New Jersey, Prentice Hall, 1997.

[16] S. B. Martin, A. K. Kar, **Axiomatic Design for the development of enterprise level e-commerce strategies**, Proceedings of ICAD2002, Second International Conference on Axiomatic Design, Cambridge, MA, 2002.

[17] G. Rowe, G. Wright, **The Delphi technique as a forecasting tool: issues and analysis**, International Journal of Forecasting Vol. 15, 1999, pp. 353–375.

[18] C. Lobo, D. S. Cochran, P. C. Lima, **Strategy based on axiomatic design**, Proceedings of The Third World Congress on Intelligent Manufacturing Processes & Systems, Cambridge, MA, 2000.

[19] Züst, R., **Einstieg ins Systems Engineering: Optimale, nachhaltige Lösungen entwickeln und umsetzen**, Zürich, Orell Füssli, 1997

[20] N. P. Suh, **Axiomatic Design – Advances and Applications,** New York: Oxford University Press, 2001.

Performance Study on Organic and Nitrogen Removal of Leachate by a two-stage Oxic-anoxic Biological Aerated Filter (OABAF) System

Pik-Chi James FUNG
Department of Civil and Structural Engineering, The Hong Kong Polytechnic University, Hong Kong

Wai-Pui Victor Lam, Li Patrick Ma,
Ecolo Environmental Industries Ltd., Unit 23, 13/F., Block A, Wah Lok Industrial Centre, 37-41 Shan Mei Street, Fotan, N.T., Hong Kong

and

Dr. Chuen-Yu John Chan
School of Environmental Science and Engneering, Sun Yat-san University, Guangzhou, P. R. China

ABSTRACT

Biological aerated filter (BAF) is an attractive process option, particularly when low land usage is required. In this study, a two-stage oxic-anoxic biological aerated filter (OABAF) reactor system was proposed for the enhancement of nitrogen removal in the treatment of simulated leachate. Laboratory experiments were conducted under various operation conditions with different carbon-to-nitrogen ratio (C/N ratio) and hydraulic retention time (HRT). Results of the present study indicated that the OABAF process reduced 57-66% NH_4^+-N removal efficiency with 85%–93% organic matter removal. It was found that when coral stone was used as filter media, this system can achieve a 97% organic removal and 94% ammonia reduction of simulated leachate at 2.4 g/hour organic loading, and 10.5 hrs HRT under the effective volume is 33% of the reactor volume; when biological plastic spores that were mixed with coral stone, this system can achieve a 86% organic removal and 62% ammonia reduction of simulated leachate at 18 g/hour organic loading, and 10.5 hrs HRT under the effective volume is 55% of the reactor volume. From these experimental data of study, it was concluded that the two-stage OABAF system was excellent in nitrogen removal performance and can be applied to treat effectively the leachate generated from refuse transfer station or landfill.

1. INTRODUCTION

Hong Kong is still one of the few metropolises which still adopt landfill method. 100% of the waste production in Hong Kong was landfilled. Transfer station is adopted as an economic mode of transporting MSW before the waste is transported to the remote landfill sites. In spite of many advantages, the generation of highly polluted liquid leachates can bring contamination of groundwater and soil contamination during landfill usage. Year by year, the risk of the cumulated landfill leachate has become a severe environmental issue. In Hong Kong, leachate is treated by activated sludge process. However, the conventional biological treatment processes always cause odour problem. Under the extreme high organic (BOD & COD) and nitrogen (NH_4-N) strengths, leachate removal efficiencies and operational conditions for treatment are usually unsatisfactory. In order to prevent the leachate in order pollution, it is imperative to improve the treated leachate in order to comply the effluent discharge standard to public sewers by an innovative method.

The proposed two-stage OABAF reactor system may be a highly effective biological treatment process that is developed on the basic technique of combined conventional activated sludge process and trickling filters in a system. It is a continuous operation of aerated biofilm reactor, where the biomass films are grown on small carrier elements (or media) that have a higher density than water, and are kept in the bottom with wastewater streams flowing up or down the reactors. This study was initially carried out under a research project, "*Pilot-scale Study on Treatment Performance of Leachate by Anoxic-oxic Biofiltration Reactor"* in September 2005, and was further carried on in a one-year master degree dissertation study in 2009. The research project was done in conjunction with Sun Yat-san University in 2009. The study was aimed to investigate and study the treatment performance on the TOC and ammonia nitrogen removal efficiencies of simulated leachate under various operating conditions through experimental works where biological plastic spores and coral were used as bio-carrier. The applicabilities and limitations of the OABAF system were also identified. The author would like to thank The Hong Kong Polytechnic University for its support.

2. MATERIALS & METHODS

2.1 Operating principles of the two-stage OABAF system

A two-stage OABAF process within both nitrification and denitrification steps were incorporated in the study. In the 1st oxic reactor, part of the carbonaceous substrates in the influent was adsorbed on the attached biomass of the media and accumulated within bacteria cells. Fast nitrification then occurred in this reactor and enabled the ammonia-poor liquid streams to go to the next reactor (1st anoxic column) with anoxic condition. The 1st anoxic reactor functioned as denitrification. Finally, the rest reactors in the subsequent stage (2nd oxic reactor and 2nd anoxic reactor) worked for purification to achieve a better effluent quality. The design of independent nitrification and denitrification can be led by the separation of heterotrophic bacteria and autotrophic bacteria. Recycling paths from 1st oxic reactor to the anoxic reactors were initially proposed. However, it was found that a single-stage system can also achieve a 98.4% organic removal and 79.4% ammonia reduction of simulated leachate at 5.4 g/hour organic loading, and 7 hrs HRT. Thus, it may conclude that the need of extra carbon source was neglected under a single-stage treatment; but, further attention needs to be undertaken for the successful utilization of a two-stage OABAF system.

The simulated leachate was used as the experimental influent. It was basically composed of a mixed carbon source and nutrient (ammonium nitrogen only). The composition contained 22.96g/L glucose, and 3.76-45.01g/L NH_4Cl. The simulated leachate was prepared everyday with concentrations of 300±5 mg/L TOC and 25-300±5mg/L ammonia nitrogen while NO_2-N and NO_3-N were negligible. Wastewater analytical tests were carried out to investigate the performance of organic and ammonium oxidations in varied concentrations of simulated leachate in this system.

2.2 OABAF system setup & operation

The study comprised two experimental phases and a design phase. *Initial Phase* received varied operating conditions to determine an optimum treatability to a four separated reactors system. *Final Phase* was then followed by using a bigger plastic reactor with different compartment and different bio-filter media.

In Initial Phase: The main system composed of four trapezoidal plastic biofilter reactors connected in series. Each reactor was 31.5cm in height with an effective volume of 3.2L. The lengths at top and at bottom were 26.5cm and 24.5cm respectively. The widths at top and at bottom were 17.5cm and 15.5cm respectively. The reactors were packed with plastic media (which is approximately 1.5–2 mm in diameter and approximately 600-900 m^2/m^3 on an average specific surface area) to a depth of about 26 cm at Preparative Phase, and then change to coral stone at Initial Phase. The density of packing media was approximately 2×10^3 kg/m^3 for plastic spore and approximately 3.6×10^4 kg/m^3 for coral stone respectively. In oxic reactors, air was provided to maintain the DO concentration of 2.5±0.5 mg/L through diffusers located at the bottom of the reactors. In contrast, DO was not provided in anoxic columns. The system was operated in two-stage cycle, i.e., step 1 and step 2, based on the direction of influent flow (see Figure 2-1). The HRT of each step was set under different flow rate in order to achieve the optimum treatment performance. The cycle operation has been repeated by switching influent flow. The activated sludge obtained from an oxic basin of Tai Po municipal wastewater treatment plant was used as the inoculating sludge in the operation. The system was operated at room temperature ranging from 19 °C to 22 °C, and averaged 20 °C.

In Final Phase: A fibre glass tank was divided into four cuboids biofilter reactors by clapboard. Small gaps were custom made intentionally, so that the water stream could flow through the gaps and enter the following reactor. At 1st stage, each reactor was 40cm in height with an effective volume of 10.3L. Its length was 40cm and the width was 21.5cm respectively. At 2nd stage, each reactor was 40cm in height with an effective volume of 8.45L. Its length was 40cm and the width was 17.6cm respectively. The reactors were packed with plastic media and coral stone to a depth of about 25 cm. The density of packing media was 3×10^4 kg/m^3. In oxic reactors, air was provided to maintain the concentration of DO of 7.5 ± 0.5 mg/L through diffusers located at the bottom of the reactors. In contrast, DO was not provided in anoxic columns (see Figure 2-2). The HRT of each step was set under different flow rate in order to achieve the optimum treatment performance. The cycle operation has been repeated by switching influent flow.

2.3 Sampling

In Initial Phase: Samples were collected from four bioreactors. Water samples were collected near bottoms (28 cm depth) at 1st stage oxic reactor and 2nd stage oxic reactor and at the surface (6 cm depth) at 1st stage anoxic reactor and 2nd stage anoxic reactor. The samples were collected from the discharging points of the reactors.

In Final Phase: The water samples were collected from four bioreactors respectively. Water samples were collected near bottoms (20 cm depth) at 1st stage oxic reactor and 2nd stage oxic reactor and at the surface (1 cm depth) at 1st stage anoxic reactor and 2nd stage anoxic reactor. The samples were collected from the discharging points of the reactors.

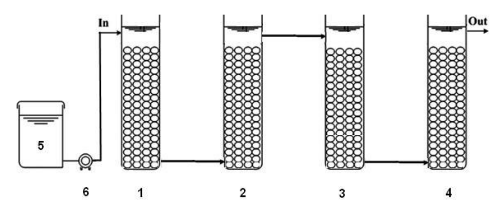


Figure 2-1 Schematic of a two-stage OABAF system: (1) 1st oxic reactor; (2) 1st anoxic reactor; (3) 2nd oxic reactor; (4) 2nd anoxic reactor; (5) influent simulated leachate tank; (6) influent pump

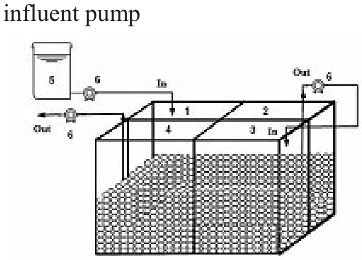


Figure 2-2 Schematic of a two-stage OABAF system: (1) 1st oxic reactor; (2) 1st anoxic reactor; (3) 2nd oxic reactor;

(4) 2nd anoxic reactor; (5) influent simulated leachate tank; (6) influent pump

3. RESULTS & DISCUSSION

3.1 Initial Phase analysis

The study at the Initial Phase was accomplished through 2 steps in 26 days. At step I (from Day 1 to Day 14), the C/N ratio, carbon and nitrogen concentrations were kept constant, different HRTs were used to find out the optimum treatment performance. The applied flow rates were ranged from 0.004-0.021 l/min, corresponding to HRT of 4-21h, respectively. The C/N ratio at this stage was fixed to 12:1 and the total effective volume/total volume ratio was 33%. At step II (from Day 15 to Day 26), the HRT was kept constant for 10.5 hours with the C/N ratio only being changed from 1:1 to 12:1. The total effective volume/total volume ratio remained at 33%. Coral stones were used as biofilter media for all the experimental works during the Initial Phase. Figure 3-1 shows the variation of ammonia, TN and TOC over time during oxic and anoxic treatment of simulated leachate. The mean values of ammonia, TN and TOC removal efficiencies in the influent and effluent of experimental works are given in Table 3-1.

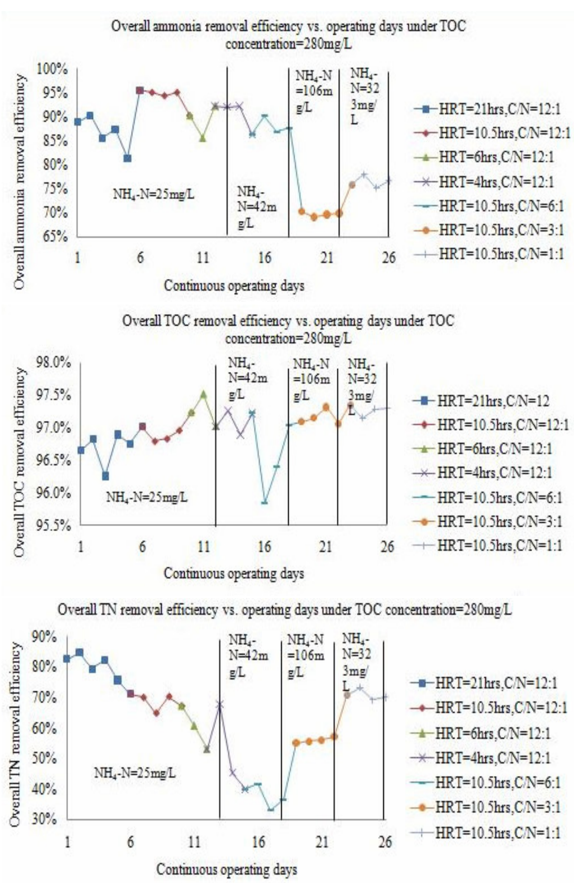


Figure 3-1 Profiles of overall ammonia, TN and TOC removal efficiency vs. operating hours of the oxic and anoxic treatment of simulated leachate in Initial Phase

Parameters	Initial
PH	7.75
Temperature (^oC)	20±2
C/N ratio	12:1
HRT (hrs)	10.5
Organic loading (g/hr)	2.4
Ammonia loading (g/hr)	0.4
Influent TOC concentration	278.4
Effluent TOC concentration	8.63
Influent NH_4-N concentration	25.2
Effluent NH_4-N concentration	1.25
Influent NO_3-N concentration	7.08
Effluent NO_3-N concentration	2.29
Influent NO_2-N concentration	0.004
Effluent NO_2-N concentration	0.015
Organic Removal Efficiency	96.9%
Ammonia Removal Efficiency	95.0%
Total Nitrogen Removal	69.0%

Table 3-1 Influent and effluent characteristics in Initial Phase

The experimental results at Initial Phase showed that:

- OABAF system is reliable and effective in removing nitrogen and organic matters - When the C/N ratio of the influent was 12:1, the ammonia nitrogen removal efficiency was consistently above 81% under all HRTs, and the final effluent ammonia nitrogen was 1.15-4.68 mg/L. The final effluent TN concentration at all HRTs of step I and step II was 12.5 mg/L and 79 mg/L on average, respectively. TOC removal efficiencies were in the range of 270–288ppm (average of 277.4ppm), which were generally high (up to 96.9±0.3%). During the 6 weeks operational period, it is found that under the total effective volume/total volume ratio of 33% and C/N ratio of 12:1 condition (when HRT is 10.5 hours), the OABAF system achieved the optimum nitrogen removal performance. 92.6% of ammonia and 60.8% of total nitrogen is removed, meanwhile, the overall TOC removal efficiency can also achieved to 96.88%.
- In all cases, there is a directly proportional relationship between C/N ratios and ammonia removal efficiency - The feed ammonia nitrogen concentration was in the range of 24-25 mg/L at step I and 25-335 mg/l at step II. C/N ratio had a significant influence on the ammonia nitrogen removal. When the C/N ratio of the influent was 12:1, the ammonia nitrogen removal efficiency was consistently above 81% under all HRTs, and the final effluent ammonia nitrogen was 1.15-4.68 mg/L. In the OABAF system, the oxic reactors acted as the main undertaker for ammonia nitrogen removal due to nitrification; its contribution to the ammonia nitrogen removal was about 72-96% at step I, however, as C/N ratio decreased, consequently, the removal efficiency decreased gradually and finally dropped to about 19% at step II. Correspondingly, the total removal efficiencies

of ammonia nitrogen were 85.2% and 71.1%, respectively. This is fairly easy to rationalize, as an increase in C/N ratios equates to a deduction in ammoniac toxicity and organic concentration in a set time period, which obviously tends to propel the activity of ammonia oxidizer.

- There is a correlation between HRT and ammonia removal efficiency - An increase in the former always corresponding with an increase in the latter. This could be explained by the following hints; at a smaller HRT, the stronger hydraulic pressure triggered biomass washout.
- A similar increase in TN removal efficiency is also observed as the flow rate increases in C/N ratios and HRT in most experiments - The final effluent TN concentration at all HRTs of step I and step II was 12.5 mg/L and 79 mg/L on average, respectively. Note here that most of the effluent TN was nitrate (nitrite concentration was almost zero) at the oxic reactors. In the operation of the two-stage OABAF system, the low effluent TN concentration would act as the benefit of its application compared with the experiment results. Comparing the average final effluent TN concentration of step I with that of step II, the final effluent TN concentration of step II was much lower than that of step I. This is important, as the main limiting factor for the oxic and anoxic treatments are likely to be the strengths of the hydraulic pressure, organic and ammonia concentration. This suggests that as C/N ratios and HRT increase, the deduction of hydraulic pressure and concentrations of organic and ammonia become a more important factor of bacteria growth.
- In 1st stage, NH_4-N in influent was converted into nitrate with over 70% on average when HRT is 10.5 hours. According to the nitrite concentration which was almost zero in the reactor, it can be inferred that nitrifiers in the 1st stage oxic reactor would dominate and nitrification was then enhanced. The enriched microorganism particles of OABAF, could contribute to the increase of nitrification efficiency. In this study, when C/N ratio is 12:1, the carbon removal efficiency is almost 95% in average. On the other hand, a decrease of the C/N ratio (Day 15 to Day 26) causes a significant change of TN and ammonia nitrogen removal parallel to the increase in TOC removal. So it may be concluded that a high C/N ratio of effluent caused predominance of nitrifying bacteria in growth competition between autotrophic and heterotrophic micro-organism population in the 1st stage oxic reactor;
- On average, 75% of influent TOC was bio-sorbed under the 1st stage process and the rests would be utilized as the substrate for denitrification reaction in the subsequent step. Most of ammonia nitrogen was not removed and the concentration of nitrite was almost zero. C/N ratio and HRT are the most critical factors controlling the distribution of denitrification activity. Any decreases of the HRT or C/N ratios leads to decrease of ammonia remove efficiency, and the parallel increase of organic removal;
- The effluent concentration of nitrate nitrogen in 1st stage anoxic reactor was about 4.65 mg/L on average. In contrast, almost unchanged effluent concentrations of TOC and NH_4-N were observed. This was due to the biosorption of organic matter and NH_4-N on the media occurring in this reactor during the operation period of the previous step. TOC bio-sorbed on the media would be utilized as the substrate for the denitrification reaction. Hence, effluent TOC concentrations decreased with time since those substrates would be locally limited within this reactor. This speculation would explain the observation that the denitrification performance was improved in 1st stage anoxic reactor. A significant decrease of nitrate concentration and the parallel slightly decrease of ammonia nitrogen indicated that most of the nitrate had been deoxidized to nitrite. According to the slightly increase of nitrite concentration, it might evidence that a portion of the nitrite had been further deoxidized into biogas;
- OABAF system showed quite a high removal for organic matters at all steps – The TOC removal efficiency did not change significantly over the range of HRTs and C/N ratios studied. This demonstrates that the removal of TOC in the biological aerated filter is mainly a physical process which is not significantly affected by HRTs or C/N ratios. The slight TOC increase at 10.5 hours and C/N ratios at 1:1, and the parallel decrease of ammonia nitrogen and total nitrogen removal efficiencies, also reflects a decrease of microbial activity.

3.2 Final Phase analysis

The study at the Final Phase was accomplished through 2 steps in 15 days. At step I (from Day 1 to Day 11), the C/N ratio, carbon and nitrogen concentrations were kept constant, different HRTs were used to find out the optimum treatment performance. The flow rates applied range was from 0.04-0.08 l/min, corresponding to HRT of 6.3-15.6h, respectively. The C/N ratio at this stage was fixed to 12:1 and the total effective volume/total volume ratio was 55%. At step II (from Day 11 to Day 15), the HRT was kept constant for 10.5 hours with the C/N ratio only being changed from 1:1 to 12:1. <u>A mixture of plastic spores and coral stones</u> were used as biofilter media during the experimental works. The total effective volume/total volume ratio remained at 55%.

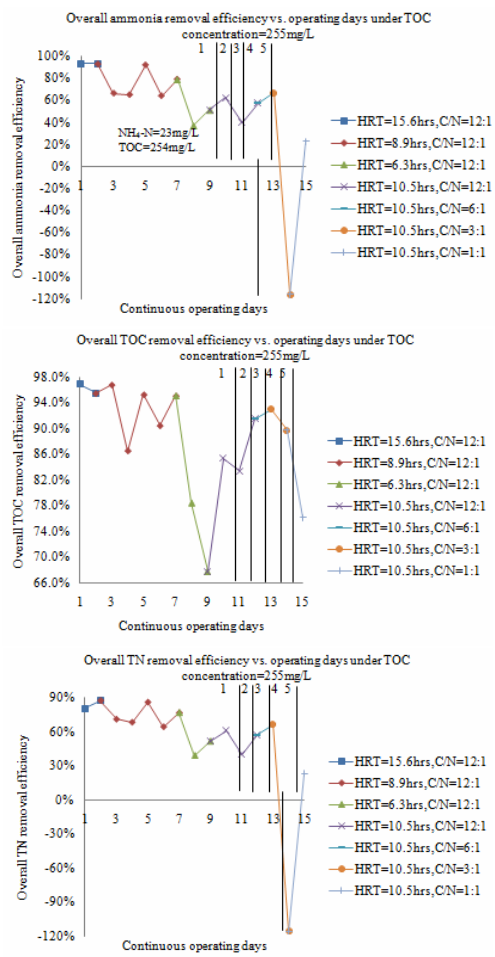


Figure 3-2 Profiles of overall ammonia, TN and TOC removal efficiency vs. operating days of the oxic and anoxic treatment of simulated leachate in Final Phase: 1. NH_4-N=23mg/L (C/N=12:1); 2. NH_4-N=23mg/L (C/N=12:1) and HRT=10.5hrs; 3. NH_4-N=44mg/L (C/N=6:1) and HRT=10.5hrs; 4. NH_4-N=85mg/L (C/N=3:1) and HRT=10.5hrs; 5. NH_4-N=303mg/L (C/N=1:1) and HRT=10.5hrs.

Parameters	Final
pH	7.75
Temperature (^{o}C)	20±2
C/N ratio	12:1
HRT (hrs)	10.4
Organic loading (g/hr)	18
Ammonia loading (g/hr)	1.5
Influent TOC concentration	256.3
Effluent TOC concentration	30.3
Influent NH_4-N concentration	22.8
Effluent NH_4-N concentration	9.6
Influent NO_3-N concentration	2.82
Effluent NO_3-N concentration	1.47
Influent NO_2-N concentration	0.005
Effluent NO_2-N concentration	0.03
Organic Removal Efficiency	94.8%
Ammonia Removal Efficiency	91.8%
Total Nitrogen Removal	82.9%

Table 3-3 Influent and effluent characteristics in Final Phase

The experiments results at Final Phase have shown that:

- During the 3 weeks operational period, it is found that under the total effective volume/total volume ratio is 55% and C/N ratio is 12:1 condition, when HRT is 10.5 hours, the OABAF system provides the optimum nitrogen removal performance. 91.8% of ammonia and 82.9% of total nitrogen is removed, meanwhile, the overall TOC removal efficiency can also achieved to 94.8%;
- A similar to Initial Phase, significant increases for all removal efficiencies are also observed as the C/N ratios and HRTs increase in most experiments - The average effluent concentration of ammonia nitrogen was 9.99 mg/L. Correspondingly, the removal efficiencies of ammonia nitrogen were 77.4%. The final effluent NO_x-N concentration was in the range of 0-3.5 mg/L and 1.35 mg/L on average. Note here that most of the effluent NO_x-N was nitrate (nitrite concentration was almost zero). The system presents quite a high removal for TOC at all HRTs and C/N ratios. It was found that the TOC in effluent was in the range of 7–50ppm (average of 27.4ppm). When increasing the C/N ratios, an apparent decrease of TOC and ammonia removal occurred;
- The lower ammonia nitrogen removal observed was most probably due to low C/N ratio inhibition. When the C/N ratios of the oxic reactors was changed from 12:1 to 1:1, the average effluent of ammonia nitrogen concentration sharply increased from 1 to 600 mg/L. This is because nitrifiers are autotrophic bacteria and heterotrophic bacteria have a higher specific growth rate than nitrifiers. When the leachate contains lots of nutrient, heterotrophic bacteria will reproduce very fast and occupy the surface of the biofilm. Therefore, a small population of nitrifiers could grow on the outer surface of biofilm, which is not favorable for ammonium removal. Moreover, the competition for oxygen between nitrifiers and heterotrophic bacteria became more intensive, because heterotrophic bacteria occupying the surface of biofilm could easily utilize the oxygen, and then cause decreased nitrification activity. The maximum nitrification rate of the reactor reached 1.3 mg/L at a HRT of 15.6 hours.
- Low oxygen tension may be another main critical factor controlling the distribution of ammonia and nitrite oxidizers in environmental systems.

4. CONCLUSION & RECOMMANDATION

The proposed a two-stage OABAF system for investigating nitrogen and organic removal in the treatment of leachate was evaluated. Based on the experiment results, the following conclusions were drawn:

- Without any physicochemical pretreatment at Initial Phase, 25–30 mg/L, 40-45 mg/L, 100-105 mg/L, and 325-335 mg/L of the raw NH_4^+-N was decreased to less than 2 mg/L, 5.5 mg/L, 33.6 mg/L and 83 mg/L, respectively. The system can achieve 68-93% NH_4^+-N removal efficiency with 95%–97% organic matter removal.
- The optimum operating condition for the OABAF system in the Final Phase was found to be of HRT = 10.5 hours and with organic loading of 18 g/hour, which corresponded to an overall ammonia removal efficiency of 65% and organic matter removal efficiency of 90%.
- The OABAF system proposed in this study was able to effectively nitrify ammonia nitrogen at high C/N ratios;
- The proposed system was very effective on removal of organic matters at all range of C/N ratios and HRTs;
- Ammonia removal via nitrification and denitrification is feasible in the biological treatment of landfill leachates which are readily occurring in decomposed solid waste environments. Results suggest that nitrification and denitrification may occur simultaneously in one OABAF system. On the other hand, studies evaluating the removal of nitrogen in the landfill are limited. So further investigation is needed to determine how other environmental conditions (i.e., temperature and gas-phase oxygen concentrations) affect ammonia removal.

References

[1] Borghei, S.M., Hosseini, S.H., 2004. The treatment of phenolic wastewater using a moving bed biofilm reactor. Process Biochemistry 39, 1177–1181.

[2] Chang W, Hong S, Park J. Effect of zeolite media for the treatment of textile wastewater in a biological aerated filter. Process Biochem 2002; 37:693–8.

[3] Zhao X, Wang Y, Ye Z. Oil field wastewater treatment in biological aerated filter by immobilized microorganisms. Process Biochem 2006; 41:1475–83.

[4] Gálvez, A., Zamorano, M., Hontoria, E., Ramos, A., 2006. Treatment of landfill leachate with aerated and non-aerated submerged biofilters. J. Environ. Sci. Health A 41, 1129–1144.

[5] Haq, I., 2003. Environmental impact assessment study: leaching of chemical contaminants from a municipal dump site Hastsal, Delhi (capital of India). Int. J. Environ. Stud. 60 (4), 363–377.

[6] Huang, L.-N., Chen, Y.-Q., Zhou, H., Luo, S., Lan, C.-Y., Qu, L.H., 2003. Characterization of methanogenic archaea in the leachate of a closed municipal solid waste landfill. FEMS Microbiol. Ecol. 46, 171–177.

[7] Kjeldsen, P., Barlaz, M.A., Rooker, R., Baun, A., Ledin, A., Christensen, T.H., 2002. Present and long-term composition of MSW landfill leachate: a review. Critical Reviews in Environmental Science and Technology 32 (4), 297–336.

[8] Berge N, Reinhart D, Townsend T. The fate of nitrogen in bioreactor landfills. Crit Rev Environ Sci Technol 2005; 35(4):365–99.

[9] Carrera J, Vicent T, Lafuente J. Effect of influent COD:N ratio on biological nitrogen removal (BNR) from high-strength ammonium industrial wastewater. Process Biochem 2004;39(12):2035–41.

[10] Di Iaconi C, Ramadori R, Lopez A. Combined biological and chemical degradation for treating a mature municipal landfill leachate. Biochem Eng J 2006; 31(2):118–24.

[11] Eding EH, Kamstra A, Verreth JAJ, Huisman EA, Klapwijk A. Design and operation of nitrifying trickling filters in recirculating aquaculture: a review. Aquacult Eng 2006; 34(3):234–60.

[12] Egli K, Bosshard F,Werlen C, Lais P, Siegrist H, Zehnder AJB, et al. Microbial composition and structure of a rotating biological contractor biofilm treating ammonium-rich wastewater without organic ca rbon. Microb Ecol 2003; 45(4):419–32.

Preliminary Study on the Model for Automatic Generation of Innovative Alternatives

Chih-ming Wu
Ph.D. Student,
Institute of Tech. Mgmt.,
Chung Hua Univ., Taiwan
d09603004@chu.edu.tw

Wen-der Yu
Professor,
Institute of Const. Mgmt.,
Chung Hua Univ., Taiwan
wenderyu@chu.edu.tw

and

Cheng Shao-tsai
Associate Professor,
Institute of Const. Mgmt.,
Chung Hua Univ., Taiwan
shaotsai@chu.edu.tw

ABSTRACT

Product/technology conceputalization is the most time-consuming and influential stage in a technology or product research and development project. The efficient generation of effective product/technology alternatives is the key to successful product/technology conceputalization. This paper presents the results of a preliminary study on the development of the Model for Automatic Generation of Innovative Alternatives (MAGIA) that offers a solution of automatic and efficient generation of effective product/technology alternatives. The proposed MAGIA integrates several existing methods including Function Modeling, TRIZ, and a specific genetic algorithm called Genetic Operation Tree (GOT), which can manupulates on the function models from an initial state to generate a final solution of a product/technology concept. In order to test the feasibility of the proposed MAGIA, an existing manhoe construction technology is selected for case study. The preliminary results show promising potentials of the proposed MAGIA for innovation of construction technologies.

1. INTRODUCTION

Generation of innovative alternatives is a critical step in technology innovation. Some approaches employed patent databases as sources of conceptual ideas [1]; others adopted the TRIZ method to systematically problem solving [2]. No matter which approaches adopted, there has no yet automatic method developed for the generation of innovative alternatives in terms of innovation objectives such as improvement of performance/function and reduction of implementation costs. The essential barrier to the automatic generation of innovative alternatives resides in the incapability of technology modeling within the requirement of an optimization model. Function Model (FM) has been the most widely adopted technology modeling scheme in many popular computer aided inovation (CAI) tools such as Goldfire Innovator® and CREAX®. The traditional optimization techniques, such as mathmetic (linear/nonlinear/inetger) programming, requires structural formation of the domain problem (e.g., function model). Non-structural optimization scheme such as Genetic Algorithms (GA) [3] may provide a solution to such barrier. In searching appropriate GA representation schemes, a recent work by Yeh and Lien [4] on the development of Operation Tree using GA has drawn our special attention. The combination of Operation Tree and GA gave birth to a new GA called Genetic Operation Tree (GOT) which provides a tree-like representation scheme for the problem domain. However, before adopting GOT to automatic FM generation, there are several problems to be resolve.

First, GOT is a tree-like representation diagram, which is different from the flow-charting diagrams such as FM; second, the objective function for product/technology needs to be defined, a guided search algorithm should be developed; third, the manupulation of GOT to generate practical innovative alternatives (in terms of FMs) instead of purely ideal concepts is challeging; forth, the implementation of generated FM in real world applications deserves equal attentions.

The resent of the paer is dedicated to resolve the abovementioned problems. A Model for Automatic Generation of Innovative Alternatives (MAGIA) that combines FM, TRIZ, and GOT is proposed to manupulate on the function models of the problem domain from an initial state to generate a final solution of a product/technology concept. In order to test the feasibility of the proposed MAGIA, an existing manhoe construction technology is selected for case study.

2. RELATED FIELDS

In this section, three fields related to the development of MAGIA are reviewed to provide required backgrounds including function modelling, TRIZ, and GOT.

1. Function Model (FM)

A function model or functional model in systems engineering and software engineering is a structured representation of the functions, activities or processes within the modelled system or subject area [5]. A function model is a graphical representation of a product/technology function within a defined scope. The purposes of a function model are to describe the functions and processes, assist with discovery of information needs, help identify opportunities, and establish a basis for determining product and service costs [6]. In systems engineering and software engineering areas a function model is created with a functional modelling perspective [7]. A functional modelling perspective of FM concentrates on describing the dynamic process. The dynamic process can be a function, transformation, activity, action, task etc. A well-known example is data flow diagrams. The perspective uses four symbols to describe a process including [8]: (1) Process — Illustrates transformation from input to output; (2) Store—Data-collection or some sort of material; (3) Flow— Movement of data or material in the process; and (4) External Entity: External to the modeled system, but interacts with it.

Another popular perspective of FM commonly adopted for patent/technology modeling is the Subject-Action-Object (SAO) model, where Function Models define the relationships between system elements in terms of the functions they perform [9]. In the SAO model, a *function* is "a *action* that directly changes or maintains a controllable or measurable **parameter** of a (material) *object*." The SAO model is depicted in Figure 1. Examples of *actions* are Move, remove, burn, weld, count, deposit, inform, rotate, hold, conduct, carry…etc. The SAO type FMs represent a system (describing the functions of a product/technology) two natural language templates: (1) Action-Object (AO), e.g., move (A) table (O); (2) Subject-Action-Object (SAO), e.g., conveyer (S) moves (A) table (O). A **parameter** is a directly measurable or controllable characteristic associated with a material *object* which is affected by a function. Examples of parameters affected by *actions* are length, area, volume, mass, density, volts, bits, joules, coulombs, temperature, roentgens..etc. A commonly way to identify the SAO FM from a product/technology is by asking the following questions: (1) What **parameter** of the *object* is controlled of changed by the action? (2) How do I measure action object?

Figure 1 SAO function model

In the proposed MAGIA, the SAO FMs are adopted as the representation scheme for the product/technology to be innovated.

2. Theory of Inventive Problem Solving (TRIZ)

The term TRIZ is the acronym for "Theory of Inventive Problem Solving" in Russian. TRIZ was developed by a Russian patent engineer Genrich Altshuller and his colleagues in the former USSR starting in 1946 [10]. Altshuller studied more than 2000,000 patents from various engineering areas and developed several important tools used for inventive problem solving, such as trend of evolution, contradiction matrix, physical contradiction resolution principles (also called inventive principles), substance and field analysis, and ideal final result. Among which, the contradiction matrix combined with 39 engineering parameters (EPs) and 40 inventive principles (IPs) are most popular techniques for application of TRIZ [2]. A comparison between the traditional vs. TRIZ problem solving processes is depicted in Figure 2. In the traditional problem-solving, the initial problem described by engineers was solved by brain-storming, where there was no methodology for systematic mapping of problem to the solutions. On the contrast, TRIZ method provides a systematic approach to model the initial domain problem by 39 EPs and represents the problem in conflicting EPs; then, the contradicting matrix is consulted and the associated IPs are recommended. The engineers then develop their solutions for the domain problem based on the IPs recommended by TRIZ.

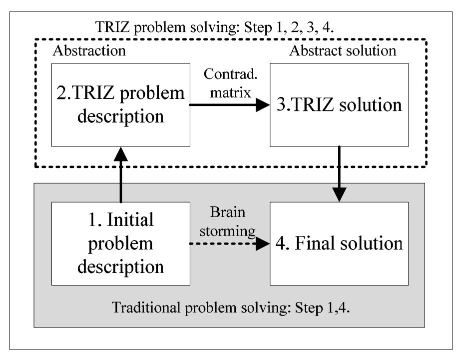


Figure 2 Traditional vs. TRIZ problem solving

In the proposed MAGIA, TRIZ method is consulted to providing the recommeded IPs for evoluation of innovative alternatives of the target product/technology.

3. Genetic Operation Tree (GOT)

Genetic Algorithm (GA) was first proposed by Holland [3]. The GA is a stochastic optimization procedure that efficiently searches a very large solution space by manipulating a set of strings (genotypes), each of which encodes a single point in the solution space. The population of strings is improved in a manner similar to Darwin's theory of natural selection. The fittest members, measured by a fitness function, attract more mates and therefore produce a greater number of offspring who inherit the superior characteristics of the parents. Offsprings are generated by randomly combining genetic characteristics from pairs of parent strings with the crossover operation. Random mutation of offsprings occasionally provides new opportunities for improvement. Genetic characteristics from the fittest members of a population are passed from generation to generation. Finally, a set of nearly optimal solutions is produced.

For a typical GA, there are four basic terms which need to be defined: (1) chromosome─a string of numbers which in some way represent a specific component of the considered object or process to be optimized; (2) genotype a collection of all chromosomes for the considered object or process; (3) population─a collection of genotypes; and (4) fitness function─overall suitability of a genotype according to the objective of the problem to be optimized. There are four basic operations of a GA include: (1) fitness evaluation─determine the fitness of the genotypes; (2) parent selection─randomly select two genotypes as parents; (3) crossover─exchange the strings in the chromosomes of the parent geneotypes; and (4) mutation─mutate some strings in the chromosomes of the parent geneotypes [11].

Genetic Operation Tree (GOT) is an extention of GAs, which applies GAs on the operation trees (OTs) as shown in Figure 3. The OTs are tree-like diagrams structured hierarchically to represent or model the domain problem.

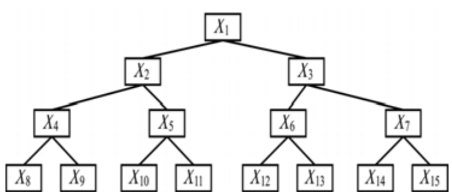


Figure 3 Operation Tree

Yeh and Lien proposed a GOT to produce self-organized formulae for predicting the strength of High-Performance Concrete [4]. In their model, the OT was encoded as a geneotype of GA. A special computation algorithm was developed to perform GA operations on the OTs they constructed. Yeh and Lien's work was mainly focused on determining the strength prediction mdoel (in terms of mathematic formulae) for High-Performance Concrete. The elements of OTs were associated with the variables, constants, and mathematic operations. The fitness measure of their problem was to minimize the difference between the estimated and known prediction values.

3. PROPOSED MODEL FOR AUTOMATIC GENERATION OF INNOVATIVE ALTERNATIVES (MAGIA)

As described previously, the objective of the current research is to generate effective product/technology FMs automatically and efficiently. In order to achieve this objective, a prototype Model for Automatic Generation of Innovative Alternatives (MAGIA) is proposed, which combines FM, TRIZ, and GOT techniques to form an integrated innovative alternatives generation method. This section describes the proposed MAGIA model in details.

1. Basic Model

The SAO FM depicted in Figure 1 is adopted as the basic model for function modelling of the target product/technology for innovation. The SAO FM will be translated into a three-level OT similar to the one as shown in Figure 3, where the leaves of OT represent the elements (*subject* or *object*) of the SAO model and the nodes in the middle level represent the *action* that the *subject* performs on the *object*. The *parameters* are recorded as the values of the leaves associated with the *subject* and *object*. The schematic diagram of the MAGIA basic model is shown in Figure 4. All target products/technologies for innovating will be represented in forms of basic model.

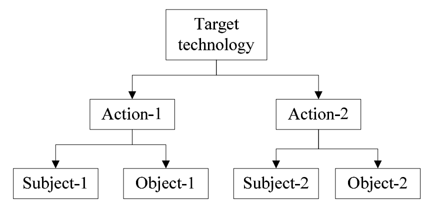


Figure 4 Basic model of MAGIA

2. GOT Encoding

The encoding of MAGIA OT into GOT is straightforward. The *Subjects*, *Objects* and *Actions* are substituted with predefined codes. The values of the elements are specific **parameters** that affect the effects of the *actions*. Generic real number strings are employed in GOT encoding of the MAGIA models. Figure 5 shows an example of the encoded genotype, where strings of Chromosome *i* are associated parameter values with the elements of the i^{th} SAO FM in of the target product/technology.

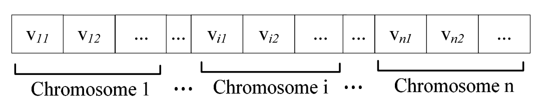


Figure 5 GOT encoding

3. Fitness Function

A specific *fitness function* based on value engineering is defined as the *Value* of the innovative alternative as in the following equation.

$$V = \frac{F}{P + C + T} \qquad (1)$$

In equation (1), F represents the value-adding function of the alternative; P is the potential problem (disliked functions) induced by the alternative; C is the estimated increased cost due to the adoption of the alternative; and T is the extra needed time to perform the work.

In real world application, it may be not easy to measure the values of *F, P, C, T* in Equation (1). In the proposed MAGIA, the contradition matrix of TRIZ is consulted and the domain experts are interviewed to provide required information of the abovementioned parameters in Equation (1).

4. Operational Procedure

A five-step operational procedure is proposed to conduct MAGIA as shown in Figure 6. The proposed procedure is described in the following:
Problem definition─Defining the domain problem; usually the root cause analaysis (RCA) method is employed to analyze and decribe the domain problem.

Funcion modelling─constructing the SAO FM of the target product/technology for innovation. At the same time, a portion of the SAO FM is selected as the scope for GOT evolution.

TRIZ consulting─consulting TRIZ contradition matrix to identify the primitive potential improvement direction: (a) as both the improvement and deteriorated EPs are known, the traditional contradiction matrix is consulted [12] and traditional TRIZ approach is adopted; (b) if only the improvement EP is known, the unitary EP approach [12] is adopted.

GOT evolution─encoding SAO FM into a GOT genotype for evolution; the GA is applied to GOT to generate improvement alternatives.

Innovative alternative evaluation─evaluating the fitness value of the generated innovative alternatives; if the fitness level is accepted or the preset No. of generations is reached, the evoluation process stops.

Alternative implementation─transforming the innovative alternative into practical application solutions.

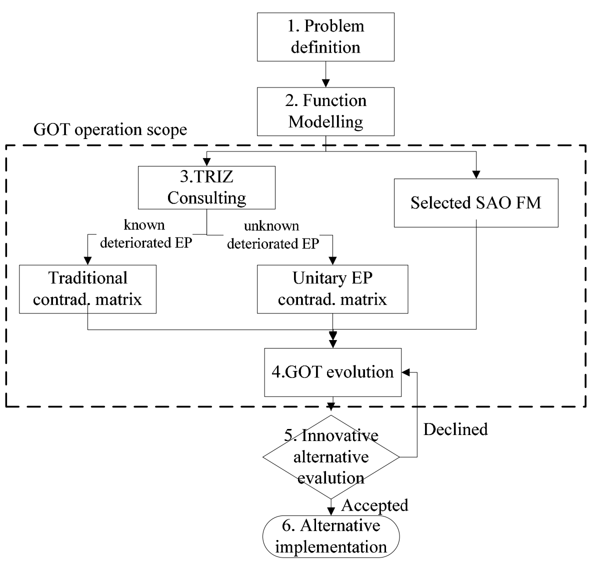


Figure 6 MAGIA operational procedure

4. DEMONSTRATED CASE STUDY

In this section, an existing manhole construction technology is selected as target for case study in order to test the feasibility of the proposed MAGIA. As Taiwan has become a developed country, many infrastructure systems are aging including highways, utilities (e.g., electricity, water supply, and gas), sewage, communication (e.g., TV, telephone) conduits, etc. As common conduits are rarely adopted in Taiwan, most of the pipeline conduits are constructed separately under roadways in the urban and suburban areas. Manholes are constructed in order to maintain the utility pipelines. It is found that the average lifetime of traditional manhole is less than two years. Such damaged manholes have become one of the major causes of deterioration of road pavement. Moreover, the deteriorated pavement surrounding manholes has contributed significant part of car and motorcycle accidents annually. According to the statistics of Public Construction Commission, there are about 2.95 millions of various manholes in Taiwan [14]. As a result, improvement of the lifecycle of manhole can not only save great amount of budget for public construction but also defer the deterioration of road pavement and improve the safety of road users. It is very desirable for public road agencies to develop an improved manhole construction technology so that the lifetime of manhole is prolonged, the quality of road pavement surface is improved, and the lifecycle cost of manhole is reduced. In a previous research, an innovative precast method has been developed to replace the traditional manhole technology [13]. The innovative alternative has been proved to be able to extend the lifetime of manholes from 1.5 years to more than five years [15].

(1) Problem Definition

According to a previous study [13], the major cause for manhole deterioration is the re-levelling of manhole covers due to re-pavement of road. As the limited time allowed for concrete curing, the structure strength of manhole is not fully developed before undertaking the loads of traffics. The structure is cracked inside the manhole and the surrounding soils of the base layer could flew into the manhole with water when it rains. As a result, the base layer of road is damaged and the road pavement is deteriorated. The root causes of manhole damaged are referred to our previous work [13]. From the previous study [13], the root cause for road pavement deterioration surrounding manhole is the drainage with soils flowing into manhole. Two causes trigger the drainage with soils are: (1) cracked manhole; and (2) rains (water). Furthermore, the insufficient strength is the main cause of cracked concrete. Since the Rains should be considered as super system (unchangeable), the solution to the problem can be solved by providing sufficient structure strength and blocking the channel of drainage into the manhole.

(2) Function Modelling

In this step, the existing manhole technology is obtained via patent search of patent databases. After state-of-the-art survey the target technology "US 3773428: Adjustable manhole support" is considered as the target technology for innovation. The target technology is considered "traditional manhole construction method (traditional method)" hereafter in this paper. There are 8 steps in the target technology: (1) preparation; (2) escavation; (3) surveying; (4) formwork; (5) concrete placement; (6) cover frame installation; (7) form removal; and (8) refill. The SAO FM for the traditional manhole construction method is shown in Figure 7.

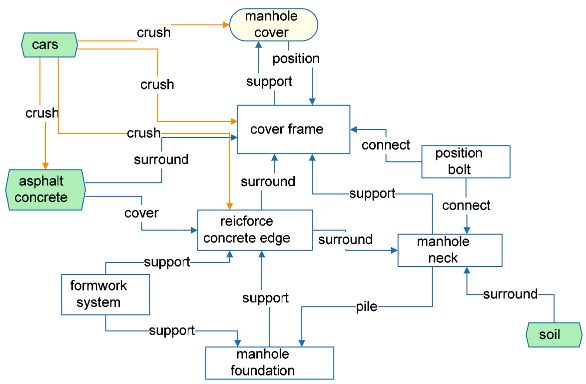


Figure 7 Complete FM fortraditional manhole technology

The FM of Figure 7 also provides information on harmful and useful *actions* between *subject* and *object*. For example, the "Reinforced concrete edge" (*subject*) performs two "support/break" (*actions*) to "Cover frame" and "Manhole neck" (*objects*). All of the *actions* provide both useful (supporting) and harmful (breaking) functions to the *objects*. This portion of SAO FM is isolated for special attention and shown in Figure 8. The fitness function is defined as shown in Equation (1). By expert judgments, the values of *F, P, C, T* in Equation (1) are estimated to be 7, 10, 6, 10 (using 1~10 scale) respectively. As a result, the fitness value of the traditional method is 0.36. Since the fitness value of the selected FM is relatively low, it needs improvement.

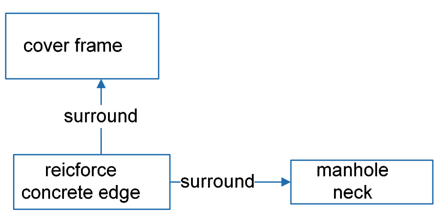


Figure 8 Selected SAO FM

(3) TRIZ consulting

The defined problem in (1) implies that the early stage strength of conrete is the key to solve the problem. As a result, the EP to be improved is "(curing) time saving". However, there is no information about the deteriorated EP. Therefore, the unitary EP approach [9] is adopted. By consulting TRIZ contradiction matrix, the top four recommended IPs for "time saving" are IP-10 (15 times, 13%), IP-35 (14 times, 12%), IP-18 (12 times, 10%), and IP-28 (10 times, 8%). Such information can be used to guide the evolution MAGIA.

(4) GOT evolution

The selected FM of Figure 8 is encoded as a specific GOT for evolution. Relevant parameter settings are: (a) population size—100; (b) cross over rate—0.9; (c) mutation rate—0.001 (1%); (d) maximum generation—75.

It is found that the evolution stopped after the 6th geneation. The initial, intermediate and final states of GOT evolution are shown in Figure 9.

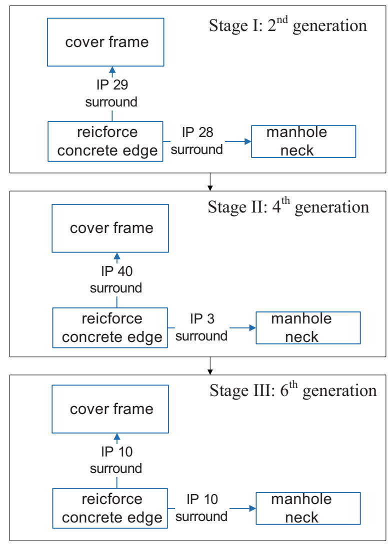


Figure 9 GOT evolution of FMs

From Figure 9, there are three stages in the evolution: (1) Stage I—the 2nd generation; (1) Stage II—the 4th generation; (1) Stage III—the 6th generation. After the 6th generation, the evolution converged to the final state shown in Figure 9. The FM of the generated innovative alternative is shown in Figure 10.

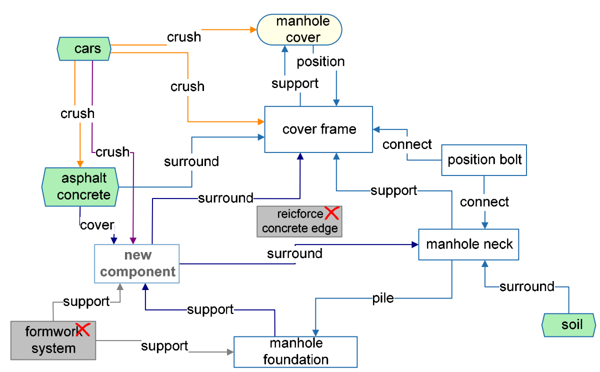


Figure 10 FM of the generated alternative

(5) Alternative Evaluation

The alternative evaluation is performed sub-quantitatively according to the fitness function of Equation (1). The evaluation results are shown in Figure 11. It shows that the generated innovative alternative improves in all criteria.

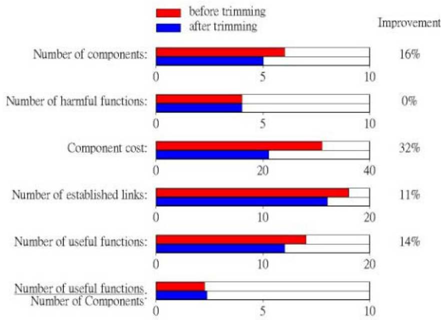


Figure 11 Evaluation of the generated alternative

(6) Alternative implementation

In this step, implementation method for the conceptual innovative technology is designed. A CAI tool, Goldfire Innovator®, is consulted with the express "How to increase/improve concrete initial strength?" The results are returned with 12 patents. Among those, 10 patents suggest using additives while the other two recommend precast or pre-stressed method. both of the recommended can be considered in real world application.

5. CONCLUSIONS

In the paper a Model for Automatic Generation of Innovative Alternatives (MAGIA) is proposed for automatic generation of product/technology concepts that was considered by engineers as the most time-consuming and costly work in a product/technology R&D project. The proposed MAGIA integrates several existing methods including Function Modeling, TRIZ, and Genetic Operation Tree, which utilizes the capability of self-evoluting capabilities of GA to generate innovative alternatives that improve the values of a product/technology. An existing manhole construction technology is selected for case study. The preliminary results show promising potentials of the proposed MAGIA for innovation of construction and other technologies.

Although the prototype of MAGIA has been developed, a more integrated and well verified system is desired to implement the proposed method and assist engineers for real world applications. This needs further researches.

6. ACKNOWLEDGEMENT

The work presented in the paper is finanacially supported by the National Science Council, Taiwan, under project No. NSC 98-2221-E-216-044. Sincere appreciations are given to the sponsor by the authors. Appreciations are also given to Prof. Yeh and Mr. Lien, the authurs of reference [4], who provide kindly help in GOT development and operations of this research.

REFERENCES

[1] Duan, H., Lin A., Li Y. and Mang, H., "CAI tool case study of reducing vibration of diesel engine lubricating oil system," *Journal of TRIZ in Engineering Design*, Vol.2, No.2, p105-114, 2006.

[2] Mohamed, Y. and AbouRizk, S. (2005). "Application of the Theory of Inventive Problem Solving in Tunnel Construction," *J. Constr. Engrg. and Mgmt.*, ASCE, Vol. 131, No. 10, 1099~1108.

[3] Holland, J. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbour, Michigan. 1975.

[4] Yeh, I-Cheng, Lien, L. C., "Knowledge discovery of concrete material using Genetic Operation Trees," *Expert Systems with Applications*, Vol. 36, pp.5807–5812, 2009.

[5] NIST, *FIPS Publication 183,* released of IDEFØ by the Computer Systems Laboratory of the National Institute of Standards and Technology, USA, 1993.

[6] Gantt, H.L., *Work, Wages and Profits*, Hive Publishing Company, Easton, PA, USA, 1974,.

[7] DOD, Systems Engineering Fundamentals. Defence Acquisition University Press, Department of Defence, USA, 2001.

[8] Wikipedia, http://en.wikipedia.org/wiki/Function_model#cite_note-9, visited 2010/05.

[9] Liu, H. C., "Design and Study of Automated Text Summarization for Extracting SAO Structures from Chinese Patent Documents," *Master Thesis*, Department of Information and Computer Science, National Chiao Tung University, Hsinchu, Taiwan, 2009.

[10] Altshuller, G. *40 principles: TRIZ keys to technical innovation*. L. Shulyak and S. Rodman, trans., Technical Innovation Centre, Worcester, MA, USA, 1998.

[11] Goldberg, D. E., *Genetic algorithms in search, optimization and machine learning*, Addison-Wesley Publishing Company, Reading, MA, USA, 1989.

[12] Liu, C. C., " A Study of TRIZ Method Improvements and Eco-Innovative Design Methods," Ph.D. Thesis, Department of Mechanical Engineering, National Cheng Kung University, Tainan, Taiwan, 2003.

[13] Yu, W. D., Wu, C. M., and Chen, W. C., "Innovating Road Manhole Technology with Systematic Technology Innovation Procedure (STIP)," Proceedings of The 2nd International Multi-Conference on Engineering and Technological Innovation (IMETI 2009), July 10~13, 2009, Orlando, Florida, USA, pp. 35~40, 2009.

[14] Public Construction Commission, "Monthly Report of Smooth Way Program," *Internal Meeting Material of PCC*, Smooth Way Program Task Force, Public Construction Commission, Executive Yuan, Taiwan, March 2010. (in Chinese)

[15] Yu, W. D., Cheng, S. T., Wu, C. M., and Chen, W. C., "Application of Systematic Technology Innovation Process to Innovate Manhole Technology," *Chinese Construction Management Journal*, No. 80, pp. 55-65. (in Chinese)

Digital Filtering for Power Systems Distance Protection Using Walsh Functions

Roberto MICHELETTI

Department of Electrical Systems and Automation, University of Pisa
Largo L. Lazzarino, 1 I-56122 Pisa, Italy

ABSTRACT

Distance relays must estimate accurately and quickly the distance to the fault even in presence of highly distorted input signals. Several digital filtering algorithms for distance relays have been proposed and some of them are used in practical applications. However nonlinear loads are becoming more and more common in electric power systems and that increases the corruption level of signals. For this reason, the study of digital filtering techniques for the extraction of fundamental components from highly corrupt voltage and current waveforms is of vital importance for the development of algorithms applied to the digital protection of transmission lines. The studied technique proposed in this paper is based on Walsh functions.

Keywords: Digital Filtering, Walsh functions, Fault location, Power systems.

1. INTRODUCTION

Several papers have been published in recent years in digital protection of power systems and in particular in protection of transmission lines because they are primarily exposed to short circuits and consequently present a higher probability of fault occurrence. Transmission lines are normally protected with distance relays; they estimate the impedance between the short circuit and the relay location using voltage and current signals. However during the fault occurrence the input signals to protective relays are contaminated with noise (harmonics and the exponentially decaying DC component), that must be rejected while must be retained only signal quantities of interest [1]-[9].

The presence of harmonics and noise determines errors when estimating the fundamental components of voltage and current necessary for the apparent impedance calculation. For this reason it is very important to implement an efficient filtering technique in order to extract the fundamental components of voltage and current and calculate precisely the apparent impedance.

The present paper presents a fault location estimator, based on Walsh functions [10]-[13], able to extract the fundamental components from faulted voltage and current waveforms taken from a typical 345 kV transmission line. In this paper it has been included a corrective factor that improves the behavior of the measuring system by compensating the error source due to the exponentially decaying DC component of the current consequent to the fault inception.

2. DIGITAL FILTERING ALGORITHM

The Fourier series of a periodic waveform f(t) may be written

$$f(t)=a_0+\sum_{n=1}^{\infty}\left[a_n\cos n\omega t+b_n\sin n\omega t\right] \qquad (1)$$

where a_0 is the mean value and:

$$a_n=\frac{2}{T}\int_0^T f(t)\cos n\omega t\,dt \qquad (2)$$

$$b_n=\frac{2}{T}\int_0^T f(t)\sin n\omega t\,dt \qquad (3)$$

and $T=2\pi/\omega$ is the period of the fundamental frequency.

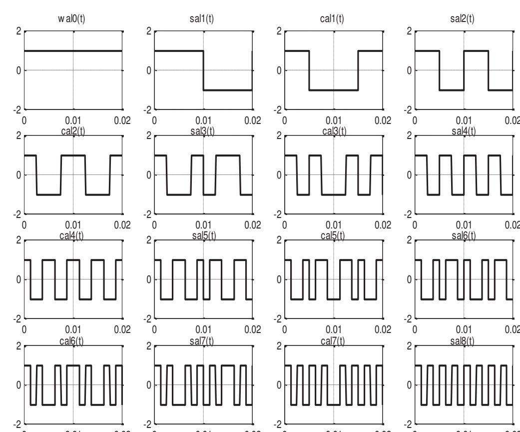


Fig. 1 Walsh functions up to the eighth order

Walsh functions, which are shown in Fig. 1 up to $sal_8(t)$, can be used as an alternative orthogonal family for the series representation of a periodic waveform according to

$$f(t) = A_0 + \sum_{n=1}^{\infty} \left[A_n cal_n(t) + B_n sal_n(t) \right] \qquad (4)$$

where

$$A_n = \frac{1}{T}\int_0^T f(t)\ cal_n(t)\, dt \qquad (5)$$

$$B_n = \frac{1}{T}\int_0^T f(t)\ sal_n(t)\, dt \qquad (6)$$

$$A_0 = \frac{1}{T}\int_0^T f(t)\ wal_0(t)\, dt \qquad (7)$$

The procedure for the Walsh series determination and the corresponding Fourier derivation is given below.
Consider a periodic voltage f(t) whose period is T at the fundamental frequency f =50 Hz. Signal f(t) is sampled with sampling frequency f_s.
The Walsh coefficients A_n and B_n are obtained according to (5), (6), (7). To yield the Fourier spectrum of f(t), A_n and B_n are first multiplied by Walsh-Fourier conversion matrices $\mathbf{F_a}$ and $\mathbf{F_b}$, respectively, to give approximations **a*, b*** to the cosine and sine coefficients of f(t):

$$\begin{aligned} \mathbf{a^* = F_a\ A} \\ \mathbf{b^* = F_b\ B} \end{aligned} \qquad (8)$$

where

$$\mathbf{A} = \begin{vmatrix} A_1 \\ A_2 \\ . \\ . \\ A_S \end{vmatrix} \qquad \mathbf{B} = \begin{vmatrix} B_1 \\ B_2 \\ . \\ . \\ B_S \end{vmatrix} \qquad (9)$$

represent the cal spectrum and sal spectrum of f(t), respectively.
The element of the Walsh to Fourier conversion matrices $\mathbf{F_a}$ and $\mathbf{F_b}$ are the Fourier coefficients of the Walsh functions $cal_m(t)$ and $sal_m(t)$, respectively.
Thus

$$a_{n,m} = \frac{2}{T}\int_0^T cal_m(t)\ \cos n\omega t\ dt \qquad (10)$$

$$b_{n,m} = \frac{2}{T}\int_0^T sal_m(t)\ \sin n\omega t\ dt \qquad (11)$$

where n is the index of row (corresponding to the harmonic order) and m is the index of column. Moreover the elements of the conversion matrix $\mathbf{F_a}$ assume the same absolute value of the conversion matrix $\mathbf{F_b}$, but their signs may differ.
The conversion matrices are essentially semi-infinite; when applied in (8) however they are truncated to SxS element square matrices. Also an additional matrix multiplication is required which compensated for the Walsh spectral truncation

$$\begin{aligned} \mathbf{K_a\ a^* = a} \\ \mathbf{K_b\ b^* = b} \end{aligned} \qquad (12)$$

where **a** and **b** are the desired cosine and sine coefficients, respectively. The simplest form of compensation occurs when $S = 2^k - 1$, i.e., the highest harmonic order in f(t) doesn't exceed the third, or the seventh, or the fifteenth,….
Then compensation takes the form of a diagonal matrix, with fixed elements

$$\left[\frac{\pi n / 2^{k+1}}{\sin(\pi n / 2^{k+1})} \right]^2 \qquad (13)$$

where n is the harmonic order.
Thus the fundamental waveform $f_1(t)$ of the signal may be written

$$f_1(t) = M_1 \sin(2\pi f t + \varphi_1) \qquad (14)$$

where

$$\begin{aligned} M_1 &= \sqrt{a_1^{\,2} + b_1^{\,2}} \\ \varphi_1 &= \tan^{-1}(a_1/b_1) \end{aligned} \qquad (15)$$

3. PROPOSED TECHNIQUE EVALUATION

In order to test the applicability of the proposed technique, a simulation of the transmission line in a faulted condition was utilized. This paper makes use of a digital simulation of faulted EHV transmission lines, whose model is described in [6] and [7]. A 250 km transmission line model was used for testing the fault distance algorithm. The series parameters of the power system are taken from a typical 345 kV transmission line; the parallel parameters of the line are neglected in this analysis.

As mentioned before, during the fault occurrence, the voltage and current waveforms possess the fundamental components with the addition of harmonics and, in particular the current waveform, a DC decaying offset.
We consider the line in faulted condition, neglecting the effects of the output impedances of power transformers; therefore the system model is linear. Moreover practical considerations such as the effects of quantization, etc. in primary system fault data are also included in the analysis, so that the data obtained are very close to those found in practice.
Fig. 2 shows typical voltage and short-circuit current waveforms.

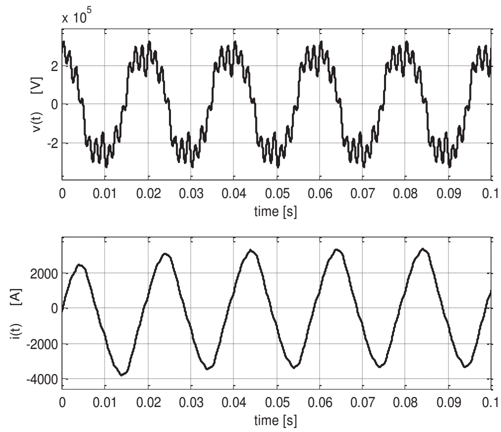


Fig. 2 Typical voltage and short-circuit current waveforms

Fig. 3 shows the components of the short circuit current, the steady state current and the exponentially decaying DC component.

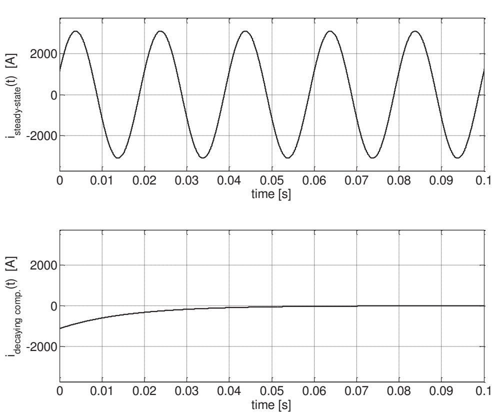


Fig. 3 Waveforms of the steady state current and the exponentially decaying DC component

The proposed procedure permits to obtain, besides the DC offset of the current I_0, the fundamental components of current $i_1(t)$ and voltage $v_1(t)$, whose waveforms, compared with the reference ones $i_{1ref}(t)$ and $v_{1ref}(t)$, obtained by the same system model in the known steady state sinusoidal condition, are shown in Fig. 4 and Fig. 5.

$$I_0 = A_0 \tag{16}$$

$$i_1(t) = I_1 \sin(2\pi f t + \varphi_{i1}) \tag{17}$$

$$v_1(t) = V_1 \sin(2\pi f t + \varphi_{v1}) \tag{18}$$

with

$$\begin{aligned} I_1 &= \sqrt{a_{i1}{}^2 + b_{i1}{}^2} \\ \varphi_{i1} &= \tan^{-1}(a_{i1}/b_{i1}) \end{aligned} \tag{19}$$

$$\begin{aligned} V_1 &= \sqrt{a_{v1}{}^2 + b_{v1}{}^2} \\ \varphi_{v1} &= \tan^{-1}(a_{v1}/b_{v1}) \end{aligned} \tag{20}$$

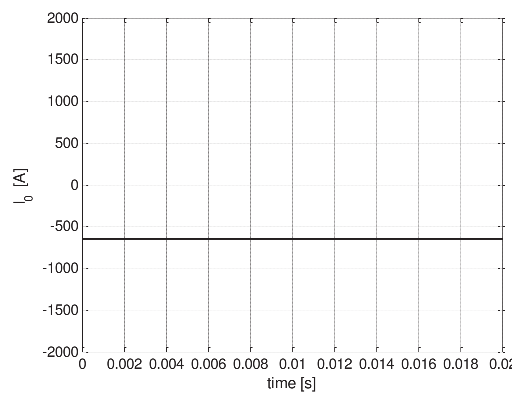


Fig. 4 DC offset of the current I_0.

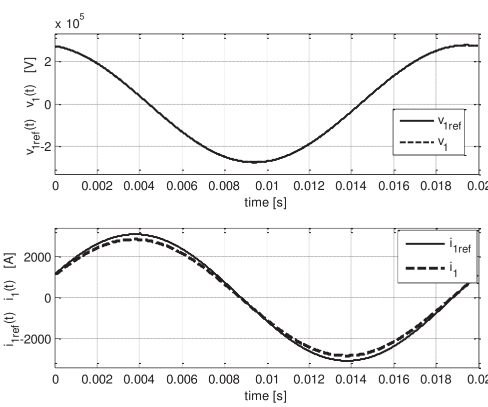


Fig. 5 Fundamental components of the voltage and current waveforms and reference waveforms.

As we can see, the waveforms of $v_1(t)$ and $v_{1ref}(t)$ are practically overlapped, while the waveforms of $i_1(t)$ and $i_{1ref}(t)$ are slightly different, due to the presence of the exponentially decaying DC component.

To obtain a more adequate match between $i_1(t)$ and $i_{1ref}(t)$, two corrective factors, k_A and k_{ph}, are introduced in order to modify both amplitude and phase angle, respectively, of the fundamental component of the current.

Corrective factors values, for given parameters of the transmission line, are function of ratios I_0/I_1 and τ/T, where τ is the line time constant; however if τ/T is in the range from 0.65 to 0.85, k_A and k_{ph} are practically functions only of ratio I_0/I_1.

Fig. 6 shows corrective factors k_A and k_{ph} as a function of ratio I_0/I_1 and for τ/T in the above mentioned range.

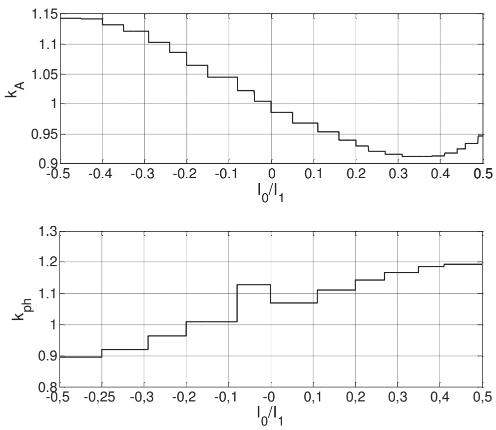


Fig. 6 Corrective factors k_A and k_{ph} as a function of ratio I_0/I_1 and for τ/T in the range from 0.65 to 0.85.

After rearrangement, the corrected fundamental component of the current $i_{1c}(t)$ may be written

$$i_{1c}(t) = I_{1c} \sin(2\pi f t + \varphi_{i1c}) \tag{21}$$

with

$$\begin{aligned} I_{1c} &= k_A I_1 \\ \varphi_{i1c} &= k_{ph}\varphi_{i1} \end{aligned} \tag{22}$$

Fig. 7 shows the waveform of the fundamental component of the current $i_{1c}(t)$ after correction, compared with the reference one $i_{1ref}(t)$.

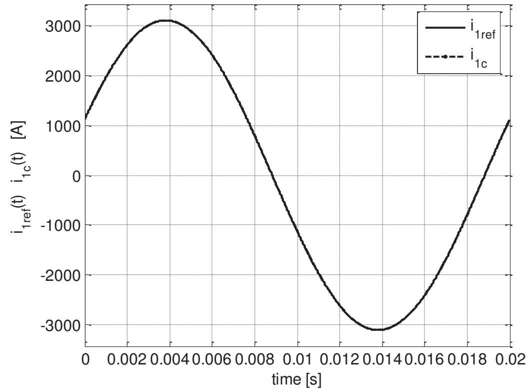


Fig. 7 Fundamental component of the current waveform after correction and reference waveform

The line reactance is given by

$$X_1 = \frac{V_1}{I_{1c}} \sin(\varphi_{v1} - \varphi_{i1c}) \tag{23}$$

Neglecting the fault reactance with respect to the line reactance, the fault distance is determined by

$$d = \frac{X_1}{x} \tag{24}$$

where x is the specific line reactance expressed in Ω/km.

Fig. 8 shows the estimated fault distance (true distance 100 km) as a function of the voltage phase angle φ_{v1} at the instant of the fault inception.

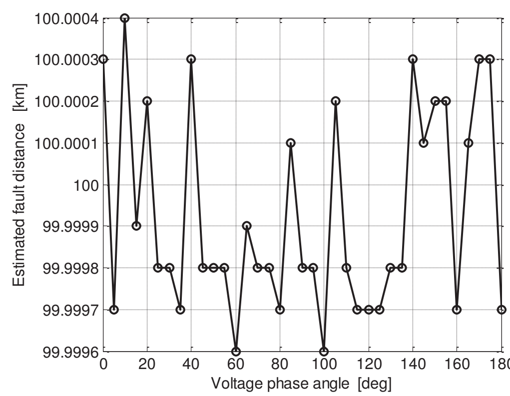


Fig. 8 Estimated fault distance as a function of the voltage phase angle φ_{v1} (true distance 100 km)

Fig. 9 shows the error in the estimated fault distance as a function of the voltage phase angle φ_{v1}.

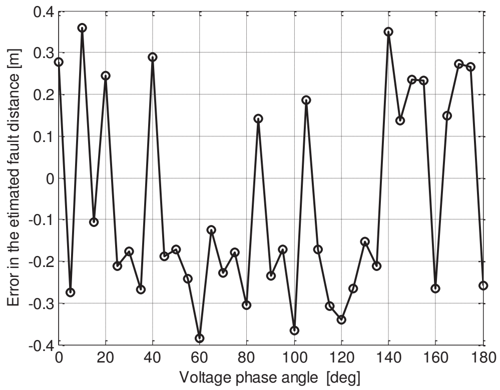


Fig. 9 Error in the estimated fault distance as a function of the voltage phase angle φ_{v1}

As observed from Fig. 8 and Fig. 9 the accuracy of the estimated fault distance is within ± 0.0004 % for any value of the voltage phase angle φ_{v1}. The uncertainty interval $\pm \Delta d$ is ± 0.4 m and doesn't depend, in our analysis in which we neglected both the influence of the line parallel parameters both the current transformers, on the fault distance itself.

4. CONCLUSION

The paper has described an algorithm for accurately locating three phase faults on high voltage transmission lines.

The digital filtering technique, based on Walsh functions, extracts the phasors of the fundamental components of voltage and current and calculates precisely the faulted line reactance and then the location of the fault.

This measurement method is fast and accurate and doesn't require expensive instrumentation; moreover it is particularly well-suited to power-frequency waveform measurements, because the synchronization of the Walsh waves with the input voltage signal, suitably conditioned, is not difficult.

The sources of errors due to voltage and current transformers have not been included in the analysis; as well, it has not been taken into account the effect of the output impedances parameters of power transformers and the parallel parameters of the line in the current calculation.

The method can be used for any type of fault on a three phase transmission line by appropriately altering the sequence networks' interconnection.

5. REFERENCES

[1] D.J., Kish, G.H., Heydt: "An introduction to power systems analysis using Walsh functions", **Proc. North American Power Symposium. IEEE Comp. Soc. Press.**, Los Alamitos, CA, USA, pp. 480-9, 1993.

[2] D.J., Kish, G.H., Heydt: "A generalization of the concept of impedance and a novel Walsh domain immittance" **IEEE Trans. on Power Delivery**, Vol. 9, n. 2, pp. 970-76, 1994.

[3] J., Sun, H., Grotstollen: "Walsh functions method for the control of active power filters", **Proc. Symposium Power Electronics Circuits. Hong Kong Polytech**, Hong Kong, pp. 112-15, 1994.

[4] F.H.J., Altuve, V.I., Diaz, M.E, Vazquez: "Fourier and Walsh digital filtering algorithms for distance protection", **Proc. IEEE Power Industry Computer Application Conference,** New York, NY, USA, pp. 423-8, 1995.

[5] R.K. Aggarwal, D.V., Coury, A.T., Johns, A. Kalam: "A practical approach to accurate fault location on extra high voltage teed feeders", **IEEE Trans. on Power Delivery**, Vol. 8, n. 3, pp. 874-883, 1993.

[6] A.T., Johns, R.K. Aggarwal: "Digital simulation of faulted EHV transmission lines with particular reference to very high speed protection", **Proc. IEE***, Vol. 123, N. 4,* pp. 353-9, 1976.

[7] D.V., Coury, H.G.F., Brito: "Digital filters applied to computer relaying", **Proc. IEEE International Conference on Power System Technology POWERCOM '98**, New York, NY, USA, pp. 1062-66, 1998.

[8] T. Lobos, J. Rezmer :" Wavelet transform for real-time estimation of transmission line impedance under transient conditions" **Electrical Engineering**, Vol. 84, n. 2, pp. 63-70, 2002

[9] K Shrivastava, D N Vishwakarma, : "Numerical Filtering Algorithms for Distance Protection of Transmission Lines : An Overview" **IEI - Electrical Engineering**, Vol. 89, pp. 19-27, 2009.

[10] M. Albu, A. Arie, S. Grigorescu; "Walsh functions based algorithm for distance protection. Relay characteristic in Walsh domain" **Conference on Precision Electromagnetic Measurements Digest CPEM1996**, pp. 53-54, 1996

[11] R., Micheletti, R., Pieri: "Walsh digital filters applied to distance protection", **IEEE Instrumentation and Measurement Technology Conference IMTC 2002**, Anchorage, AK, USA, pp. 487-490, 2002.

[12] R. Micheletti, R. Pieri: "Walsh digital filtering algorithm for high speed distance protection with improved behavior", **XVII IMEKO World Congress: Metrology in the 3rd Millennium***,* Dubrovnik, Croatia, pp. 847-850, 2003.

[13] R. Micheletti, R. Pieri:" Digital fault location estimator for power lines based on Walsh functions" **10th IMEKO TC-7 International Symposium**, Saint-Petersburg, Russia., pp. 419-423, 2004

NANOPOROUS NICKEL FOR ELECTROCHEMICAL ENERGY CONVERSION

Yong X. GAN *
Department of Mechanical, Industrial and Manufacturing Engineering
College of Engineering, The University of Toledo
2801 W. Bancroft St., Toledo, OH 43606, USA

ABSTRACT

Biodegradable products such as fruits, vegetables and milk have recently been studied as sustainable energy sources. In this work, preparation and electrocatalytic reactions of nanoporous materials in biodegradable fluids made from agricultural products were studied. In order to obtain high efficiency catalysts, electrochemical etching was conducted to selectively extract metallic elements from alloys to form porous structures. Electrocatalytic properties of the porous electrodes were characterized. Comparative studies on the electrochemical activities of the nanoporous metallic electrodes with bulk metallic catalysts were performed. It is found that nanoporous structures with high electroactive surface areas can be obtained through controlled electrochemical etching in diluted acids. The current density at the nanoporous electrode is three times higher than that of the bulk electrode with the same chemical compositions. A microfuel cell consisting of nanoporous nickel and platinum electrodes is made and tested.

Keywords:

Microfuel cell, Electrochemical energy conversion, Nanoporous metal, Catalysis; Electrochemical dealloying, Biomass, Sustainable energy.

1 INTRODUCTION

Nanoporous materials have been studied for a variety of applications, for example gas sensors [1], self-assembling templates [2, 3], hydrogen adsorbents [4, 5]. Nanoporous materials have played an important role in energy-related applications such as catalysts, and semiconductors in dye-sensitized solar cells [6]. There are increasing interest in using nanoporous materials for catalysis. For example, the photocatalytic activity of alkaline earth metal doped nanoporous TiO_2 was investigated for the degradation of bisphenol-A [7]. The alkaline earth metals used are magnesium and barium. The sol-gel technique was used to dope Mg^{2+} and Ba^{2+} into the parent nanoscale TiO_2. X-ray diffraction and FTIR analysis confirmed that Ba^{2+} was retained only on the surface of the TiO_2 as $BaCO_3$, while Mg^{2+} was kept in the substitutional sites.

There are several kinds of nanoscale catalysts according to the functionalities. Photocatalyst and electrocatalyst are two commonly used types. Titanium dioxide nanoporous electrodes for photocatalysis were produced by the re-anodization method [8]. Manganese, chromium and cobalt were doped into nanoporous anatase titanium dioxide. It is found that the photocatalytic efficiency is significantly enhanced due to the presence of Mn^{2+} and Cr^{3+} in the nanoporous TiO_2 scaffold. Mn^{2+} is the most efficient dopant, while the incorporation of Co^{2+} shows the least efficiency. TiO_2 supported bi-metallic catalysts (PtRu, PtFe) were used for methanol oxidation [9]. Yang and Xu [10] prepared nanoporous amorphous manganese oxide as an electrocatalyst for oxygen reduction in alkaline solutions. The amorphous manganese oxide catalyst is expected to have potential applications in metal-air batteries and low-power fuel cells. It is found that the amorphous manganese oxide catalyzes a two-electron reduction, converting molecular oxygen into HO_2^-. High concentration of lattice defects and active sites of the amorphous material are considered as the reasons for significant catalytic activity. Besides metal oxides, gold nanoparticles with a tubular structure are also used as catalysts towards the oxidation and the reduction of hydrogen peroxide (H_2O_2) [11]. As compared with a bulk gold electrode, the gold nanoelectrode has much higher sensitivity.

Nanoporous carbon were prepared using silica particle as the template [12, 13]. Such nanoporous carbon was used as electrochemical catalyst support for direct methanol fuel cell (DMFC) applications. For example, a PtRu (1:1) alloy was supported on the nanoporous carbon [13]. The structural properties of the carbon support played an important role in the metal disper-

*Tel: 419-530-6007; Fax: 419-530-8206; Email: yong.gan@utoledo.edu.

sion and the performance of the DMFC. The support with large pore size and high surface area was favorable for high dispersion of PtRu catalyst and easy formation of the triple-phase boundary. The catalysts with higher metal dispersion and structural integrity showed higher catalytic activities in the methanol electro-oxidation and the DMFC performance test. Heterogeneous catalysts for electrochemical reduction of CO_2 were developed from rhenium complexes. Such complexes were deposited onto a solid support consisting of nanocrystalline TiO_2 film on glass [14]. It was revealed that the nanoporous nature of the TiO_2 substrate increased the number of redox sites per unit surface area.

Noble metals such as palladium and gold as nanoclusters can be incorporated into the nanoporous domains of gel-type functional resins for oxidation catalysis in liquid phases [15, 16]. The surface properties of platinum catalysts based on various nanoporous matrices such as zeolite, silicate and niobiosilicate were investigated [17]. The locations of Pt in the nanopores and the interactions of metal-support were found dependent on the chemical composition and structure of the matrices. Pt may be electroplated onto nanoporous metals such as gold to obtained low Pt loading electrocatalysts for proton exchange membrane (PEM) fuel cells [18].

Nanoporous materials may be made via various methods. Chemical etching is a easy way to make nanoporous metals [19-21], while co-evaporation may be used for preparation of nanoporous metal oxides [22]. Nanoporous Au specimens with ligament-like structures were fabricated by dealloying [19]. Smith, Maaroof and Gentle [20] prepared nanoporous gold films through de-alloying $AuAl_2$. The nanoporous gold has void densities between 45% and 65%. Zhou, Jin and Xu [21] proposed another way to fabricate nanoporous gold film. The starting materials are copper and gold alloys thin films obtained through vacuum deposition. The formation of ultra-thin nanoporous gold films from the Au-Cu alloy films involved chemical etching by hydrochloric acid or by nitric acid.

It is well known that preparing nanoporous metals via chemical etching is very difficult to control in view of the reaction kinetics. For example, the change in the concentration of the acids, and the temperature fluctuation in the etchant solutions have significant effect on the ligament morphology, void density and strength of the nanoporous metals. During the process of pore formation, the capillary effect causes surface relaxation and affects the dimension of the nanoporous metals [23]. Another drawback of chemical dealloying is the environmental pollution due to the frequent use of high concentration, volatile enchants such as concentrated hydrochloric acid and/or nitric acid. One of the objectives of this work is to develop a better controllable dealloying process than chemical etching. Preliminary research results of electrochemical dealloying in low concentration acids to generate nanoporous nickel and gold will be presented. Electrocatalytic property of the nanoporous Ni in agricultural products for sustainable energy applications will be evaluated.

2 MATERIALS AND EXPERIMENTAL METHODS

In this part, the materials used in the work will be described first. Experimental setup will be introduced second. Th procedures for electrochemical dealloying will be given. Finally, characterization of the electrocatalytic property will be presented.

2.1 Materials

Copper-nickel wire with the diameter of 0.5 mm was purchased from Alfa Aesar. The composition of the Cu-Ni wire is Cu:Ni=55:45 wt%. Platinum with the diameter of 0.25 mm and the purity of 99.997 wt% was also purchased from the same supplier. Other chemicals such as hydrochloride acid, sulfuric acid, acetone with ACS purities were ordered from Alfa Aesar. Biodegradable materials such as tomato, milk, chocolate drink and grape fruit juice were used in characterizing the electrocatalytic behavior of the nanoporous Ni.

2.2 Experimental Setup

In this work, a three electrode system was set up for both electrochemical dealloying and electrocatalytic property characterization. The three electrodes are the Ag/AgCl reference electrode, the work electrode, and the platinum counter electrode.

2.3 Electrochemical Dealloying

Electrochemical etching is widely used for metallurgical specimen preparation, surface polishing and machining. Recently, this technique has been studied for selective removal of alloy elements from metallic alloys. In this work, removal of Cu from Cu-Ni alloy by electrochemical etching in diluted HCl and H_2SO_4 was conducted following the steps as described below.

First, HCl aqueous solution with the concentration of 9 wt% was made. The H_2SO_4 used has the concentration of 25 wt%. Second, the Cu-Ni was made into a work electrode. The electrode was assembled with the Ag/AgCl reference electrode, and the platinum counter electrode in the cell containing HCl or H_2SO_4 solution. Under controlled potentials generated by positive cyclic voltammetrical (CV) scan, selective dissolving of metal elements was achieved. As a result, nanoporous Ni was obtained. During the CV scan, changes in anodic current at the work electrodes were recorded. Such information will be used to explain the electrochemical dealloying mechanisms.

2.4 Electrocatalytic Property Characterization

The electrocatalytic activity of the nanoporous Ni in the biodegradable substances of fresh tomato, milk, chocolate drink, and grape fruit juice, respectively, was characterized by cyclic voltammetry (CV). First, we tested the electrochemical behavior of the nanoporous nickel electrode and a bulk Ni electrode in grape fruit juice. The electroactive surface area effect of the nanoporous Ni is revealed. Tomato was also used to form a cell to examine the effect of surface area on the catalytic oxidization behavior of the nanoporous Ni electrode. Subsequently, the electrocatalytic behaviors of the nanoporous Ni in the milk and chocolate drink(milk) were characterized. During electrocatalytic property test, the same three-electrode cell was used. The nanoporous nickel electrode was used as the catalyst support electrode (work electrode or anode). The platinum wire was used as the counter electrode, and the Ag/AgCl electrode was used as the reference electrode. The initial concentration of KCl filled in the reference electrode is 3M. The potential of the Ag/AgCl electrode versus the standard hydrogen electrode was calculated to be equal to 0.199 V. The Electrochemical Analyzer was used to provide the CV scan functions. The CV scanning range was controlled over a wide potential range of 0.0 V $\sim$ 1.5 V. The initial potential was 0.0 V. The polarity was set as positive. The scan rate was 0.1 V/s. The data acquisition interval was 0.001 V. The quiet time was 2 s, and the sensitivity was 10^{-3} (A/V).

3 RESULTS AND DISCUSSION

3.1 Cyclic Voltammograms of Dealloying

The current density-potential (I-V) response associated with the electrochemical dealloying of copper from the Cu-Ni alloy was obtained. Fig. 1(a) is the cyclic voltammogram showing the current density change at the Cu-Ni alloy electrode in the 25% H_2SO_4 solution, while Fig. 1(b) is the relationship between the specific current and the potential. The results in Fig. 1(a) reveals that in the potential region of less than 0.3 V, the main reaction is the dissolution of Ni into the electrolyte. With the increase of potential, Ni is passivated, and the dissolving of copper becomes the main reaction.

It is noted that which element will be removed for the CuNi alloy through electrochemical dealloying depends on the potential level. Since Ni is less noble than Cu, in the low potential range from 0 to 0.3 V, Ni will dissolve first. Nanoporous copper scaffold is obtained. At higher potential levels, the newly formed surface oxide layer of NiO (passive film) protects the Ni from dissolution, the more noble metal, copper, is the element being removed from the alloy. Therefore, nanoporous Ni is obtained in the potential range from 0.3 to 1.5 V.

In the HCl electrolyte, the reaction of Ni dissolution occurs in a wide potential range up to 0.52 V because NiO passive film is more difficult to form in the presence of Cl^-. This can be seen from the results shown in Fig. 2(a), the cyclic voltammogram showing the current density change at the Cu-Ni alloy electrode in the 9% HCl. At the higher potential level greater than 0.52 V, copper dissolution becomes the main reaction. Fig. 2(b) is the relationship between the specific current and the potential, showing the same trend of transition from Ni dissolution to Cu dissolution at 0.52 V.

3.2 Effect of Nanopores on Oxidation Behaviors

In order to examine the effect of nanopore on the electrocatalytic behaviors of electrodes in biodegradable fluids, comparative studies on nanoporous nickel and bulk nickel were conducted. The nanoporous electrode was prepared by electrochemical dealloying of the CuNi alloy wire with the diameter of 0.5 mm in the HCl electrolyte as described in Section 3.1. The bulk nickel electrode is made of a pure nickel wire with the diameter of 1.0 mm. Electrochemical catalysis behaviors of both bulk Ni and nanoporous Ni in the grape fruit juice were studied by cyclic voltammetry and the results are shown in Figs. 3(a) and (b). It is noted that the current density at the nanoporous electrode, as shown in Fig. 3(b), is calculated based on the initial area of the alloy before dealloying. The cyclic voltammogram of the bulk Ni as shown in Fig. 3(a) has the reaction controlled feature as the porous Ni electrode. However, the current density at the nanoporous Ni is as 4 times high as that at the bulk nickel electrode, which reveals the nanoporous electrode is more effec-

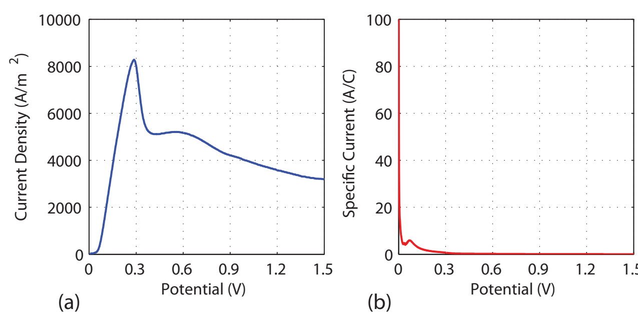


Figure 1. CYCLIC VOLTAMMOGRAMS ASSOCIATED WITH ELECTROCHEMICAL DEALLOYING OF Cu-Ni IN 25% H_2SO_4 SOLUTION: (a) I-V RESPONSE; (b) SPECIFIC CURRENT CHANGE.

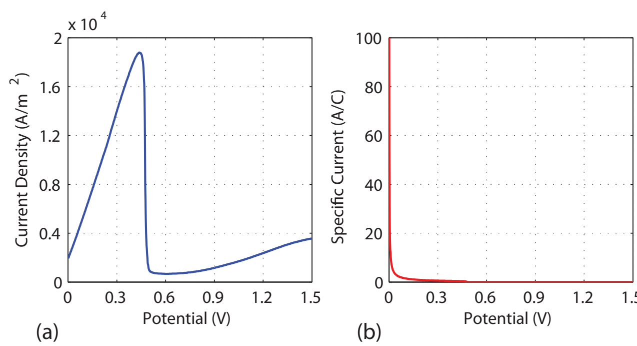


Figure 2. CYCLIC VOLTAMMOGRAMS ASSOCIATED WITH ELECTROCHEMICAL DEALLOYING OF Cu-Ni IN 9% HCl SOLUTION: (a) I-V RESPONSE; (b) SPECIFIC CURRENT CHANGE.

tive than the bulk Ni in electrochemical catalytic oxidation of the grape fruit juice.

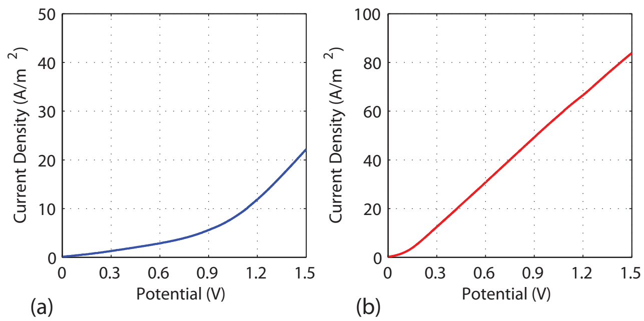


Figure 3. CYCLIC VOLTAMMOGRAMS OF Ni ELECTRODES IN GRAPE FRUIT JUICE: (a) WITHOUT PORE; (b) WITH PORES.

The electrocatalytic behaviors of the above used Ni electrodes without nanopores and with nanopores in fresh tomato were studied. At the same potential, the current density at the bulk nickel electrode shown in Fig. 4(a) is lower than that at the nanoporous nickel electrode shown in Fig. 4(b). This trend is the same as found in the grape fruit juice case. The oxidation of tomato on the bulk electrode has three distinct stages as mark by “A”, “B”, and “C” on the cyclic voltammogram of Fig. 4(a). In stage “A” and “C”, oxidation reactions are the controlled processes. In stage “B”, the diffusion controlled feature is shown. At the nanoporous Ni electrode, only the reaction controlled behavior in the tomato juice was found, which is revealed by the cyclic voltammetry results in Fig. 4(b).

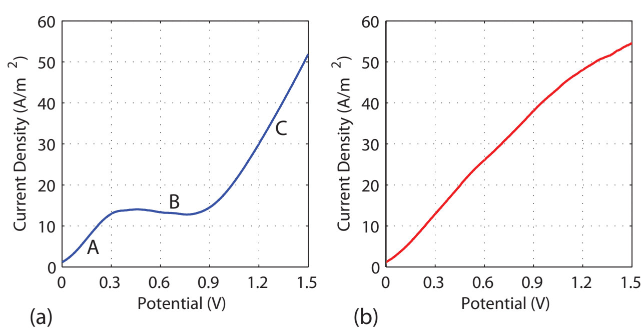


Figure 4. CYCLIC VOLTAMMOGRAMS OF Ni ELECTRODES IN TOMATO: (a) WITHOUT PORE; (b) WITH PORES.

3.3 Oxidation Behaviors during the Reversed Scan

Reversed cyclic voltammetry scans were performed to show the reversibility of the nanoporous Ni for electrochemical catalytic oxidation in biodegradable fluids. Figs. 5(a) and (b) are the cyclic voltammograms of nanoporous Ni in the chocolate drink. The solid blue line in Fig. 5(a) reveals the electrochemical oxidation features of the chocolate drink at the nanoporous Ni electrode in the forward scan. The dotted red line in Fig. 5(a) shows the electrochemical oxidation behaviors of the chocolate drink at the nanoporous Ni electrode in the reversed scan. The oxidation reaction controlled features as revealed in both forward scan and backward scan are the same, which means that the reversibility is well kept. Peak “A” as shown in Fig. 5(b) represents the start of electrochemical oxidation of the chocolate drink on the nanoporous Ni at the potential around 0.3 V.

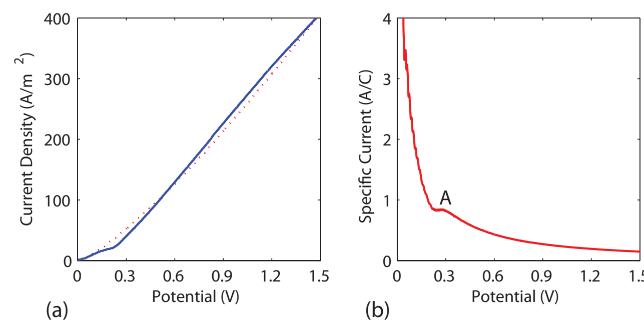


Figure 5. CYCLIC VOLTAMMOGRAMS OF THE NANOPOROUS Ni ELECTRODE IN CHOCOLATE DRINK: (a) I-V RESPONSE; (b) SPECIFIC CURRENT CHANGE.

3.4 Effect of Fermentation

The grape fruit juice was sealed in a glass bottle and kept at the room temperature for ten days. Fermentation was observed as revealed by the following signs. The original red color is faded and some white precipitates were found at the bottom of the bottle. Typically, fermentation also results in the formation of alcoholic products, which may cause change in electrochemical oxidation behaviors of the fluid. To verify this, the cyclic voltammogram of nanoporous Ni in the fermented grape fruit juice is shown in Fig. 6. As compared with the catalytic behavior of nanoporous Ni in unfermented grape fruit juice shown in Fig. 3(b), the current density in fermented fluid is about one third higher. This indicates that fermentation facilitates the electrochemical catalytic oxidation of the fluid on the porous Ni electrode. Such results are helpful in recycling stale juices and drinks for green fuel applications. To recycle these discarded fruits and drinks and convert them into useful energy are meaningful to save energy and clean the environment.

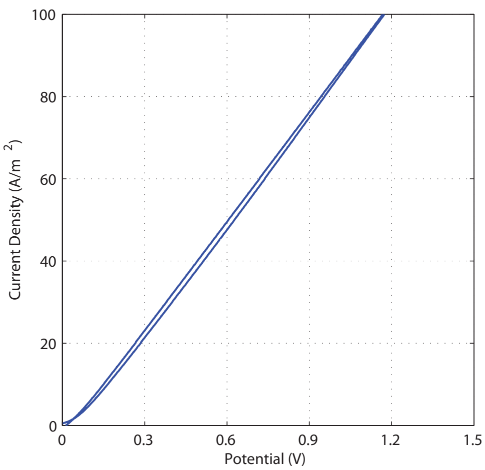


Figure 6. CYCLIC VOLTAMMOGRAM OF THE NANOPOROUS Ni ELECTRODE IN FERMENTED GRAPE FRUIT JUICE.

3.5 Voltage Generated by the Porous Ni

Finally, we made a microfuel cell consisting of a nanoporous nickel anode, white milk solution and a platinum cathode. The plot of voltage v.s. time is given in Figure 7. The stable output of voltage is about 0.135 V. Obviously, in order to build a practical power generation device, many of such cells are needed.

4 CONCLUSIONS

Electrochemical dealloying and electrocatalytic behavior of nanoporous nickel in biodegradable fluids for green fuel applications were performed. High electroactive surface areas in nanoporous nickel were obtained. Based on the preliminary research, the following concluding remarks can be made.

Studies on electrochemical dealloying of the Cu55Ni45 alloy in the 25% sulphuric acid reveal that in the potential region of less than 0.3 V, the main reaction is the dissolution of Ni into the electrolyte. With the increase of potential, Ni is passivated, and the dissolving of copper becomes the main reaction. Therefore, which element is removed for the CuNi alloy through electrochemical dealloying is dependent on the level of the potential used. Since Ni is less noble than Cu, in the low potential range from 0 to 0.3 V, Ni dissolves first. Nanoporous copper scaffold is obtained. At higher potential levels, the newly formed surface oxide layer of NiO (passive film) protects the Ni from dissolution. The more noble metal, copper, is the element being removed from the alloy. Nanoporous Ni is obtained in the potential range from 0.3 to 1.5 V.

In the 9% HCl electrolyte, the reaction of Ni dissolution from the Cu-Ni alloy occurs in a wide potential range up to 0.52 V because NiO passive film is more difficult to form in the presence of Cl^-. At the higher potential level greater than 0.52 V, copper dissolution becomes the main reaction. The transition from Ni dissolution to Cu dissolution at 0.52 V is revealed by cyclic voltammetry.

Comparative studies on nanoporous and bulk specimens show that a nanoporous electrode is more effective than a bulk electrode in electrochemical catalytic oxidation of biodegradable fluids. The cyclic voltammogram of the bulk Ni has the reaction controlled feature. The current density at the nanoporous Ni is as 4 times high as that at the bulk nickel electrode.

Fermentation changes the electrochemical oxidation behaviors of biodegradable fluids, facilitating the electrochemical catalytic oxidation of the fluids on the nanoporous nickel electrode. The cyclic voltammogram of the nanoporous Ni in the fermented grape fruit juice shows that the current density in fermented fluid is about one third higher than that in the unfermented fluid, which indicates that the fermentation makes the oxidation of the fluid on the nanoporous Ni electrode even easier. Such preliminary results may be helpful in recycling stale fruits or drinks for green fuel applications.

ACKNOWLEDGMENT

This work was supported by the research start-up fund and the Summer Faculty Research Fellowship from The University of Toledo.

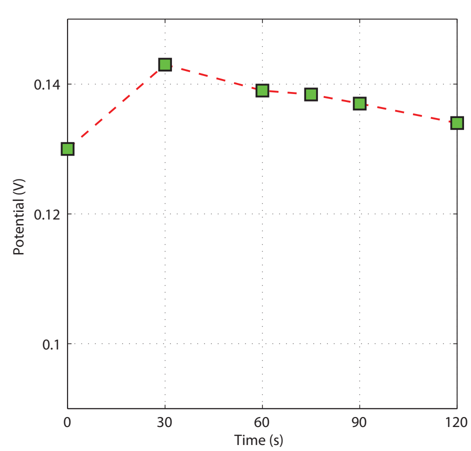


Figure 7. VOLTAGE OF NANOPOROUS Ni/WHITE MILK/Pt CELL.

REFERENCES

[1] T. I. Gorbanyuk, A. A. Evtukh, V. G. Litovchenko, V. S. Solntsev, “Nanoporous silicon doped by Cu for gas-sensing applications”, **Physica E: Low-dimensional Systems and Nanostructures**, Vol. 38, No. 1–2, 2007, 211–214.

[2] S. V. Mattigod, G. E. Fryxell, K. E. Parker, “Anion binding in self-assembled monolayers in mesoporous supports (SAMMS)”, **Inorganic Chemistry Communications**, Vol. 10, No. 6, 2007, pp. 646–648.

[3] R. Kumar, M. Chaudhari, S. K. Chakarvarti, “Morphology and collective I-V characteristics of template synthesized nano metal-semiconductor heterojunctions”, **Materials Letters**, Vol. 61, No. 7, 2007, pp. 1580–1582.

[4] P. Bénard, R. Chahine, “Storage of hydrogen by physisorption on carbon and nanostructured materials”, **Scripta Materialia**, Vol. 56, No. 10, 2007, pp. 803–808.

[5] J. J. Vajo, G. L. Olson, “Hydrogen storage in destabilized chemical systems”, **Scripta Materialia**, Vol. 56, No. 10, 2007, pp. 829–834.

[6] S. Pavasupree, Y. Suzuki, S. Pivsa-Art, S. Yoshikawa, “Synthesis and characterization of nanoporous, nanorods, nanowires metal oxides”. **Science and Technology of Advanced Materials**, Vol. 6, No. 3–4, 2005, pp. 224–229.

[7] N. Venkatachalam, M. Palanichamy, B. Arabindoo, V. Murugesan, “Alkaline earth metal doped nanoporous TiO_2 for enhanced photocatalytic mineralisation of bisphenol-A”, **Catalysis Communications**, Vol. 8, No. 7, 2007, pp. 1088–1093.

[8] E. B. Gracien, J. Shen, X. R. Sun, D. Liu, M. Li, S. D. Yao, J. Sun, “Photocatalytic activity of manganese, chromium and cobalt-doped anatase titanium dioxide nanoporous electrodes produced by re-anodization method”, **Thin Solid Films**, Vol. 515, 2007, No. 13, pp. 5287-5297.

[9] M. Hepel, I. Dela, T. Hepel, J. Luo, C. J. Zhong, “Novel dynamic effects in electrocatalysis of methanol oxidation on supported nanoporous TiO_2 bimetallic nanocatalysts”, **Electrochimica Acta**, Vol. 52, No. 18, 2007, pp. 5529-5547.

[10] J. Yang, J. J. Xu, “Nanoporous amorphous manganese oxide as electrocatalyst for oxygen reduction in alkaline solutions”, **Electrochemistry Communications**, Vol. 5, No. 4, 2003, pp. 306-311.

[11] M. Delvaux, A. Walcarius, S. Demoustier-Champagne, “Electrocatalytic H_2O_2 amperometric detection using gold nanotube electrode ensembles”, **Analytica Chimica Acta**, Vol. 525, No. 2, 2004, pp. 221-230.

[12] A. N. Vasiliev, L. V. Golovko, V. A. Povazhny, E. Zlotnikov, J. Chen, J. G. Khinast, “Functionalized nanoporous carbon as a catalyst for Suzuki coupling reactions”, **Microporous and Mesoporous Materials**, Vol. 101, No. 3, 2007, pp. 342-347.

[13] P. Kim, H. Kim, J. B. Joo, W. Kim, I. K. Song, J. Yi, “Preparation and application of nanoporous carbon templated by silica particle for use as a catalyst support for direct methanol fuel cell”, **Journal of Power Sources**, Vol. 145, No. 2, 2005, pp. 139-146.

[14] F. Cecchet, M. Alebbi, C. A. Bignozzi, F. Paolucci, “Efficiency enhancement of the electrocatalytic reduction of CO_2: fac-[Re(v-bpy)$(CO)_3$Cl] electropolymerized onto mesoporous TiO_2 electrodes”, **Inorganica Chimica Acta**, Vol. 359, No. 12, 2006, pp. 3871-3874.

[15] C. Burato, P. Centomo, G. Pace, M. Favaro, L. Prati, B. Corain, “Generation of size-controlled palladium(0) and gold(0) nanoclusters inside the nanoporous domains of gel-type functional resins: Part II: Prospects for oxidation catalysis in the liquid phase”, **Journal of Molecular Catalysis A: Chemical**, Vol. 238, No. 1–2, 2005, pp. 26-34.

[16] B. Corain, C. Burato, P. Centomo, S. Lora, W. Meyer-Zaika, G. Schmid, “Generation of size-controlled gold(0) and palladium(0) nanoclusters inside the nanoporous domains of gel-type functional resins: Part I: Synthetic aspects and first catalytic data in the liquid phase”, **Journal of Molecular Catalysis A: Chemical**, Vol. 225, No. 2, 2005, pp. 189-195.

[17] I. Sobczak, J. Grams, M. Ziolek, “Surface properties of platinum catalysts based on various nanoporous matrices”, **Microporous and Mesoporous Materials**, Vol. 99, No. 3, 2007, pp. 345-354.

[18] R. Zeis, A. Mathur, G. Fritz, J. Lee, J. Erlebacher, “Platinum-plated nanoporous gold: An efficient, low Pt loading electrocatalyst for PEM fuel cells”, **Journal of Power Sources**, Vol. 165, No. 1, 2007, pp. 65-72.

[19] M. Hakamada M. Mabuchi, “Mechanical strength of nanoporous gold fabricated by dealloying”, **Scripta Materialia**, Vol. 56, No. 11, 2007, pp. 1003-1006.

[20] G. B. Smith, A. I. Maaroof, A. Gentle, “Homogenized Lorentz-Drude optical response in highly nanoporous conducting gold layers produced by de-alloying”, **Optics Communications**, Vol. 271, No. 1, 2007, pp. 263-268.

[21] H. Zhou, N. Jin, W. Xu, “New approach to fabricate nanoporous gold film”, **Chinese Chemical Letters**, Vol. 18, No. 3, 2007, pp. 365-368.

[22] T. Tesfamichael, N. Motta, T. Bostrom, J. M. Bell, “Development of porous metal oxide thin films by co-evaporation”, **Applied Surface Science**, Vol. 253, No. 11, 2007, pp. 4853-4859.

[23] D. A. Crowson, D. Farkas, S. G. Corcoran, “Geometric relaxation of nanoporous metals: The role of surface relaxation”, **Scripta Materialia**, Vol. 56, No. 11, 2007, pp. 919-922.

Selection of the best Biomass-to-Bioenergy route for its implementation in the European Energy sector. An Integrated Efficiency, Economic and Environmental Analysis

Anna Sues, Hubert J. Veringa
Chemical Engineering and Chemistry, Environmental Technology Group
TU/e Eindhoven University of Technology
Eindhoven, North Brabant 5600 MB, the Netherlands

ABSTRACT

Biomass availability is rather limited in Europe and, hence, it is of crucial importance to determine the optimal biomass-to-energy conversion pathway. This selection is somehow complex as there could be antagonistic motivations coming from industrial stake-holders, politicians, scientists or the society. Consequently, the aim of this paper is to present different biomass-to-biofuels alternatives that follows various economic, environmental and/or social drivers. Results are also compared with European Directives 2001/77/EC and 2009/28/EC. In General, maximizing bio-electricity over other biofuels turns out to be the best economical and environmental option. Combined with solar and wind energy, about 31% of the electricity production by 2020 could be renewable, i.e., 10 points higher than the target of Directive 2001/77/EC. If biomass is conducted to SNG production, fossil natural gas imports could be reduced by 1.63 EJ/yr in 2020, although this alternative implies higher costs and less CO_2 savings than the previous bio-electricity solution. In case of promoting Fischer-Tropsch fuels, the share of biofuels in transport will be 9.5%, which is slightly below the 10% share target of Directive 2009/28/EC. H_2 is disregarded as feasible option for transport due to several technological barriers, although it would lead to substantial CO_2 savings at a moderate price. Conversely, methanol results in the worst environmental solution as CO_2 emissions are larger than those of conventional fossil fuels.

Keywords: Biofuel, biomass, gasification, efficiency, Life Cycle Analysis (LCA), techno-economic evaluation.

1. INTRODUCTION

Among all renewable technologies, biomass is the only source that could be used to produce biofuels for the transport sector. However, unlike wind or solar energy, biomass availability is limited and, hence, it is of crucial importance to select the most convenient biomass-to-biofuel conversion route. Nowadays, 2nd generation biofuels, and gasification technology in particular, are gaining interest for its implementation in the medium-term future due to its potentially higher efficiency and lower cost and CO_2 emissions. In this paper we present the evaluation of five biofuels (i.e., Syntethic Natural Gas (SNG), methanol (MeOH), Fischer-Trospch fuels (F-T), H_2 and electricity) for their introduction in the European Energy sector by 2020. The most convenient biomass-to-bioenergy technology will be selected following technological, economic and environmental drivers.

In Europe, forest and agricultural residues are stochastically distributed, leading to definite areas where the concentration of biomass differs substantially among them. In some cases, those areas do not correspond with the established country borders and, hence, new "bio-borders" are suggested to maximize the amount of biofuels that can be later produced in Europe. This "re-drawing" is especially sensitive for the processes that require a relatively large scale in order to operate at a more competitive price (e.g., Fischer-Tropsch or methanol plants). SNG and hydrogen production are profitable at medium plant sizes, thus giving a combined national and "bio-borders" scenario. Similarly, since logistics costs have a major impact in electricity final price, borders follow also a combined scenario. Once the "bio-border" for different biofuels and biopower production is determined, the next step is to identify which alternative is the best for the European energy market. Answering this question is somehow complex as the society, the scientific community, industry, or the politicians have their own motivations. In this paper, biofuels or bioelectricity implementation within European countries is discussed under the several scenarios presented in Table 1. When co-firing is prioritized (i.e., scenarios I-A, II, VI), coal is replaced up to 10% in weight basis. In other scenarios, less coal is replaced as biomass is primarily consumed for biofuels production.

***Table 1**: Biofuels and bioelectricity scenarios, which takes into account preferences from industry, politicians, scientists or the society.*

Code	Preference	Description
I-A	Max. bioelectricity production[a]	A fraction of biomass is used for co-firing and the rest is send in new BIGCC[d]
I-B		All biomass is used in new BIGCC
II	Potential biofuels introduction in the medium-term	A fraction of biomass is used for co-firing. The rest is consumed in new FT plants, and remainings for SNG production.
III	Max. biofuels[b] share in transport	Biomass is primarily consumed in new FT plants. The rest is used for SNG plants.
IV	Max. FT-diesel (Oil companies)	Biomass is primarily consumed in new FT plants. The rest is used for cofiring.
V	Max. SNG production	Biomass is primarily consumed in new SNG plants. The rest is used for cofiring.
VI	Hypothetical[c] max CO_2 reduction	Our initial hypothesis is that maximal CO_2 reduction is obtained when a fraction of biomass is used for co-firing and the rest for SNG plants. However, later analyses will confront our initial hypothesis.
VII	Max. Hydrogen production	Biomass is primarily consumed in new H_2 plants. The rest is used for cofiring.
VIII	Max. MeOH production	Biomass is primarily consumed in new MeOH plants. The rest is used for cofiring

(a): According to the European Directive 2001/77/EC, about 21% of the electricity must be produced from renewable sources by 2010 in EU25.

(b): According to the European Biofuels Directive (2003/30/EC), the share of biofuels in transport should achieve the target of 10% by 2020.

(c): The EU countries have committed to reduce greenhouse gas emissions during the first Kyoto commitment period 2008-2012 by 5% compared to the 1990 reference year (COM 2006).

(d): "New plants" refers to new BIGCC (Biomass Integrated Combined Cycle) plants that will operate on 100% biomass (i.e., no co-firing).

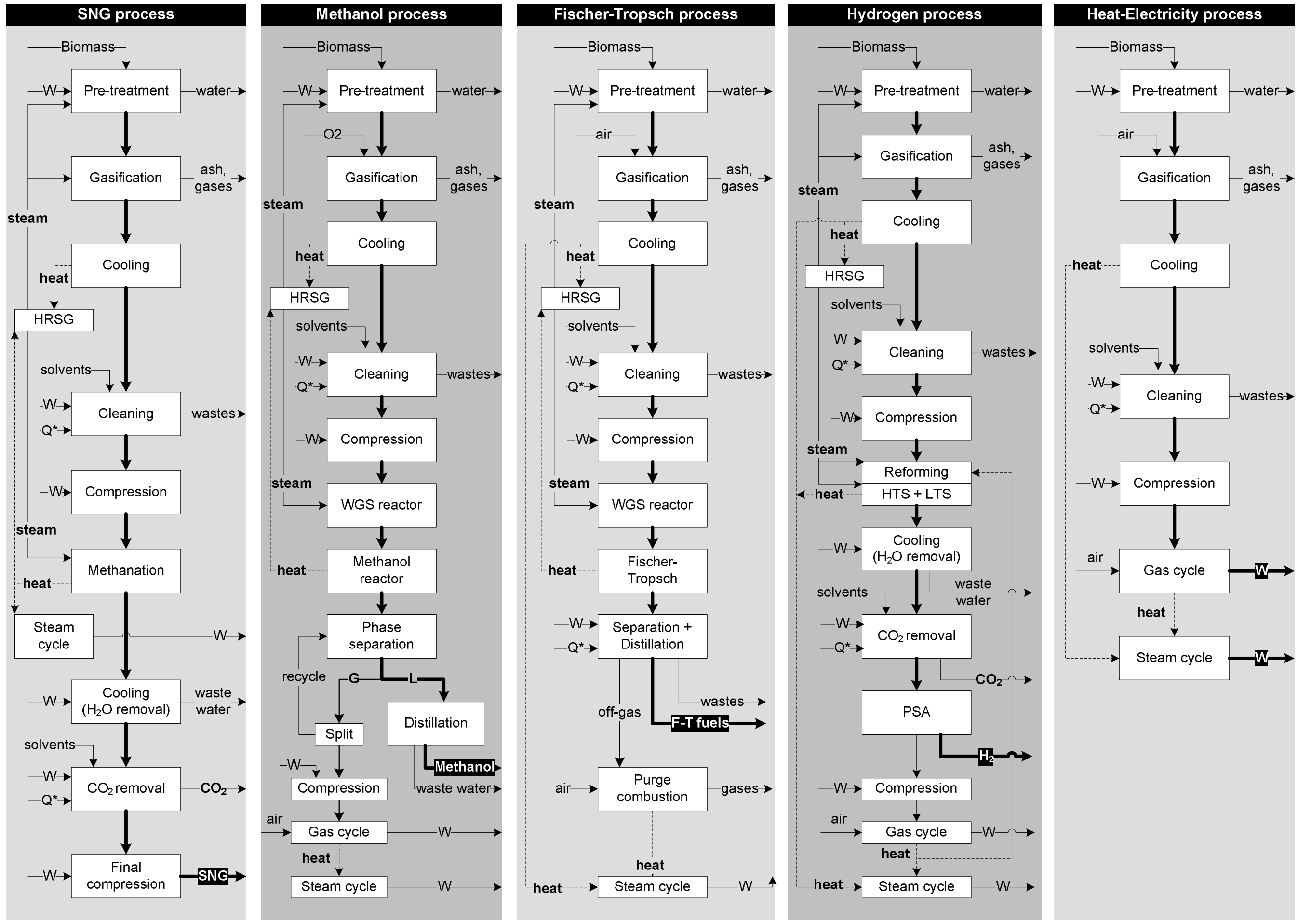


Figure 1: *Overview of the 5 biowaste-to-biofuels conversion routes (W and Q represent work and heat flows, respectively) [1].*

2. BIOFUELS TECHNOLOGY SELECTION

The five different biomass-to-biofuels routes have been modeled in Aspen Plus. The corresponding diagram block of each process is depicted in Figure 1. In all cases, pre-treatment is required to adjust particles size and moisture content to 10%. Gasification is the core operation unit for the 5 biofuels chains although the working conditions, the oxidizing agent (i.e., steam, air or O_2) and its design are different among them. After gasification, biomass is converted into the so-called syngas, a mixture of mainly CO and H_2, although other gases such as CO_2, CH_4, C_2H_6 or C_2H_4 are also produced. In subsequent units, undesired by-products (mainly H_2S, NH_3, and HCl) are removed in order to avoid deactivation of the catalysts used downstream, damage engines, boilers or turbines, as well as to minimize SO_x and NO_x formation when burning remaining unconverted gases. In all the production chains, H_2S is removed at 55^{o}C and 1bar in a MDEA scrubber system, and the solution is regenerated in a stripping column working at 110^{o}C and 2 bar. MDEA is selected due to its low energy requirement for regeneration and higher selectivity over H_2S compared to CO_2. In fact, at this stage of cleaning, CO_2 removal is not desired as this gas is a reactant for some downstream catalytic reactors. NH_3 is removed in a subsequent unit using a H_2SO_4 solution. HCl and other halogens can be removed by injecting sodium or calcium-based powdered absorbents into the gas streams and by removing them in the de-dusting stage (e.g., cyclones). After the cleaning stages, the H_2:CO ratio is adjusted in a WGS (water-gas-shift) reactor by adding a specific amount of superheated steam. Outlet gases then undergo a series of catalytic reactions in a specific reactor for each biofuel. Later stages of the process comprise upgrading and final compression in order to achieve the specifications imposed by the market.

For the electricity generation process, syngas leaving the gasifier is cleaned and then send to a combine Brayton-Rankine cycle. A single expansion turbine is used for the Brayton cycle where exhaust gases are expanded from 1200^{o}C and 15 bar to 1 bar and ~ 605^{o}C. This remaining hot stream is used to provide part of the heat required for the Rankine cycle, in which three steam turbines are operated with inter-heating. In this cycle, steam is expanded from 200 bar to ~ 0.07 bar.

Moreover, for a better overall process efficiency, heat supply and demand are carefully matched so that more high quality heat is left to produced superheated steam that will be later used for electricity production (Rankine cycles), steam gasification, biowastes drying, and H_2:CO adjustment in the WGS reactors. In particular, a considerable amount of heat is recovered after gasification as the syngas needs to be cooled down prior to cleaning and compression stages. Another source of heat is taken during cooling of methanation and Fischer-Trospch or methanol reactors. However, extra fuel is still needed to cover the energy demand of the biofuels plant. In previous studies, we have evaluated to burn either an extra amount of biomass or natural gas to meet energy requirement of the plants. Our results show that the best configuration is to burn an extra amount of biomass and take electricity from the grid.

3. OWN MULTIDIMENSIONAL "3-E" MODEL: INTEGRATION OF EFFICIENCY, ECONOMIC AND ENVIRONMENTAL PARAMETERS

3.1 Model definition

Sustainability of a process is rather difficult to evaluate and quantify as many factors are involved and they are also given in different units. For instance, environmental impact is measured in ppm (or gCO_2eq/kg biofuel for the case of global warming evaluation), whereas production costs are given in €/$GJ_{biofuel}$, and efficiency is calculated as the ratio of energy output divided by the energy input (i.e., MW_{output}/MW_{input}). Hence, conversion factors are required if we intend to give one unique parameter in order to be able to communicate with scientists, legislators and economists at once.

In our model, mass and energy balances from Aspen Plus simulations are used to calculate exergetic and energy efficiency of the 5 biofuels conversion routes (i.e., SNG, MeOH, FT, H_2 and electricity), as shown in the left-hand sequence of Figure 2. Aspen Plus simulations are coupled to Aspen Icarus in order to calculate the biofuels production price "ex-works'. This price is determined by fixing and investor internal rate of return (IRR) of 12%, and with 50% of capital leverage (i.e., 50% of the TCI is borrowed from banks). Final end-user price is obtained by adding logistic and distribution costs. Those calculations are iterated for different plants sizes (i.e., from 1 to 5000 MW_{fuel}) and 24 European countries in order to determine the optimal plant scale and location for each biofuel. In the last stage, environmental impact of each configuration is integrated into the economic evaluation by calculating an ecotax value (i.e., €/ton CO_2 that would equalize biofuels and conventional fossil fuel price).

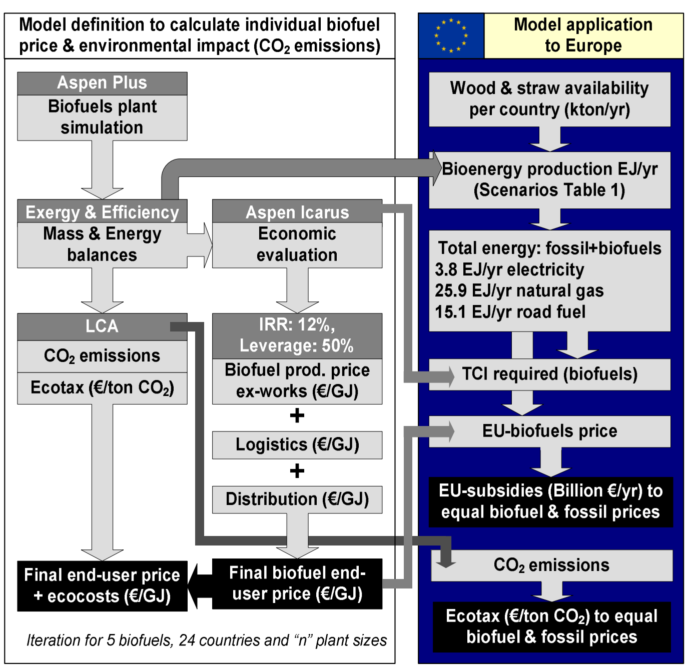


Figure 2: Schematic representation of our model.

3.2 Model application to Europe

The multidimensional model of previous section 3.1 is applied to Europe to select the best scenario of Table 1. For that purpose, several stages are needed (see Figure 2, right-hand sequence). Firstly, forestry and straw residues availability within 24 European countries is calculated from different values found in literature [2-7]. Subsequently, total biofuel and/or bioelectricity generation in Europe is determined by applying efficiency values from our model. Produced biofuels can substitute part of the fossil fuels consumption by 2020, although their share will be different for each scenario in Table 1. In any case, fossil energy is still needed in order to meet the energy demand by 2020, whose values have been stipulated in the report of Mantzos et al [8]. In particular, we assume that biofuels and bioelectricity can be introduced in 3 sectors:

- Electricity production from coal (i.e., **3.77 EJ/year**). This figure implies that about **8.34 EJ/yr** of coal is consumed.
- Natural gas consumption for several uses (e.g., electricity, heating, or other industrial applications), which accounts for **25.90 EJ/year**
- Fossil fuel consumption in the road transport, which sums **15.10 EJ/year**. This figure includes public road transport, private cars, motorcycles, and trucks.

Optimal biofuels plant scales are deducted from the iteration of our model in section 3.1. This parameter, together with biomass availability across Europe, fixes the number of plants that could be built in each scenario in Table 1, which in turn, allows the calculation of total capital investment (TCI). Following the profitability analysis of our model (i.e., IRR equal to 12%), final end-user prices for each biofuels and country are obtained. However, since biofuels prices are always higher than conventional fossil fuel prices, governmental subsidies would be needed to equalize prices and do not charge the final consumer. An alternative to public assistance is to calculate an ecotax (i.e., €/ton CO_2), which would notably penalize fossil fuels as their CO_2 emissions are notably higher.

4. RESULTS

Results are presented following the structure of previous section 3. Hence, the first section of this paper is dedicated to the individual values for each biofuel, whereas the second part includes results of the different scenarios in Table 1.

4.1 Specific biofuels efficiency, final price and ecotax

Figure 3 depicts the final biofuel 'end-user' price for a specific country and at different plant sizes. However, it should be mentioned that the represented biofuels prices accounts only for Austria, which is in the "average" side together with Eastern countries. In effect, biofuels production turns out to be economically more competitive in Southern, Baltic countries and UK, biofuels, whereas it is notably more expensive in Northern and Scandinavian states. This Figure 3 is completed by representing the effect of scale on the efficiency of the processing plants (i.e., right y-axis). As observed, both parameters are intrinsically connected as higher efficiencies are translated into lower Total Production Costs (TPC). Moreover, according to the William's equation, TCI does not follow a straight line when increasing the plant size. On the contrary, it is represented by an exponential relationship with a scaling factor in the range of 0.4 to 0.8. Conversely, logistics are directly dependent on the scale, whereas distribution costs are assumed to be constant (i.e., 3.61 €/GJ_{SNG}, 3.44 €/GJ_{FT}, 4.32 €/GJ_{FT}, 10 €/GJ_{H2}, 0.01€/GJ_{elec}). Figure 3 also identifies the optimal plant scale for each biofuel. As observed, electricity and SNG production is more profitable at lower sizes (i.e., 100 and 200 MW_{fuel} respectively), followed by H_2 (i.e., 500 MW_{H2}), whereas MeOH and FT-fuels generation require larger plants (i.e., 1000 MW_{fuel}). Comparison among biofuels reveals that SNG and electricity prices are on the lowest side (i.e., 19 and 23 €/GJ respectively for Austria). Conversely, FT-fuel, MeOH and H_2 are more expensive and yield close values (i.e., 26, 27 and 28 €/GJ). However, it should be mentioned that H_2 price is much higher than SNG due to its intensive distribution costs. Similar conclusions can be drawn in terms of exergetic efficiency. In effect, SNG and electricity production is more efficient (i.e., up to 45.5% for both biofuels), followed by MeOH and FT (i.e., up to 43.9%). Hydrogen efficiency (i.e., 42.5%) is notably penalized by the compression requirements of the pipelines distribution system. Results for the rest of European countries follow similar trends, although, as aforementioned, quantitative values are different for each country. Same analysis is done for straw residues in Figure 4, where it is observed that prices are always higher due to lower efficiencies and larger pre-treatment and logistics costs.

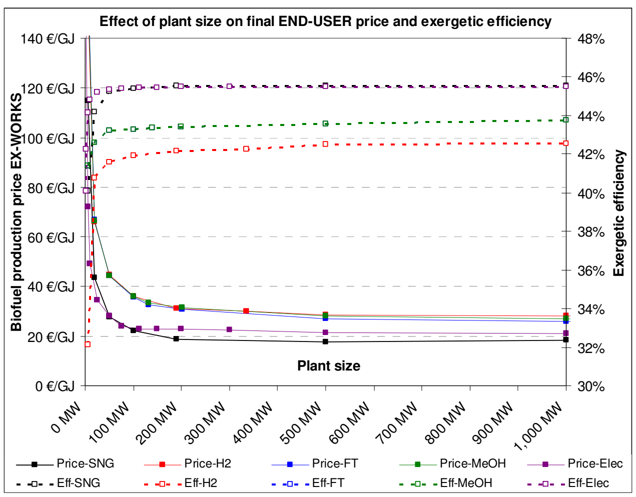


Figure 3: Final prices and exergy efficiency for wood-based biofuels.

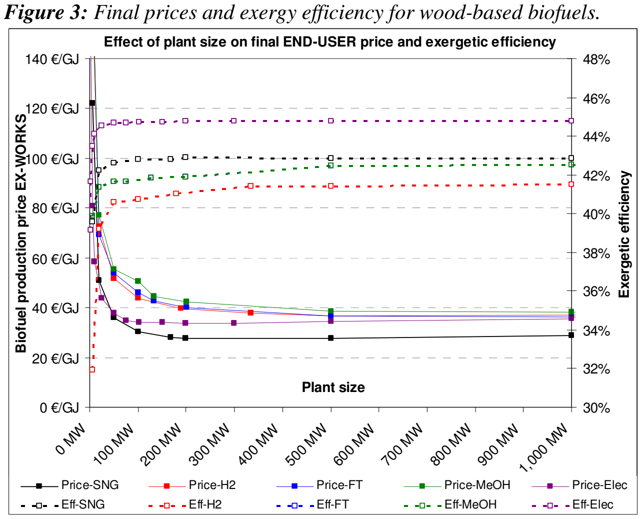


Figure 4: Final prices and exergy efficiency for straw-based biofuels.

When biomass availability is taken into account, it is observed that few countries have enough biomass to feed any biofuel plant at its optimal production scale. This assertion is especially sensitive for FT and MeOH production as both biofuels require large amount of biomass to run plant scales of 1000 MW_{fuel}. In effect only Spain, France, Italy, Austria, Germany, Poland, Sweden, Finland, UK, Hungary and Romania can feed either wood or straw-based biofuels plants using own biomass sources (see Figure 5 and 6). Alternatively, biomass could be also imported from nearby regions, as done in next section 4.2. However, in some cases, CO_2 emissions could exceed the threshold of fossil fuels.

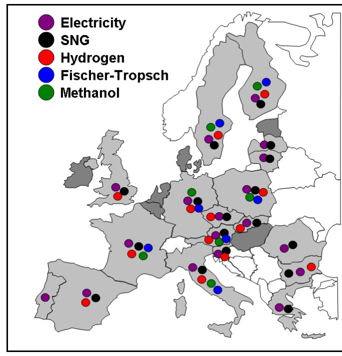


Figure 5. *The colors identify that at least 1 plant can be fed with 'national' forest residues.*

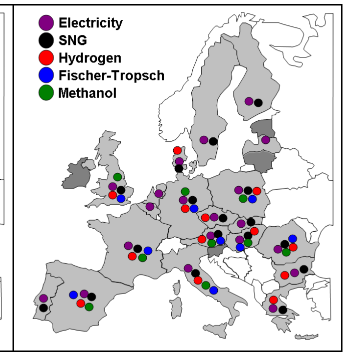


Figure 6. *The colors identify that at least 1 plant can be fed with 'national' straw residues.*

For the case that biomass is not imported, relative ecotax (i.e., €/ton CO_2) are found in the range presented in Table 2. This ecotax does not correspond to the Carbon taxes that are levied from companies exceeding the limits of CO_2 emissions. In fact, is a 'virtual' value that we have calculated in order to charge CO_2 emissions and equalize biofuels and fossil fuels prices. Hence, it should be added in top of production costs. Calculated ecotax values are notably high as biofuels and fossil fuels price difference is rather substantial. Bio-electricity production turns out to be an exception as, for some countries, bio-based price is cheaper than coal-based power price (i.e., all countries except Austria, Bulgary, Baltic states, Sweden, Finland and France). Concerning SNG production, Sweden is the only country where bio-based SNG is cheaper than fossil natural gas.

Table 2: *Ecotax ranges for wood and straw-based biofuels (€/ton CO_2).*

⇩Biofuel / Feedstock ⇨	Forest wastes	Straw
Electricity	0 to 53	0 to 32
SNG	51(*) to 208	71(*) to 373
FT-fuels	152 to 420	114 to 640
MeOH	198 to 675	215 to 900
H_2	9 to 181	8 to 233

(*) Ecotax for Sweden is 0 as SNG is cheaper than fossil gas.

On the other hand, the LCA analysis reveals that less CO_2 is emitted (in terms of kg-eq CO_2/GJ_{fuel}) during SNG production, followed by electricity and H_2 generation. Notably higher CO_2 emissions are released for FT and MeOH mainly because a larger amount of biomass needs to be transported and biofuel distribution is done by means of trucks, which consume fossil diesel. In countries where biomass availability density is (i.e., kton/km^2) is scarce, CO_2 emissions exceed those of the corresponding fossil fuel. In effect, wood-based MeOH and/or FT production in Sweden and Finland release more CO_2 than fossil diesel as biomass collection distances are considerable.

Same observation applies for straw-based MeOH and FT generation in Poland, Germany and Spain. This fact could be minimized by importing biomass from nearby regions, as suggested in section 4.2.

4.2 Biofuels introduction into the European Energy market

The model of Figure 2 is applied in this section to determine maximum bioenergy production when all available forest and straw residues are consumed, and following the scenarios of Table 1. However, values of each scenario correspond to the maximum theoretical production that could be achieved if all available biomass could be purchased. In effect, unlike solar and wind energy, biomass is normally owned by individuals or holdings, which ultimately decide its final application and price.

For an optimal utilization of biomass sources in electricity generation, two '*short-terms'* scenarios are analyzed. In the first case (i.e., **scenario I-A**), it is assumed that 10 wt% of the coal consumed in power plants is substituted by the corresponding amount of straw, and the remaining part together with forestry residues are used in potentially new BIGCC plants operating at the optimal scale of about ~100 MW_{el}. Electrical efficiency (η_{el}) of co-firing plants is negatively affected by 4% when introducing 10 wt% of straw. Efficiencies of new bio-based BIGCC stations attain lower values, i..e, 42% and 36% for wood and straw (see Figure 3 and 4). In the second case (**scenario 'I-B'**), biomass sources are fully consumed in potential new BIGCC plants. Hence, in this case, no co-firing is envisaged and energy efficiencies of coal plants are not affected (i.e., 45.2% [8]). Table 3 presents the share of renewable electricity generated in these 2 scenarios. In both tables, "*x*" and "*y*" represent the wood and straw fraction that is used in co-firing plants. By definition, in scenario '**I-B',** "*x*" and "*y*" fractions are both 0%. In all cases, final electricity outcome (renewable + fossil) should at least equal 3.77 EJ/yr, which accounts for the predicted coal-based electricity generation by 2020 [8].

Table 3: *Share of renewable energy (i.e., bio-electricity or biofuel production divided per total energy[(a)]) for the each scenarios of Table 1.*

Scenarios	Biomass in co-firing	% renewable energy in electricity	natural gas	road fuel	total
I-A	x=0,y=26%	**28.9 %**	0.0%	0.0%	3.0%
I-B	x=0,y=0	**34.3 %**	0.0%	0.0%	3.0%
II	x=0,y=26%	5.1%	0.4%	**8.1%**	3.4%
III	x=0,y=0	0.0%	0.3%	**9.5%**	3.3%
IV	x+y=5.7%	1.2 %	0.0%	**9.5%**	3.3%
V	x+y=1.9%	0.1 %	**6.3%**	0.0%	3.7%
VI	x=0,y=26%	5.1 %	**5.5%**	0.0%	3.7%
VII	x+y=4.6%	0.3 %	0.0%	**10.3%**	3.5%
VIII	x+y=6.3%	1.1 %	0.0%	**5.6%**	2.0%

(a) Total energy equals to 3.77 EJ/yr of coal-based electricity, 25.9EJ/yr of natural gas and 15.1 of fossil road fuels by 2020 in Europe.

According to results from Table 3, the share of renewable electricity produced from biomass (i.e., 34.3%) is the largest for the second scenario '**I-B',** in which no co-firing is planned. In effect, notably less coal is needed to fulfill total power outcome of 3.77 EJ/yr. Conversely direct co-firing (i.e., '**I-A**') has the limitation of 10% coal substitution by biomass. Combined with solar and wind energy, about 31% of the electricity production by 2020 could be renewable, i.e., 10 points higher than the target of Directive 2001/77/EC.

The third **scenario 'II'** refers to maximizing the introduction of biomass into the energy market with minor changes in the actual infrastructure. In this case, about 10 wt% of coal consumed in existing coal-fired power plants is replaced by the corresponding straw amount. The remaining biomass fraction is then used for new bio-based FT-plants which would operate at the optimal scale of 1000 MW_{fuel}. In **scenarios 'III'** and '**IV'** FT-fuels production is prioritized and biomass leftovers are used for either SNG or cofiring respectively. As expected, higher biofuel share is obtained in the III and IV (i.e., 9.5%) as more biomass is available for FT-production. The 9.5% biofuels share is slightly behind the 10% European target established in the Directive 2009/28/EC. Maximal SNG production, at the optimal scale of 200 MW_{SNG}, is obtained in **scenario "VI"** (i.e., SNG share of 6.3%). For this analysis, all available wood and straw residues are primarily converted into SNG plants, whereas the leftovers are used in co-firing stations. In the **scenario 'V'**, preferences are exchanged and, thus, less biomass is ready for SNG production (i.e., SNG share of 5.5%). If H_2 and MeOH (i.e., **scenario 'VII'** and **'VIII'**) are produced for road transport, the corresponding biofuels shares attain 10.3 and 5.6 % respectively. The share of H_2 is notably high due to the expected better efficiencies of fuel cell cars (FCV). However, it should be mentioned that the introduction of H_2 in the energy market is predicted for a long-term future. Therefore, is highly probable that biomass would be already consumed in other processes, thus, reducing the possibility of producing bio-H_2.

When comparing the share of renewable energy in the total energy (see last column in table 3), it is observed that MeOH, followed by maximal electricity (i.e., '**I-A'** and '**I-B'**) and FT-fuels production (i.e., '**II'** to '**IV'**) lead to the lowest "total renewable share", whereas SNG attain the highest value. However, when taking into account that a diesel-fuelled car has an efficiency of about 22%, and natural gas-fuelled power plants work at less than 60% efficiency, the conclusion is then different, as more "useful" renewable energy is produced in the '**I-A**' and '**I-B**' cases. On the other hand, a LCA analysis reveals that the maximal renewable share in scenario '**V**' does not correspond to the maximal CO_2 savings.

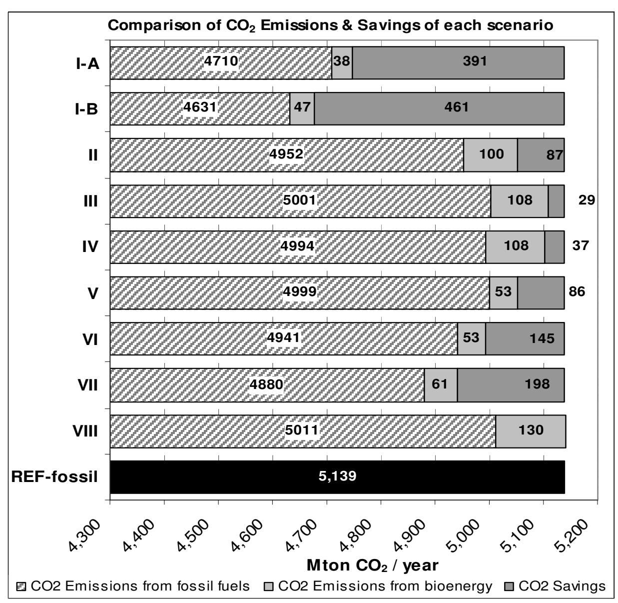


Figure 6: *CO_2 emissions (from fossil & biofuels) and savings.*

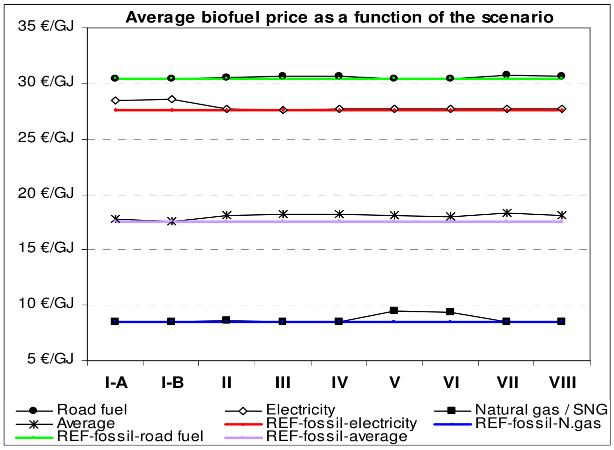


Figure 7*: Average fossil/biofuel prices for each scenario.*

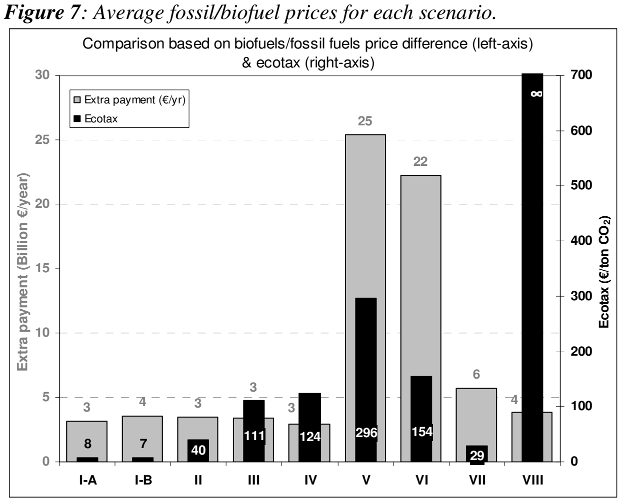


Figure 8*: Annual extra payment (grey bars) due to biofuels and fossil price differences. Calculated ecotax (€/ton CO_2) is given by black bars.*

In effect, as shown in Figure 6, bio-electricity production is a better alternative from an environmental point of view as ~ 361 to 491 Mton CO_2 are avoided annually. The '**V**' scenario is now the third best candidate with CO_2 savings of 145 Mton/yr. Average scenarios biofuels/fossil fuel prices are also compared in Figure 7, where biofuel prices are assumed to be free from taxation. Annual extra payment, which is due to difference in biofuels and fossil prices, is given by grey bars in Figure 8. As observed, average electricity price difference is notably higher in the "**I-B**" than in the '**I-A**' scenario. This extra costs account for 3-4 Billion € annually for the 24 European countries. In effect, bio-based electricity generation, especially when produced in new BIGCC plants, is more expensive than conventional coal power production. Black bars of Figure 8 represent the ecotax that would equalize fossil and biofuels prices (in €/ton CO_2). In this case, the "**I-B**" scenario implies the lowest ecotax (i.e., 7 €/ton CO_2), closely followed by the "**I-A**" case (i.e., 8 €/ton CO_2). H_2 would be the subsequent second best alternative although, as aforementioned, H_2-fuelled FCV will not be commercialized in the coming 10-20 years. Moreover, prioritizing co-firing over extra biofuels production is a better solution (i.e., "**II**" versus '**III**' and '**IV**' scenarios for FT, and "**VI**' versus '**V**" for SNG). Methanol (i.e., 'VIII') is completely disregarded as not only its price is high, but also it pollutes more than fossil diesel due to the low efficiency of MeOH-fuelled FCV.

5. CONCLUSION & DISCUSSION

Individual biofuels evaluation reveals that SNG and electricity yield the highest exergetic efficiencies when using wood as feedstock (i.e., ~ 45.5%). This statement is translated into the lowest biofuel prices (i.e., 19-27 and 23-34 €/GJ for SNG and electricity respectively in Austria) and required ecotax. However, when the analysis is extended to consume all available forest and straw residues in Europe for energy purposes, bioelectricity production turns out to be the best alternative from an economic and environmental point of view. In effect, about 391 to 461 Mton CO_2 are saved each year when biomass is used in either co-firing or new BIGCC plants (i.e., '**I-A**' scenario) or all biomass for new BIGCC plants (i.e., '**I-B**') respectively. The corresponding biofuels and fossil prices differences are also the lowest for both scenario, i.e., 3-4 Billion €/year respectively, with ecotax lying in the range of 7-8 €/ton CO_2. It is also observed that co-firing is preferred over extra biofuels production when the aim is to increase CO_2 savings.

On the other hand, if bioelectricity is summed to solar and wind energy, about 31% of the electricity production by 2020 could be renewable, i.e., 10% points higher than the target of Directive 2001/77/EC. In case of prioritizing FT-fuels production (i.e., scenarios '**III**' or '**IV**'), the share of biofuels in transport will be 9.5%, which is slightly below the 10% share target of Directive 2009/28/EC.

REFERENCES

1. Sues, A., M. Jurascik, and K.J. Ptasinski, *Exergetic Evaluation of 5 Biowastes-to-Biofuels Routes via Gasification* Energy, 2009. **Submitted for revision**.
2. Asikainen, A., et al., *Forest Energy Potential in Europe (EU27)*, in *Working Papers of the Finnish Forest Research Institute*. 2008, Finnish Forest Research Institute, Joensuu Research Unit: Helsinki (Finland).
3. EEA, *How much bioenergy can Europe produce without harming the environment?* 2006, European Environment Agency (EEA): Copenhagen (Denmark).
4. Siemons, R., et al., *Bio-Energy's role in the EU Energy Market. A view of developments until 2020*. 2004, BTG biomass technology group BV: Enschede (the Netherlands).
5. Nikolaou, A., M. Remrova, and I. Jeliazkov, *Biomass availability in Europe*, in *Lot 5: Bioenergy's role in the EU Energy Market*. 2003, Centre for Renewable Energy Sources (CRES). BTG Czech Republic s.r.o. ESD Bulgaria Ltd: Pikermi Attiki (Greece).
6. Wakker, A., et al., *Biofuel and bioenergy implementation scenarios: VIEWLS WP5, modelling studies* 2005, Energy Research Centre of the Netherlands (ECN). Chalmers University of Technology: Petten (the Netherlands).
7. de Wit, M. and A.P.C. Faaij, *Biomass Ressources Potential and Related Costs*, in *REFUEL Work Package 3*. 2008, Copernicus Institute. Utrecht University: Utrecht (the Netherlands).
8. Mantzos, L., et al., *European Energy and transport Trends to 2030 (Appendix2)*, in *Summary Energy Balances and Indicators*. 2003, European Comission. Directorate General for Energy and Transport: Luxembourg (Luxembourg).

Analysis on the impacts of the transmission line capacity expansion on clean power system development to realize clean energy policies

Jinxu Ding and Arun Somani [1]

Abstract— **As global warming has become an issue that all human beings have to address, it is time to figure out how to reform our power system so that we can reduce the CO_2 emission meanwhile stimulating the development of clean energy, such as wind energy. Because wind energy is not so dispatchable as fossil energy and it is always located in remote areas far away from high-power-demand areas, it is necessary to set up new transmission lines to transmit the clean power to these power-hungry areas. In this paper, we use a modeling framework (designed by ourselves [3]) to do sensitivity analysis on the transmission line capacity expansion. In our modeling framework, we use linear and mixed integer linear programming to model the fossil and wind power capacity expansion problems and also model the energy policy (e,g, RPS). Our results show that transmission line capacity expansion associated with wind power generation systems play a crucial role in the process of long-term clean power system development. It means that the capacity expansion of wind/fossil power generation system is very sensitive to the variation of the associated transmission line capacity expansion (TLCE). For example, a small increasement of TLCE percentage out of its associated wind/fossil generation systems (from 10% to 25% in North Dakota) can reduce requirements for the capacity expansion of wind/fossil power generation systems by 69.23% for wind power generation systems and by 84.33% for fossil power generation systems. The quantitative results from our model discover that the similar relationships can be found in other states that need to be planned with renewable energy. This discoverage can help decision makers, ISO, utility companies to do investment, management or development for renewable energy and to realize clean energy policies.**

Keywords: Clean power system, Energy policy, RPS (Renewable Portfolio Standard), Transmission line capacity expansion, Operation research, Sensitivity analysis, Green House Emission

I. Introduction

As a major component of energy industry, power generation is a primary source that produces CO_2 emission. It is right time to find out how to reduce the green house gas emission by replacing fossil power with clean power. In the initial phase of the replacement, it is necessary to design a strategy about how to stimulate clean power development in the regions with rich renewable sources, such as wind energy in the Midwest area of America. But, without building new transmission line systems across the area in America, it is impossible to transmit the wind power to other regions with the high power demand. Thus, transmission line capacity expansion becomes a bottleneck problem that must be solved if we want to stimulate clean power development. The similar problem can also be found in many countries that need to develop clean power system across a large region.

In order to understand the impacts of new transmission line capacity expansion on the clean power development and find solutions for the problem, we use a modeling tool designed by ourselves [3] to do sensitivity analysis on the parameter about transmission line capacity limitation on the whole clean power system development. The main function of the tool is to find solutions for strategy-level long-term investment planning problems of energy infrastructure. The tool is comprised of optimization models that minimize investment cost of fossil and clean power capacity expansion meanwhile meeting the power demand of the planned region. Because of the fluctuation of renewable energy, it is important to consider fossil energy as a backup to avoid possible blackout. In our tool, we also consider using an energy storage system (such as heat tank [4]) to store the surplus wind energy. The stored clean energy can be released later to maintain the balance between power demand and supply. The constraints in the optimization models include (1) we need to satisfy the power demand of the region; (2) we need to satisfy some standards about the requirement of clean power market-share percentage by some future years in the region. For example, RPS [1] is a state-level clean power regulation that requires the increased production of power from renewable energy sources, such as wind, solar, biomass, and geothermal. RPS policy generally imposes an obligation on electricity supply companies to produce a specified fraction of their electricity from renewable energy sources [1]. With the above optimization objective and constraints, we can apply some optimization techniques of operation research (e.g. linear programming) to the above clean power development problems.

Some similar works have been done in this area. A related work about operation research for energy modeling is WinDS (Wind Deployment Systems Model) [5] developed by SEAC (Strategic Energy Analysis Center) of NREL (National Renewable Eenergy Lab). It is a multiregional and multitime-period linear programming model embedded with the Geographic Information System (GIS). This model focuses on the market issues about transmission access and cost, and the fluctuation of wind power. There are also other energy planning models. For example, the All-Modular Industry Growth Assessment (AMIGA) model [6] is a comprehensive economic model of energy markets. None of the above related works focuses on the conventional/renewable energy infrastructutre reform and planning. Moreover, our model is comprised of linear or mixed integer linear programming models that can describe decision making problems in energy domain better than the above single-type optimization models. For example, weather a region should buy clean power from other regions or not buy can be better described as an integer variable. The major contributions of our paper include we use the modeling tool designed by ourselves to do the optimization computation and present the computation results, which discover the relationships between fossil/wind power capacity expansion and its associated transmission capacity expansion. We also do the sensitivity analysis for the parameter about transmission line capacity expansion (TLCE)

[1] Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011, U.S.A. Email: {jxding,arun}@iastate.edu. This work is funded by **NETSCORE-21** (EFRI-RESIN: 21st Century National Energy and Transportation Infrastructures: Balancing Sustainability, COsts, and REsiliency) research project Award #0835989 from NSF-EFRI (National Science Foundation - Office of Emerging Frontiers in Research and Innnovation)

and discover that the capacity expansion of wind/fossil power generation system is very sensitive to the variation of the associated transmission line capacity expansion (TLCE). For example, a small increasement of TLCE percentage out of its associated wind/fossil generation systems (from 10% to 25% in North Dakota) can reduce requirements for the capacity expansion of wind/fossil power generation systems by 69.23% for wind power generation systems and by 84.33% for fossil power generation systems. The quantitative results from our model discover that the similar relationships can be found in other states that need to be planned with renewable energy. This discoverage can help decision makers, ISO, utility companies to do investment, management or development for renewable energy and to realize clean energy policies.

II. Model

In this work, we do the clean power planning for five states (**ND**:**N**orth **D**akota, **SD**:**S**outh **D**akota, **NE**:**NE**braska, **MN**:**M**i**N**nesota, **IA**:**I**ow**A**) of the Midwest area in America because this whole region has rich wind energy [2]. We define each state as a sub-region, which has existing fossil power capacities and some potential wind energy. For each sub-region, we set up a HLM (hour level model) that is responsible for doing optimization computation at hour level because the balance between power demand and power supply must be planned at finer temporal resolution. We also design a YLM_i (year level model) that is responsible for doing optimization computation for fossil or wind power capacity expansion in region i because this kind of planning should occur at year level. Because power demands of continuous days in a sub-region do not have much difference, we choose m typical days from each month of every year between 2010 and 2049. For one year, there will be $M = 12 \times m$ typical days. This can also keep the computation volume of each sub-region in a reasonable range. In each typical day with 24 hours of region i, the HLM_i is computed for each hour. The computation results of HLM_i mainly include power supply from fossil power, wind power and also stored surplus power because of the fluctuation of wind energy. The $M \times 24$ typical day's computation results of HLM_i are accumulated together to form the year level results, which are used to find the percentage value of the wind power supply out of the total power supply in the sub-region. The clean power percentage value is used to compare with the specified percentage value in RPS_i policy of the region i. If it is less than the value indicated by RPS_i, the YLM_i is solved to find how to do the wind power capacity expansion for the current year in region i such that the RPS_i policy can be realized. Then, the HLM_i models are solved again with the new capacity expansion solved by YLM_i. If the RPS_i is still not realized in region i, the YLM_i model will be solved again for more clean power capacity expansion according to the distinct constraints of region i and then the HLM_i will be solved again to check whether RPS_i policy is realized. The HLM_i and YLM_i models will not be solved for the next year until RPS_i policy is realized for the current year in region i. In this way, our modeling framework [3] can help find out how to do fossil and wind power planning with finer temporal resolution for a long time period such as 40 years. In this section, we present the nomenclature defined and the models designed for the modeling framework.

A. Nomenclature

The following terms and symbols will be used in the design for HLM_i and YLM_i models in Section II-B and II-C.

(A1) Sets and Indices :

T	the set of hours in a day from 1 to 24
t	$t \in T$ for hour-level model, $t \in Y$ for year-level model
k	a future hour, $k \in T$
Y	the set of years from 2010 to 2049
y	the current year, $y \in Y$
z	a future year, $z \in Y$
S	the set of load sub-regions in Midwest area
i	a load subregion $i \in S$
j	a load subregion $j \in S$

(A2) Decision variables :

CE_{it}^{wp}	the capacity expansion of wind power plants of load sub-region i in period t [MW]
CE_{it}^{fp}	the capacity expansion of fossil power plants of load sub-region i in period t [MW]
CE_{it}^{wh}	the capacity expansion of heating storage of wind power plants of load sub-region i in period t [MW]
PS_{it}^{fp}	the power supply from fossil power plants of load sub-region i in period t [MW]
PS_{it}^{wp}	the power supply from wind power plants of load sub-region i in period t [MW]
PS_{it}^{wh}	the power supply from the energy storage of wind power plants of load sub-region i in period t [MW]
PB_{it}	the clean power bought from other load sub-regions (with stored surplus wind) to load region i in period t. In YLM, it is PB_{iy}. [MWh]
BB_{it}	the binary variable that indicates whether the load sub-region i needs to buy power from other regions [MWh]
R^{wh}	the percentage of stored power released from energy storage of wind power plants

(A3) Parameters about the cost of operation:

OC_{iy}^{wp}	the operation and management cost rate of a wind power plant of load sub-region i in year y [$/MW]
OC_{iy}^{fp}	the operation and management cost rate of a fossil power plant of load sub-region i in year y [$/MW]
OC_{iy}^{h}	the operation and management cost rate of heating storage of load sub-region i in year y [$/MW]

(A4) Parameters about the cost of investment and transmission:

IC_{iy}^{wp}	the investment cost of wind power plants of load sub-region i in year y [$/MW]
IC_{iy}^{fp}	the investment cost of fossil power plants of load sub-region i in year y [$/MW]
IC_{iy}^{h}	the investment cost of heating storage of load sub-region i in year y [$/MW]
TC_{iy}^{w}	the cost of transmission lines corresponding to wind power capacity expansion of load sub-region i in year y [$/MW]
TC_{iy}^{fp}	the cost of transmission lines corresponding to the fossil power capacity expansion of load sub-region i in year y [$/MW]

(A5) Objective function variables (part 1):

PV_{it}	the total price volatility caused by the difference between power supply and power demand in S of load sub-region i in period t

(A6) Objective function variables (part 2):

CO_{it}	the total CO_2 emission cost of load sub-region i in period t [$]
IC_{it}	the total investment cost of wind and fossil power plant capacity expansion of load sub-region i in in period t [$]
OC_{it}	the total operation and management cost of wind and fossil power plant of load sub-region i in period t [$]
TC_{it}	the total cost of transmission lines built up for transmitting wind power from power plants to its closest existing power grids of load sub-region i in in period t [$]
TF_{iy}	the transmission factor of load sub-region i in year y. It expresses the limitation of transmission systems on the wind power that can be transmitted online in real time. $TF_{iy} \in (0, 1]$

(A7) Parameters :

RPS_{it}	the percentage of clean power in the total power supply of load sub-region i in period t
DR^w_{it}	the discount rate of funding invested on wind energy development of load sub-region i in period t
DR^f_{it}	the discount rate of funding invested on fossil energy development of load sub-region i in period t
PD_{it}	the power demand of load sub-region i in period t [MWh]
CO^{fp}_{it}	the cost of CO_2 emission of fossil power plants of sub-region i in period t [$/MW]
PN_{it}	the power that can be bought from the neighbouring sub-regions of i in the period t
EC^{fp}_{iy}	the existing fossil power plant capacity of load sub-region i in year y [MW]
PO^{wp}_{it}	the output power generated by the wind turbines in wind power plants .It is the minimal value among the existing wind power plant capacity and the total available wind power of load sub-region i at time t
SP^{wh}_{it}	the surplus power stored in the energy storage system of wind power plants of load sub-region i in period t [MW]
EC^{wp}_{iy}	the existing wind power plant capacity of load sub-region i in year y [MW]
η^h_{iy}	the transformation effciency rate of heating storage of load sub-region i in period t.
TWP^A_{it}	the total potential wind power that can be captured by wind trubines of load sub-region i in period t [MW]. In year level model, it is TWP^A_{iy}.

B. YLM (Year Level Model)

The YLM_i (year level model) in [3] is responsible for doing optimization computation of wind power development in sub-region i at year level. It is shown in the linear programming model **1**. It mainly focuses on satisfying the clean power market share requirement of RPS [1] policy in sub-region i by doing fossil or wind power capacity expansion at year level. Its objective is to minimize the cost of investment for capacity expansion, transmission line cost, CO_2 emission and operation cost. In this model, the constraint (1b) is to guarantee the balance of power demand and supply in sub-region i at year level. The constraints of (1c,1d) require that the power supply from fossil/wind power source are upperbounded by the sum of existing fossil/wind power capacity and the capacity expansion of fossil/wind power in year y. The constraint of (1e) requires that the power supply from wind energy storage system is upperbounded by the product of its storage efficiency and the sum of existing capacity of the storage system and its capacity expansion in year y. The constraint (1f) means that the total existing and expanded wind capacity should not be more than the total potential wind power available in the year y. The constraint (1g) requires that the percentage of clean power out of total power supply must be greater than or equal to a specified fraction in RPS policy for sub-region i in year y. The constraint of (1h) means that the energy storage system capacity expansion is upperbounded by its associated wind power capacity expansion and the transmission line capacity. If the power supply from wind turbines cannot be transmitted online in real-time, it will be stored in the associated energy storage system. The constraint (1i) means that the power supply from wind power plants and storage systems is upperbounded by the total existing and expanded capacity of wind power plants and storage systems and their associated transmission line capacity. All equations and inequations of the model **1** should hold for $\forall i, j \in S, y \in Y$.

$$\mathbf{min}\ IC_{iy} + CO_{iy} + TC_{iy} + OC_{iy} \tag{1a}$$

$$\mathbf{s.t.}$$

$$PS^{fp}_{iy} + PS^{wp}_{iy} + PS^{wh}_{iy} + PB_{iy} = PD_{iy} \tag{1b}$$

$$PS^{fp}_{iy} \le EC^{fp}_{iy} + CE^{fp}_{iy} \tag{1c}$$

$$PS^{wp}_{iy} \le (EC^{wp}_{iy} + CE^{wp}_{iy}) \times TF_{iy} \tag{1d}$$

$$PS^{wh}_{iy} \le \eta^h_{iy} \times (EC^{wh}_{iy} + CE^{wh}_{iy}) \tag{1e}$$

$$CE^{wp}_{iy} \le (TWP^A_{iy} - EC^{wp}_{iy}) \tag{1f}$$

$$(PS^{wp}_{iy} + PS^{wh}_{iy} + PB_{iy}) \ge RPS_{iy} \times (PS^{fp}_{iy} + PS^{wp}_{iy} + PS^{wh}_{iy}) \tag{1g}$$

$$(EC^{wh}_{iy} + CE^{wh}_{iy}) \le (EC^{wp}_{iy} + CE^{wp}_{iy}) \times (1 - TF_{iy}) \tag{1h}$$

$$PS^{wp}_{iy} + PS^{wh}_{iy} \le (EC^{wp}_{iy} + CE^{wp}_{iy} + EC^{wh}_{iy} + CE^{wh}_{iy}) \times TF_{iy} \tag{1i}$$

$where:$

$$CO_{iy} = CO^{fp}_{iy} \times PS^{fp}_{iy} \tag{1j}$$

$$TC = TC^w_{iy} \times (CE^{wp}_{iy} + CE^{wh}_{iy}) \times TF_{iy} + TC^{fp}_{iy} \times CE^{fp}_{iy} \tag{1k}$$

$$IC = \frac{(IC^{fp}_{iy} \times CE^{fp}_{iy})}{DR^f_{iy}} + \frac{(IC^{wp}_{iy} \times CE^{wp}_{iy})}{DR^w_{iy}} + \frac{(IC^{wh}_{iy} \times CE^{wh}_{iy})}{DR^w_{iy}} \tag{1l}$$

$$OC = (OC^{fp}_{iy} \times (EC^{fp}_{iy} + CE^{fp}_{iy})) + (OC^{wp}_{iy} \times (EC^{wp}_{iy} + CE^{wp}_{iy})) + (OC^{wh}_{iy} \times (EC^{wh}_{iy} + CE^{wh}_{iy})) \tag{1m}$$

C. HLM (Hour Level Model)

The HLM model in [3] is to minimize the imbalance caused by the power demand variation and wind power fluctuation and the CO_2 emission caused by fossil power generation at hour level. The HLM is a mixed integer linear

name	state	basic case	case 1	case 2	case 3
TF	ND	0.1	0.15	0.2	0.25
	SD	0.1	0.15	0.2	0.25
	NE	0.3	0.35	0.4	0.45
	MN	0.2	0.25	0.3	0.35
	IA	0.1	0.15	0.2	0.25

TABLE I
SUMMARY OF THE TF VALUES IN THE FOUR CASES OF SENSITIVITY ANALYSIS

programming model because the sub-region i needs to decide whether it needs to buy power from other sub-regions in the case that its local power demand cannot be satisfied. The HLM is shown in Model **2**. The constraint of (2b) is to guarantee the balance of power demand and supply in sub-region i at hour level. The constraint of (2c) means that the power supply from storage systems is upperbounded by the stored wind power of last hour. It also depends on the storage system efficiency. The constraint (2d) means that the power that can be bought by sub-region i at hour t is upperbounded by the total power that can be provided by other sub-regions. The binary variable BB_{it} is used to determine whether sub-region i needs to buy power from other sub-regions in order to keep power balance at hour t. The constraint of (2f) means that the power supply from wind power depends on the minimal value among the existing wind power capacity and the total wind power that can be captured by wind turbines at hour t. It is also impacted by the transmission factor. The constraint of (2g) expresses the impact of transmission capacity limitation on the power supply from wind power. Here, we develop wind power with higher priority than fossil power. In the case that power demand cannot be satisfied by the total supply from bought power and local wind-power generation and storage systems, the fossil power will be used as a backup as long as the RPS policy is satisfied. However, because the wind energy fluctuates and it needs to be transmitted by new-built expensive transmission lines, the transmission capacity is not necessarily equal to the related wind power capacity expansion as long as RPS policy is satisfied. The impact of the transmission limitation will be analyzed in Section III. All equations and inequations of the hour-level model **2** should hold for $\forall i \in S, t \in T$.

$$\mathbf{min}\ PV_{it} + CO_{it} \tag{2a}$$

$$\mathbf{s.t.}$$

$$PS_{it}^{fp} + PS_{it}^{wp} + PS_{it}^{wh} + PB_{it} = PD_{it} \tag{2b}$$

$$PS_{it}^{wh} \leq \eta_{it}^{h} \times SP_{i(t-1)}^{wh} \times R_{it}^{wh} \tag{2c}$$

$$PB_{it} \leq BB_{it} \times PN_{it} \tag{2d}$$

$$PS_{it}^{fp} \leq EC_{it}^{fp} \tag{2e}$$

$$PS_{it}^{wp} \leq min\{EC_{it}^{wp}, TWP_{it}^{A}\} \times TF_{it} \tag{2f}$$

$$PS_{it}^{wp} + PS_{it}^{wh} \leq (EC_{it}^{wp} + EC_{it}^{wh}) \times TF_{it} \tag{2g}$$

$$\mathbf{where:}$$

$$PV_{it} = 1 - \frac{PS_{it}^{fp} + PS_{it}^{wp} + PS_{it}^{wh}}{PD_{it}} \tag{2h}$$

$$SP_{it}^{wh} = PO_{it}^{wp} - PS_{it}^{wp} \tag{2i}$$

$$PO_{it}^{wp} = min\{EC_{it}^{wp}, TWP_{it}^{A}\} \tag{2j}$$

$$CO_{it} = CO_{it}^{fp} \times PS_{it}^{fp} \tag{2k}$$

III. Sensitivity Analysis

In this section, we do a sensitivity analysis on the transmission factor TF to find out the importance of the transmission line limitation in the process of developing clean power systems. We analyze the impact of transmission capacity change on the fossil and wind power capacity expansion. The table I shows the values of TF in four cases that are discussed in Section III. The tables of (II-VI) present the computation results of wind, fossil power generation capacity expansion and the associated wind energy storage system capacity expansion from 2010 to 2049. From the results, we can see that the capacity expansions of wind/fossil power systems decrease as the values of TF (transmission factor) rise from the the basic case to the case 3. For example, in the basic case of Table II, the **GRPY** (the **G**eometric mean of the growth **R**ate of wind power generation system **P**er **y**ear) is 7.59% in ND from 2010 to 2049. In the case 1 of the same table, the value of **GRPY** becomes 6.58% as we increase TF from 0.1 to 0.15. The reduction of **GRPY** means that the requirement on the wind power generation capacity is reduced in ND during the 40 years because we increase the transmission line capacity associated with wind power generation system. If TF is 0.1, there will be at most 10% of the wind power (generated by the wind turbines), which can be transmitted online in real time. The left power can be stored in the associated energy storage systems. The higher value of TF, the more wind power generated by wind turbines can be transmitted online in real time. Because of the high investment cost of transmission line capacity expansion and the related high maintain cost, we use TF as a parameter to limit the transmission line capacity expansion as long as the RPS goals can be realized in a state. In Table II, we can see that the value of **WPCI** (the Wind Power Capacity Installed) by 2049 decreases from $13308MW$ in the basic case to $9214MW$ in the case 1 in ND. The similar decrease can also be found in the values of **FPCI** (Fossil Power Capacity Installed Capacity) and **WESI** (Wind Energy Storage system Installed Capacity) in Table II. Especially, the values of **WESI** also decrease because more wind power can be transmitted online in real time and we do not need to store them. But, they are not 0 because we still store part of surplus wind power that cannot be transmitted online in real time in order to deal with the wind energy fluctuation. Moreover, the values of **WPP** (Wind Power Percentage) increase as the values of TF rise from the basic case to the case 3 in Table II because more wind power gets online in real time. The $\mathbf{RPS_i}$ is the required clean power fraction of the state i by year y. The above discussions also apply to the tables of III, IV, V and VI.

We summarize the above discussions in the figures of (1-5) to show the impact of transmission factor (associated with wind power generation and storage system) perturbations on the fossil/wind power generation and storage system capacity accumulatively installed in the five states by 2049. From the results of the figures of (1-5), we can see that the transmission line capacity limitation have a significant effect on the capacity expansions of fossil/wind power generation and storage systems during the period from 2010 to 2049. When more wind power can be transmitted online in real time, the required capacity expansions of fossil/wind power generation and storage systems decrease. For example, in Figure 1, if we increase the transmission factor TF from 0.1 to 0.25, the accumulatively installed fossil power capacity in ND by 2049 can be reduced by 84.33% ($(34430 - 5393)/34430 = 84.33\%$). The related

TF	Year	WPCI [MW]	FPCI [MW]	WESI [MW]	WPP %	RPS_{ND} %
basic case 0.1	**2010**	767	5289	0	4.37	1.61
	2020	2342	5622	788	13	12.74
	2030	4902	6953	2068	18.84	18.42
	2040	7432	10988	6583	24.6	24.38
	2049	13308	34430	11803	30	30
	GRPY	7.59%	4.92%	27.18%	5.06	7.79
case 1 0.15	**2010**	767	5289	0	4.85	1.61
	2020	1703	5289	468	13.5	12.74
	2030	3367	5289	1300	19.17	18.42
	2040	5104	6370	4296	26.84	24.38
	2049	9214	16458	7680	30.58	30
	GRPY	6.58%	2.95%	25.78%	4.83	7.79
case 2 0.2	**2010**	767	5289	0	5.33	1.61
	2020	1391	5289	312	13.88	12.74
	2030	2535	5289	884	19.31	18.42
	2040	4974	6287	2103	27.22	24.38
	2049	7173	8721	5504	32.1	30
	GRPY	5.9%	1.29%	24.71%	4.71	7.79
case 3 0.25	**2010**	767	5289	0	5.92	1.61
	2020	1183	5289	208	14.21	12.74
	2030	2119	5289	676	20.4	18.42
	2040	3991	5289	1612	28.37	24.38
	2049	4095	5393	3071	33.2	30
	GRPY	4.39%	0.05%	22.86%	4.52	7.79

TABLE II
THE IMPACTS OF TRANSMISSION LINE CAPACITY LIMITATION ON THE FOSSIL AND WIND POWER CAPACITY EXPANSION IN ND

TF	Year	WPCI [MW]	FPCI [MW]	WESI [MW]	WPP %	RPS_{SD} %
basic case 0.1	**2010**	288	2962	0	1.92	1.61
	2020	1744	2962	728	13.15	12.74
	2030	3408	2962	1560	19.08	18.42
	2040	6112	4459	2912	24.79	24.39
	2049	7902	7143	7005	30	30
	GRPY	8.86%	2.28%	25.49%	7.3	7.79
case 1 0.15	**2010**	288	2962	0	2.88	1.61
	2020	1224	2962	468	13.5	12.74
	2030	2264	2962	988	19.74	18.42
	2040	3127	2993	2658	26.07	24.39
	2049	3127	2993	2658	32.52	30
	GRPY	6.31%	0.03%	22.41%	6.41	7.79
case 2 0.2	**2010**	288	2962	0	3.84	1.61
	2020	912	2962	312	13.91	12.74
	2030	1744	2962	728	20.35	18.42
	2040	2503	2993	2003	28.78	24.38
	2049	2503	2993	2003	35.69	30
	GRPY	5.7%	0.03%	21.52%	5.88	7.79
case 3 0.25	**2010**	288	2962	0	4.8	1.61
	2020	808	2962	260	14.61	12.74
	2030	1432	2962	572	21.33	18.42
	2040	2087	2993	1565	29.51	24.38
	2049	2087	2993	1565	37.26	30
	GRPY	5.21%	0.03%	20.76%	5.4	7.79

TABLE III
THE IMPACTS OF TRANSMISSION LINE CAPACITY LIMITATION ON THE FOSSIL AND WIND POWER CAPACITY EXPANSION IN SD

reduction on wind power capacity expansion by 2049 can be 69.23% ($(13308 - 4095)/13308 = 69.23\%$) and the related energy storage system installed-capacity can be reduced by 73.98% ($(11803 - 3071)/11803 = 73.98\%$). The similar reductions can also be found in the figures of (2-5).

IV. Conclusion

In this paper, we use the modeling tool developed by ourselves to present the computation results about the relationship between fossil/wind power capacity expansion and its associated transmission line capacity expansion of five states in Midwest of America from 2010 to 2049. Our quantitative results show that the transmission line capacity associated with wind power capacity expansion plays an

TF	Year	WPCI [MW]	FPCI [MW]	WESI [MW]	WPP %	RPS_{NE} %
basic case 0.3	**2010**	153	7552	0	1.19	0.66
	2020	777	7552	312	7.59	7.5
	2030	1817	7552	832	14.85	14.81
	2040	3481	7552	1664	23.42	22.62
	2049	5494	8051	2670	31.5	30
	GRPY	9.62%	0.16%	22.42%	8.76	10.28
case 1 0.35	**2010**	153	7552	0	1.39	0.66
	2020	777	7552	312	8.25	7.5
	2030	1713	7552	780	15.35	14.81
	2040	3065	7552	1456	23.83	22.62
	2049	4851	7802	2349	31.87	30
	GRPY	9.27%	0.08%	22.02%	8.36	10.28
case 2 0.4	**2010**	153	7552	0	1.67	0.66
	2020	673	7552	260	8.9	7.5
	2030	1505	7552	676	16.12	14.81
	2040	2753	7552	1300	24.52	22.62
	2049	4417	7552	2132	33.15	30
	GRPY	9.01%	0%	21.72%	7.96	10.28
case 3 0.45	**2010**	153	7552	0	2.58	0.66
	2020	673	7552	260	9.5	7.5
	2030	1401	7552	624	17.11	14.81
	2040	2545	7552	1196	25.7	22.62
	2049	4001	7552	1924	33.98	30
	GRPY	8.73%	0%	21.4%	6.83	10.28

TABLE IV
THE IMPACTS OF TRANSMISSION LINE CAPACITY LIMITATION ON THE FOSSIL AND WIND POWER CAPACITY EXPANSION IN NE

TF	Year	WPCI [MW]	FPCI [MW]	WESI [MW]	WPP %	RPS_{MN} %
basic case 0.2	**2010**	1805	13905	0	4	1.41
	2020	4612	16454	3610	16.9%	16.65
	2030	8623	20249	6482	26.05	26.04
	2040	11597	21303	7969	28.3	28.15
	2049	14824	22935	9582	30.16	30
	GRPY	5.55%	1.29%	26.48%	5.32	8.16
case 1 0.25	**2010**	1805	13905	0	4.8	1.41
	2020	3786	15120	2788	17.8	16.65
	2030	7129	17881	5046	26.26	26.04
	2040	9393	18092	6177	29	28.15
	2049	12010	18970	7487	30.8	30
	GRPY	4.98%	0.8%	26.11%	4.88	8.16
case 2 0.3	**2010**	1805	13905	0	5.7	1.41
	2020	4427	15186	1311	18.28	16.65
	2030	7472	16933	2833	27.55	26.04
	2040	9448	16933	3821	30.11	28.15
	2049	11625	17284	4910	31.3	30
	GRPY	4.89%	0.56%	24.35%	4.45	8.16
case 3 0.35	**2010**	1805	13905	0	6.31	1.41
	2020	4029	14836	1112	19.01	16.65
	2030	6701	16082	2448	28.89	26.04
	2040	8365	16082	3280	31.02	28.15
	2049	10255	16222	4225	31.92	30
	GRPY	4.56%	0.4%	23.87%	4.24	8.16

TABLE V
THE IMPACTS OF TRANSMISSION LINE CAPACITY LIMITATION ON THE FOSSIL AND WIND POWER CAPACITY EXPANSION IN MN

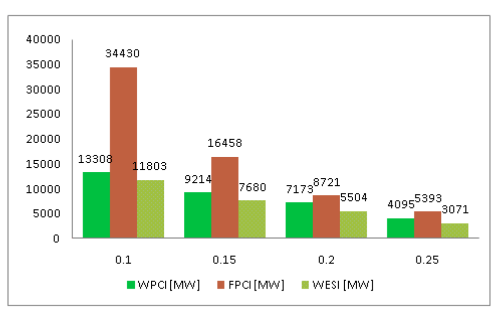


Fig. 1. The impact of transmission factor TF (associated with wind power generation and stroage system) perturbations (from 0.1 to 0.25) on the fossil, wind and storage system capacity accumulatively installed in ND by 2049

TF	Year	WPCI [MW]	FPCI [MW]	WESI [MW]	WPP %	RPS_{IA} %
basic case 0.1	**2010**	3053	14212	0	5.11	0.66
	2020	3714	14248	3343	10.86	7.5
	2030	6234	15936	5477	14.99	14.81
	2040	11345	20007	10032	22.77	22.62
	2049	17599	25326	15661	30.05	30
	GRPY	4.59%	1.49%	28.1%	4.65	10.28
case 1 0.15	**2010**	3053	14212	0	6.6	0.66
	2020	3410	14256	178	11.8	7.5
	2030	4681	14384	3979	15.55	14.81
	2040	7672	16126	6486	23.72	22.62
	2049	12169	19584	9998	31.22	30
	GRPY	3.61%	0.83%	26.64%	4.06	10.28
case 2 0.2	**2010**	3053	14212	0	7.2	0.66
	2020	3053	14212	0	12.02	7.5
	2030	4970	14256	958	16.11	14.81
	2040	7469	14988	5975	24.35	22.62
	2049	9368	15846	7259	32.13	30
	GRPY	2.92%	0.28%	25.6%	3.91	10.28
case 3 0.25	**2010**	3053	14212	0	7.98	0.66
	2020	3053	14212	0	12.78	7.5
	2030	4301	14212	624	16.88	14.81
	2040	6832	14662	1890	25.26	22.62
	2049	8174	15548	5060	33.2	30
	GRPY	2.56%	0.23%	24.45%	3.72	10.28

TABLE VI
THE IMPACTS OF TRANSMISSION LINE CAPACITY LIMITATION ON THE FOSSIL AND WIND POWER CAPACITY EXPANSION IN IA

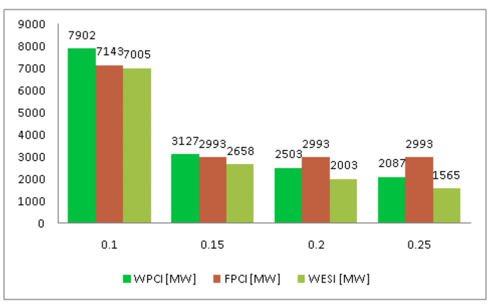


Fig. 2. The impact of transmission factor TF (associated with wind power generation and stroage system) perturbations (from 0.1 to 0.25) on the fossil, wind and storage system capacity accumulatively installed in SD by 2049

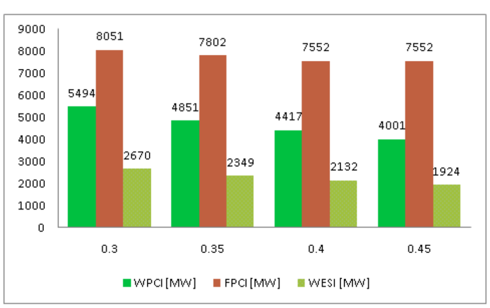


Fig. 3. The impact of transmission factor TF (associated with wind power generation and stroage system) perturbations (from 0.3 to 0.45) on the fossil, wind and storage system capacity accumulatively installed in NE by 2049

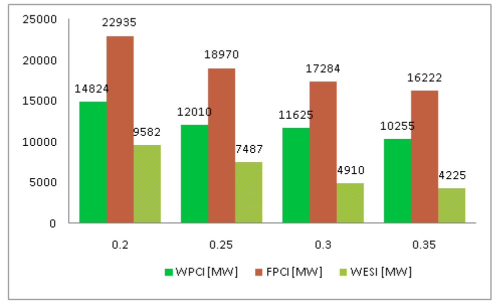


Fig. 4. The impact of transmission factor (associated with wind power generation and stroage system) perturbations (from 0.2 to 0.35) on the fossil, wind and storage system capacity accumulatively installed in MN by 2049

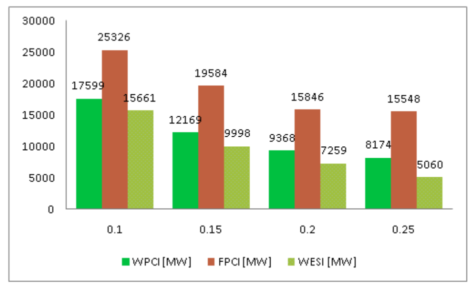


Fig. 5. The impact of transmission factor TF (associated with wind power generation and stroage system) perturbations (from 0.1 to 0.25) on the fossil, wind and storage system capacity accumulatively installed in IA by 2049

important role in the process of development of fossil and wind power systems. We do a sensitivity analysis for the transmission factor TF, which describes the transmission line capacity limitations on the wind power systems. The quantitative results show that a small perturbation on TF can have huge impacts on the development of fossil, wind and energy storage systems. The results in our paper have shown that the decision makers, wind power generators, ISO, utilities and investors should consider the economic benefits (reducing investment on fossil and wind power generation systems) and environmental benefits (reducing CO_2 emission) of building new transmission lines in remote areas with rich wind energy. The primary contribution of this paper is to use optimization techniques of operation research to do sensitivity analysis on the impacts of the transmission line capacity limitation on clean power system development in order to realize clean power policy RPS.

References

[1] Renewables portfolio standard http://www.awea.org/policy/rpsbrief.html

[2] Wind power map of the United States http://rredc.nrel.gov/wind/pubs/atlas/maps/chap2/2-01m.html

[3] Jinxu Ding and Arun Somani, "A long-term investment planning model for mixed energy infrastructure integrated with renewable energy", *the 2nd Annual IEEE Green Technologies Conference,* Fort Worth, Texas, April 15-16, 2010

[4] Thermal energy storage, http://www.nrel.gov/csp/troughnet/thermal_energy_storage.html

[5] Background of WinDS Model http://www.nrel.gov/analysis/winds/background.html

[6] The AMIGA Modeling System, http://amiga.dis.anl.gov/Tutorial1.pdf

Risk Assessment for Transboundary Spreading of Imbalances in Relation to Reduced Generation Reserves

Arturas KLEMENTAVIČIUS, Virginijus RADZIUKYNAS, Neringa RADZIUKYNIENE
Lithuanian energy institute
Breslaujos 3, LT-44403 Kaunas
LITHUANIA
arklemen@mail.lei.lt; virginijus@mail.lei.lt; rutkauskaite@mail.lei.lt

ABSTRACT
Paper presents a new balancing approach presuming curtailment of normal-as-usual generation reserves (secondary and tertiary reserves). The curtailment diminishes the robustness of a control area to resolve the eventual contingencies (emergencies), i.e. increases the risk of transboundary transfer of imbalance to adjacent systems. Nevertheless, if risk increment is small, it might be rather a rational clue to save balancing cost in proportion to the curtailed schedule of reserves.
The risk assessment procedure is presented based on building the contingency (state, imbalance) trees for major normal states of control area and evaluation of probabilities of transboundary imbalances. Therein the critical risk value is necessary to be established, as it will play a role of point of reference. The procedure is also based on building the risk function of reduced reserves. The function sets interdependence of reserve sizes and risk values.
The idea of a new balancing approach presumes the replacement of curtailed secondary and tertiary reserves by tertiary and emergency reserves, respectively.
Paper presents analysis of two hypothetical contingency trees, with gradual curtailment of generation reserves in the same normal state of a control area. Also two hypothetical cases of reserve replacement are imitated to demonstrate the applicability of an approach.

KEY WORDS
Imbalance, reserves, contingency, risk, probability, and control area

1. Introduction

The balancing problem of a transmission system operator (TSO) is to maintain the balance of demand and supply in its control area. The balance is being kept when the forehand schedule of transboundary load flows (interchanges with other control areas) is observed. Thus the TSO shall cope with area control errors (ACE), i.e. with momentary deviations from this schedule (imbalances) [1]. To cope with ACE means to clear (settle, eliminate) the imbalances: to localize them within the control area thus preventing their spreading beyond the control area to adjacent and further control areas in the synchronous zone.

The negative imbalances (shortages in power supply) mostly can be cleared with increase of fast-response generation, either own (in the control area in question) or from neighbouring areas. The TSO must schedule (book) the fast-response generating capacities in adequate amounts: overcapacities will cause overpay in favour of capacity providers, while undercapacities (lesser capacities) will be insufficient to clear certain probable imbalances, with resulting inadvertent and unacceptable spreading to other control areas.

The fast-response generating capacities are generation reserves. For the purpose of this paper, the reserves herein are classified into 3 major categories: secondary reserves (minute reserves), tertiary reserves (minute and hour reserves) and emergency reserves (of various timescales). In reality, there are a number of different names of reserve categories [2].

The balancing herein covers three reserve categories: balance reserves, secondary and tertiary, and emergency reserves. In fact, the balancing and emergency control are closely associated since a prolonged and progressing imbalance, unless cleared, might worsen an aggravated state and lead to emergency.

The TSO balancing cost depends mainly on reserve capacity (regulation-up capacity) and balancing energy (if the imbalances occur and this capacity is activated for energy production). Whereas the size of reserve capacity is scheduled (booked), the less is a schedule, the less is a cost. Irrespective of whether the TSO pays to reserve providers or provides those reserves itself, the balancing cost finally is born by customers. Therefore the general interest of customers is to bring down the balancing cost. Usually, the regulators and competition authorities monitor the balancing activity of TSO.

By now the reserves were sized pursuant to various considerations, and *n-1* condition was prevailing. Several balancing mechanisms based on minimum price bids in balancing/reserve capacity market were proposed and applied. Here two recent approaches are noteworthy: 1) “timescale” integration of balancing market and intraday market [3] and 2) “spatial” integration of balancing market as:

- pooling of reserves;
- sharing of reserves;
- shift from area control to regional control [4].

Nonetheless, these approaches do not presume the revision of reserve sizes in an individual control area: sizes seem to be the same as before integration.

By now the reserves also have been usually sized without risk assessment for the transboundary spreading of imbalances. It means that the relationship between risk value and reserve size was not analysed. In fact, the reserve curtailment shall negatively affect the capability of control area to clear the imbalance and to prevent the resulting spreading of imbalance (or emergency) to adjacent control areas. Such a negative effect seems to be larger if *n-k* contingencies are included into consideration (increased number of contingencies shall enlarge both number and size of imbalances).

The paper presents idea to curtail the secondary and/or tertiary reserves and, maybe, emergency reserves deliberately, for cost reasons, to trade off against increased risk of non-cleared imbalances (and transboundary spreading). Such risks should be assessed for normal and reduced sizes of reserves. Here the maximum allowable risk should be set as an operational constraint for a specific control area. The curtailed reserves could be replaced by reserves of other categories.

The following definitions and acronyms are used in the paper:

curtailed reserves – reserve capacities excluded from initial (normal) reserve capacities

reduced reserves – available reserves capacities after curtailment of reserves

EC – emergency control
ER – emergency reserve
SC – secondary control
SR – secondary reserve
TC – tertiary control
TR – tertiary reserve
TSO – transmission system operator.

2. Risk Assessment Procedure

Imbalance occurs as a result of contingency, i.e. an inadvertent triggering event. Emergency herein is qualified as a severe contingency.

Contingency impairs a normal state of control area and moves it to a disturbed state, say, alerted (with lesser imbalance) or aggravated (with larger imbalance) state. If not counteracted and compensated for within a control area, the imbalance goes across the borders of the area and affects other power systems. It is assumed that TSO performs the counteraction and compensation measures by activation of scheduled balance and emergency reserves.

The risk of transboundary spreading of imbalances depends on 3 factors:

- normal state **X0** of control area,
- eventual relevant contingencies,
- capability of reserves to counteract contingencies and compensate for imbalances.

Such a risk might be conceived as a probability that a control area in a normal state **X0** will cause a transboundary spreading of imbalance. The assessment of such a risk is not an easy thing. The eventual relevant contingencies should be identified, filtering out those with fewer eventualities and less relevance.

It is a matter of willingness and convention to set the interval of acceptable risk with regard to technical, economical considerations, system stability and other considerations. The upper limit of the interval shall correspond to the maximum allowable risk. The lower limit could correspond to minimum significant risk.

Identification of contingencies is a focal point in risk assessment procedure. It needs the building of contingency tree as illustrated below.

2.1 Contingency Tree in Case of Normal-as-Usual Reserves

As TSO schedules (books) reserves for a next day cycle, the normal state of a control area is conceived as a day-cycle operation plan. Such a normal state **X0(i,j,k)** is defined by

- generation schedule (layout and pattern) **i**,
- network topology **j**,
- load schedule (layout and pattern) **k**.

If the schedule **r** of balance and emergency reserves SR, TR and ER is included into an normal state, it might be denoted as **X0(i,j,k,r)**. Once such a state is defined, its contingency tree shall be worked out. Concomitantly, this state itself is a root of a tree (see Fig.1).

The tree consists of vertices representing specific contingencies (states), and branches standing for aggravating transitions from one state (contingency) to another state (with additional contingency). Contingency can be a multiple event (for example, consisting of two consecutive or simultaneous outages).

Since a wide variety of specific contingencies and their chains might be forecasted for one normal state with parameters **i,j,k,r**, the TSO shall identify only relevant contingencies. To some extent, such identification is a matter of art and heuristics. But this informal component of identification shall be subsequently combined with formal component – probability calculation (or evaluation) for each relevant contingency.

The illustrative example of tree building is presented in Fig.1.

TSO starts with identification of *n-1* contingencies **X1, X2, X3** which come out of normal state **X0(i1,j1,k1,r1)**. Vertices **X1, X2, X3** represent not only contingencies, but also the respective disturbed states of the control area and even momentary imbalances of those states.

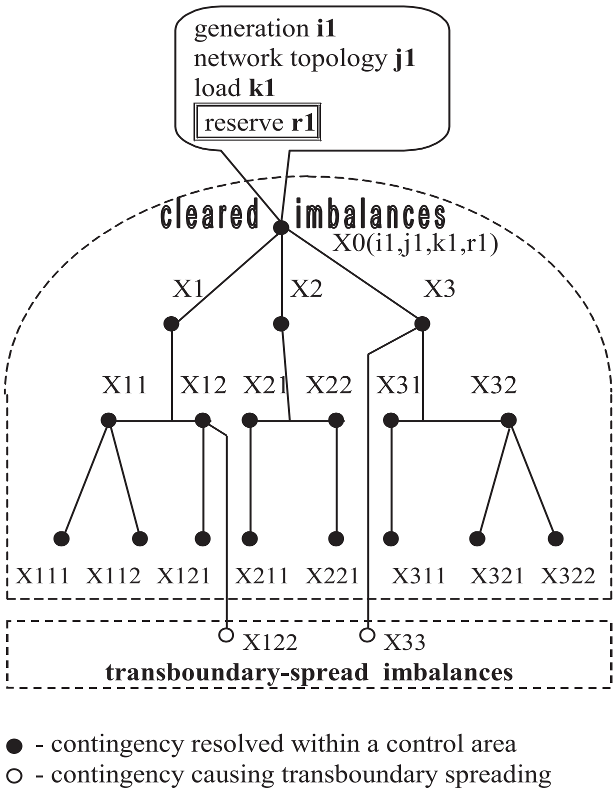


Fig.1. Illustrative Contingency Tree for Normal State **X0(i1,j1,k1)** of a Control Area with Normal-as-Usual Reserves **r1**

Further the TSO assigns (evaluates) probability of occurrence for each vertex (contingency, state, imbalance), e.g. p(**X1**), and imitates the counteraction measures against contingency by means of reserves **r1**. Herein the **r1** stands for normal-as-usual sizes of reserves, i.e. the capacities of available SR, TR and ER are in pursuance with conventional sizing rules and suffice to counteract rather large disturbances. In case of success the momentary imbalance is cleared (compensated for) and transboundary transfer will not have taken place. This result is inscribed into a vertex by a filled ball (e.g., **X1**).

If appropriate, the TSO proceeds with the same procedure for *n-2* contingencies. As in illustrative case, it identifies the contingencies **X11** and **X12,** which can occur in the disturbed state **X1,** and others (**X21, X22, X31, X32** and **X33**). The new disturbed state, say **X11**, contains both contingencies – preceding **X1** (still actual) and new one **X11**.

As depicted in Fig.1, TSO figures out that there is one contingency **X33** that will not have been resolved by reserves **r1**. It will cause transboundary spreading of imbalance. A blank ball of vertex X33 notes such result.

If appropriate, the TSO continues the same procedure for *n-3* contingencies. It finds out one more contingency **X122** that will not have been resolved by reserves **r1**.

To separate the resolved and unresolved contingencies, two zones are overlaid on the contingency tree:

1) CLEARED IMBALANCES;

2) TRANSBOUNDARY-SPREAD IMBALANCES.

Consequently a tree is configured so as to place the vertices of resolved contingencies into upper zone, and unresolved ones – into lower zone (Fig.1).

The risk **R** of transboundary spreading of imbalances for normal state **X0(i1,j1,k1,r1)** is calculated as:

$$\mathbf{R}_{(\mathbf{i1,j1,k1,r1})} = \mathbf{p(X122)+p(X33)}, \qquad (1)$$

where **p(X122)** and **p(X33)** are probabilities of contingencies (disturbed states, imbalances) **p(X122)**, **p(X33)**, respectively.

2.2 Contingency Tree in Case of Reduced Reserves

Further the TSO imitates curtailment of normal-as-usual reserves **r1** and examines capabilities of the reduced reserves **r2**, **r3**,..,**r7** to resolve the same contingencies for the same normal state **X0(i1,j1,k1)**. Illustratively, it was found that:

– curtailment by 10% will not diminish the capability of reduced reserves **r2** (90%): no new one contingency unresolved;

– curtailment by 20%, similarly, will not diminish the capability of reduced reserves **r3** (80%);

– curtailment by 30% will preclude the reduced reserves **r4** (70%) from resolving the contingency **X121**;

– curtailment by 40% (**r5=**60%) will not bring to a disability to resolve any new contingency ;

– curtailment by 50% will preclude the reduced reserves **r6** (50%) from resolving the contingency **X12**;

– curtailment by 60% will preclude the reduced reserves **r7** (40%) from resolving the contingencies **X112**, **X211**, **X322**.

This deterioration of reserve capabilities from **r1** (100%) to **r7** (40%) is presented in Fig. 2.

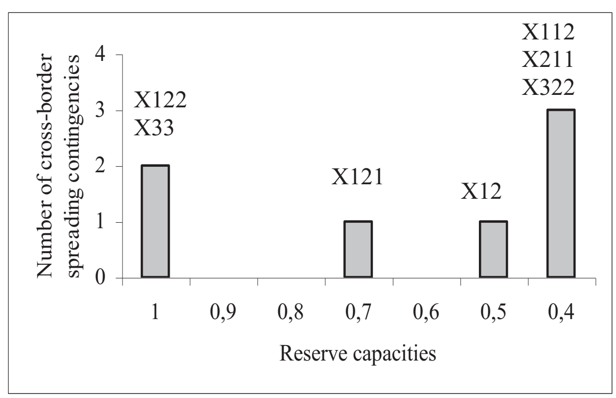


Fig.2. Dependence of Unresolved Contingencies on the Size of Reserved Capacities for Contingency Tree in Fig.1

Thus the TSO gets to know that the robustness of its control area is equal to the risk value $\mathbf{R}_{(\mathbf{i1,j1,k1,r1})}$ when normal-as-usual reserves **r1**, i.e. 100 % of dedicated reserves, are available.

As demonstrated, the curtailment of reserves from 100% to 40% will lead to 5 new unresolved contingencies and will reshape the zones of cleared imbalances and transboundary-spread imbalances. This reshaping is presented in Fig.3.

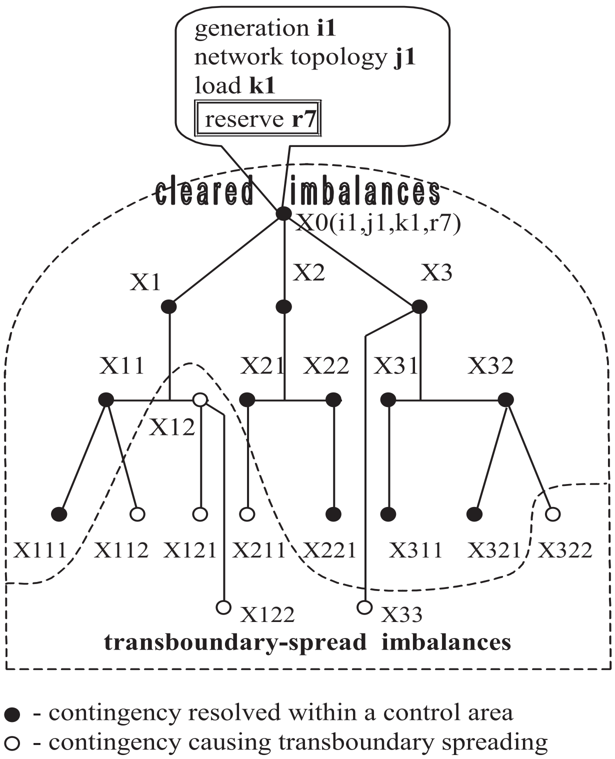


Fig.3. Illustrative Contingency Tree for Normal State **X0(i1,j1,k1)** of a Control Area with Reduced Reserves **r7** (40% of Normal-as-Usual Reserves)

As seen from Fig.3 and Fig.1, the reshaping is significant: the border between zones has shifted significantly inwards the zone of cleared imbalances. The increased risk **R** of transboundary spreading of imbalances for normal state **X0(i1,j1,k1,r7)** will be:

$$R_{(i1,j1,k1,r7)} = p(X122) + p(X33) + p(X121) + p(X12) + p(X112) + p(X11) + p(X322). \qquad (2)$$

2.3. Cohesion of Reserve Size and Risk

As presented in Section 2.2, the TSO gets from **X0(i1,j1,k1,r1)** to normal state **X0(i1,j1,k1,r7)** gradually, through successive reductions of **r1** to **r2**, **r3**,…, **r7**. Using expressions analogous to (1) or (2), operator finds the respective risk values for those intermediate reserve sizes **r2**,…,**r6**. Thus operator has sufficient data to draw up a function, which links reserves and risks (Table 1).
Nevertheless, risk function and entire risk assessment would get practical applications if risk function were intersected by upper limit of acceptable risk. Such a limit would be a critical value for presuming the trade-offs between reserving cost and risks. In effect, without having critical risk as point of reference, operator can hardly make a decision whether reduced reserve, say, **r7** is acceptable.

Table 1. Illustrative Risk Function of Reduced Reserves

Size of reserve	Risk of transboundary spreading of imbalances	Comments on risk value
r1=100%	$R_{(i1,j1,k1,r1)}$	risk might be considered insignificant
r2=90%	$R_{(i1,j1,k1,r2)}$	ibidem
r3=80%	$R_{(i1,j1,k1,r3)}$	ibidem
r4=70%	$R_{(i1,j1,k1,r4)}$	minimum significant risk
r5=60%	$R_{(i1,j1,k1,r5)}$	ibidem
r6=50%	$R_{(i1,j1,k1,r6)}$	critical risk
r7=40%	$R_{(i1,j1,k1,r7)}$	unacceptable risk

Exemplarily, for the risk function case in Table 1, the range of acceptable risk is assumed to extend from $R_{(i1,j1,k1,r4)}$ to critical risk $R_{(i1,j1,k1,r6)}$. However, $R_{(i1,j1,k1,r7)}$ exceeds critical risk and corresponds to insufficient robustness of control area.
Ideally, all normal states of control area {**X0(i1,j1,k1)**, **X0(i2,j2,k2)**, …} shall be imposed the same critical risk value.

3. Replacement of Curtailed Reserves

The idea of risk deterioration in exchange of curtailed reserves, with control of risk value against critical risk, presumes the backup for reduced reserves and replacement of the curtailed reserves by other reserve categories. Herein there are two major types of replacement: 1) TR replaces the curtailed SR; 2) ER replaces the curtailed SR and TR. Both types are considered in sections 3.1 and 3.2, respectively.

3.1. Replacement of Curtailed Secondary Reserve

The rationale of this replacement is to curtail the SR schedule, e.g. for a next day cycle, in hope to replace the curtailed capacities by TR which is scheduled in normal-as-usual size. Nevertheless, if the tertiary reserve must fulfill its function in tertiary control and cannot back up the reduced SR, it can be done by emergency reserve. The functioning of reserves and aggravation of a control area's state is illustrated by hypothetical case in Fig.4. Therein the situations of control area are laid out clockwise in 6 graphs, in aggravating order.
Situation 1 corresponds to normal balancing practice. Both secondary and tertiary reserves are scheduled by the TSO in normal-as-usual sizes.
Situation 2 depicts transition to a new balancing approach. TSO curtails the SR schedule and books only minor part of it, while TR is scheduled as previously, but divided into two portions. The larger one is equal to the curtailed portion of SR and is dedicated for its eventual

replacement. Thus the TSO is saving a balancing cost incurred by curtailed SR.

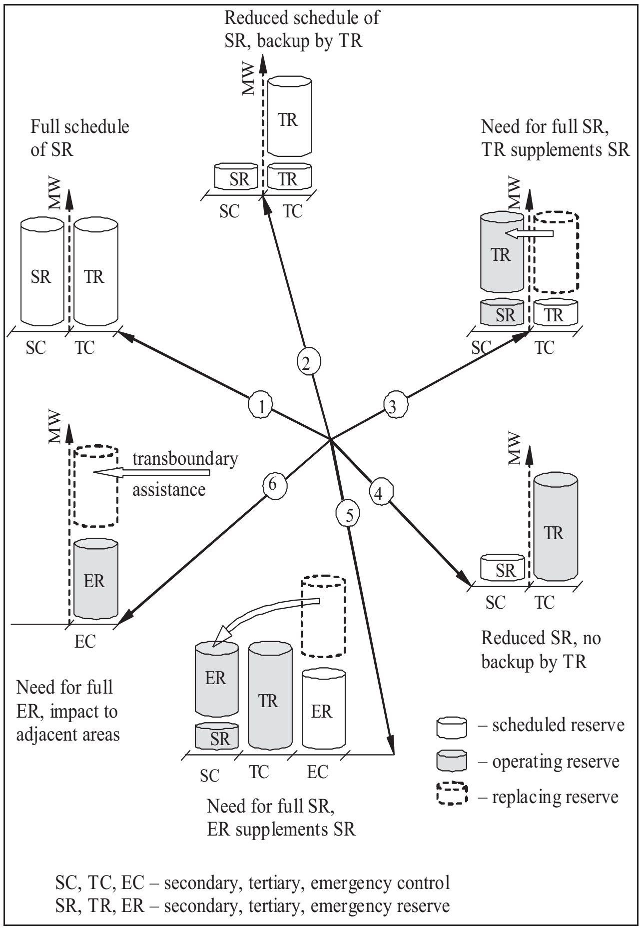


Fig.4. Operation of reserves for replacement of curtailed secondary reserve

Situation 3 reflects a occurrence of contingency, for instance significant load deviation equal to capacity of normal-as-usual SR. Loosely referring to contingency tree in Fig.3, it might be contingency **X2**. The momentary imbalance is cleared collectively by the reduced SR and larger portion of TR. Hence TR is providing assistance to secondary control.

Situation 4 represents an end of secondary control cycle and beginning of tertiary control cycle. It is assumed that since cause of contingency **X2** persists, the entire imbalance must be taken over by TR. Thus TR takes upon itself a new part of imbalance previously cleared by reduced SR in situation 3. The reduced SR is released from counteraction to **X2,** so as to be prepared for eventual new contingency. Nevertheless, there will be no backup from TR for the released SR. The situation can be considered as alerted.

Situation 5 introduces additional contingency, say, **X21** (Fig. 3). It might be a generation outage also as large as capacity of normal-as-usual SR. As in preceding situation 3, the reduced SR is insufficient to clear the imbalance. The backup is provided by ER in amount of curtailed SR. The imbalance is cleared collectively by reduced SR and a portion of ER. It decreases the preparedness of ER to compensate for an eventual emergency. The situation is aggravated.

Situation 6 represents a final and worst aggravation of situation 5. Following a strong contingency **X211**, the normal-as-usual emergency reserve is needed, however only part of it is available. The contingency **X211** cannot be resolved due to momentary deficiency of ER, and, consequently, the imbalance (emergency) will spread to adjacent control areas. It will be make those areas to share the clearance of imbalance, i.e. to provide a transboundary assistance to the control area in question. Maybe, the area could resolve this contingency by undertaking intelligent load shedding. Anyway, the situation 6 is a serious violation of parallel operation rules.

3.2. Replacement of Curtailed Secondary and Tertiary Reserves

The rationale of this replacement is to curtail the TR schedule, additionally to curtailed SR schedule (as in section 3.1), for larger savings of balancing cost. Thereby TSO hopes to provide concurrent backup, when required, for reduced SR&TR from normal-as-usual ER. The functioning of reserves and aggravation of a control area's state is illustrated by hypothetical case in Fig.5 where the aggravation process is split into in 4 graphs (clockwise).

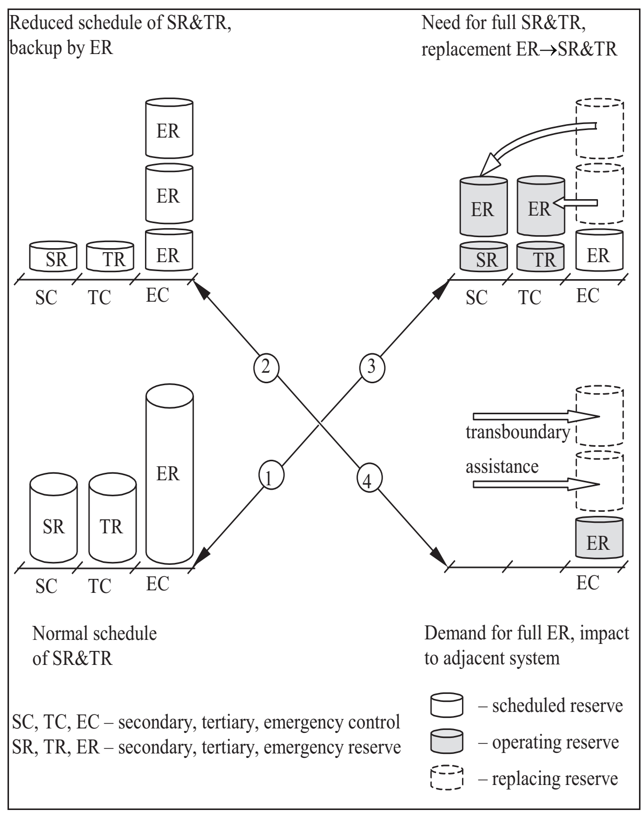


Fig.5. Operation of reserves for replacement of curtailed both secondary and tertiary reserves

Situation 1 is the same as situation 1 in Fig.4.

In situation 2, the TSO addresses to a new balancing approach and curtails both SR and TR schedules. To pass through the eventual contingencies, which would require those missing portions of SR and TR, TSO divides the capacity of ER into 3 portions. 2 of them are equal to the curtailed portions and are dedicated for eventual replacements. The 3rd portion is left for emergency control.

The situation 3 represents occurrence of two contingencies in succession, both as large as the curtailed reserves. First one, contingency **X2** (Fig.3), is resolved by one portion of ER in secondary control cycle and in the following tertiary control cycle (collectively with curtailed SR and curtailed TR, respectively). Second contingency **X21** needs the second portion of ER in order to be resolved in a new secondary control cycle. Thus both portions of ER enter into domains of secondary and tertiary control. In result, both imbalances **X2** and **X21** are cleared, but the robustness of control area relies further on the third portion of ER. The situation is vulnerable.

Situation 4 represents the worst development of preceding situation 3. Control area experiences a strong contingency **X211**, which could be resolved only with normal-as-usual ER, but, unfortunately, it is not available. Such deficiency of ER leads to transboundary imbalance, even larger as in previous hypothetical case (Fig.4).

4. Conclusions

1. Risk assessment for major normal states of a control area (power systems), as proposed in the paper, yields the risk values (probabilities) for eventual events of imbalance transfer to adjacent areas (power systems). Here the *n-k* type events (contingencies) shall be included into consideration.

2. The critical risk value (maximum allowable) should be laid down for a control area in order to distinguish between acceptable and unacceptable risk. Reliance on such a value allows the transmission system operator for some curtailments in schedules of normal-as-usual generation reserves. It brings savings in balancing cost in exchange to the increased but still acceptable risk of transboundary spreading of imbalances.

3. Once a new balancing approach as defined in conclusions 1 and 2 is undertaken, the tertiary and emergency reserves should be shared for secondary and tertiary control to replace the curtailed secondary and tertiary reserves, respectively.

References

[1] *Coordinated operational planning* (UCTE OH–Policy 4, final policy 2.0 E, 03.05.2006, available at http://www.ucte.org).

[2] *Current balance management in Europe* (Report issued by ETSO, December 2003, available at http://www.etso-net.org).

[3] *Benefits and practical steps towards the integration of intraday electricity markets and balancing mechanisms.* (Report prepared by Frontier Economics Ltd and Consentec for the European Commission, London, December 2005, available at http://ec.europa.eu/energy/gas_electricity/studies/electricity_en.htm).

[4] *Balance management harmonisation and integration* (4th Report issued by ETSO, January 2007, available at http://www.etso-net.org).

Modeling & Simulating Public Health Stemming from the Evolution in the Electric Energy Portfolio of Southeastern Virginia

Dr. Joshua G. Behr
Virginia Modeling, Analysis, & Simulation Center (VMASC)
Old Dominion University, Norfolk VA 23435

and

Dr. Rafael Diaz
Virginia Modeling, Analysis, & Simulation Center (VMASC)
Old Dominion University, Norfolk VA 23435

ABSTRACT

Southeastern Virginia is expected to follow the national trend towards investment in renewable energy approaches as part of its electricity portfolio. The evolution in portfolio mix away from traditional methods, such as centralized nuclear and coal-fired generation, and towards cleaner- and renewable-source generation will impact public health. This research proposes a System Dynamics (SD) approach to better understand the intersection between potential near-future electric energy portfolios and public health. Fossil fuel-based emissions have long been recognized as contributing to deleterious health conditions and hospital admissions among adults and children. The inclusion of such health-related externalities is essential in assessing the costs of electric energy portfolio combinations. This research follows three steps: 1) identification of electric energy generation sources that have the potential to be included within a near-future electric energy portfolio, 2) defining the public health costs associated with each of these sources, and 3) modeling the interactions among these energy sources and public health to produce a best-feasible energy portfolio solution. Proffered is an optimization model that efficiently explores the complex decision space inclusive of the nature of the energy sources, storage, and transmission, as well as the public health.

Keywords: Electric Energy, Renewable Energy, Public Health, Community Health, Coal, Nuclear, Solar Photovoltaic, Wind, System Dynamics

1. INTRODUCTION

Conventional methods of electricity production derived from natural gas, oil, and coal, while traditionally abundant, reliable, and affordable, emit into our environment (air, water, and soil) the pollutants sulfur dioxide (SO_2), nitrogen oxide (NO_X), carbon dioxide (CO_2), and mercury. Within the United States, the generation of electricity alone contributes substantially to the overall amount of these pollutants emitted into our environment [1]. Each of these pollutants has been recognized to have a detrimental impact on several dimensions of public health ranging from chronic conditions such as respiratory ailments (e.g., asthma), disease such as lung cancer, and neurological disorders (consumption of mercury-laden fish).

The production cost of fossil-based generation of electricity has been known to fluctuate greatly over relatively short periods. Short-term instability and uncertainly in oil and natural gas prices may stem from geo-political machinations as well as both speculated and real disruptions in the supply chain. The cost of renewable energy production, many forms of which may be produced domestically if not relatively close to the point of consumption, may be insulated from the sever price swings associated with a heavy reliance on fossil fuels [2]. Renewable energy production methods, due to advances in engineering and technology, are increasingly becoming recognized as reliable. Yet despite recent efficiencies, renewable-source generated electricity is still at a competitive disadvantage cost-wise relative fossil-based energy production. Many have argued, though, that this disadvantage stems from the fact that the market cost for electricity generated from fossil fuel does not take into account the externalities of energy production relating to poor air, water, and ground quality and the attendant public health issues [3]. That is, external to the compilation of the price of electricity to the consumer are the quality of life and medical costs imposed upon the communities. Internalizing these externalities into the cost of fossil-based electricity would, intuitively, lessen the price competitive advantage of fossil- over renewable-generated electricity. Although the identification and quantification of these factors external to those parties explicitly involved in the transaction may be difficult and counterintuitive to established approaches, the inclusion of these externalities into the cost equation is necessary for informed policy formation [4][5][6]. Modeling an energy system to identify the social and environmental impacts of different combinations of energy generation, transportation, and distribution has been identified as an important goal in reducing uncertainly and informing policy [7].

2. RESEARCH QUESTION

Virginia as a whole, and Southeastern Virginia specifically, can be expected over the next several decades to mirror the trend, seen nationally across many other states, towards experimentation and application of innovative renewable energy approaches as part of its electricity portfolio. Within Southeastern Virginia, an evolution in portfolio mix away from traditional methods of generation and towards cleaner- and renewable-source generation can be expected to impact the externalities associated with public health, thus affording the region a 'health dividend' [8]. What is uncertain are the various impacts on public health stemming from different energy portfolio mixes. Thus, this research presents an approach to model and simulate the impacts on public health for the likely near-future electric energy portfolios within Southeastern Virginia.

3. POLICY ENVIRONMENT

The Environmental Protection Agency's Clean Air Interstate Rule (CAIR), targeting largely the eastern half of the United States (east of the Mississippi), compels states to permanently reduce NO_X and SO_2 emissions to predetermined caps by the years 2009/2010 and 2015. Virginia's NO_X reductions, relative 2003 levels, constitute roughly a 53 percent decrease by 2009 and a 61% decrease by 2015 (1.5m and 1.3m tons, respectively). SO_2 reductions are 45 percent by 2010 and 73 percent by 2015 (3.6m tons and 2.5m tons, respectively). The Clean Air Mercury Rule (CAMR), national in scope, sets for petroleum-based electricity facilities allowances that may be bought and traded in efforts to meet targets set for 2010 and 2018. Conceptually, Environmental Portfolio Standards (EPS), common in many states, set targets for the percentage of electricity sold within the state that must be produced by renewable sources. The State of Virginia itself has set a nonbinding Renewable Portfolio Standard (RPS) for utilities with the state-wide goal of 12 percent of energy sales to be derived from renewable energy by 2022 (more precisely, 12 percent of 2007 base year sales by 2022) [9].

A majority of US States are now covered by renewable energy/electricity standard (RES). These standards contain targets for the proportion of a state's portfolio of electricity sources that must be generated by means of renewable energy. A region's evolution towards a more diversified portfolio of energy sources may be viewed as laudable due to a string of expected benefits derived from such evolution. These benefits may include less dependence on foreign oil, decreased degradation of the environment inclusive of global warming, and a range of community and environmental health issues [10, 11]. However, given the potential diversity in the composition of a region's electric energy generation and distribution system, there may be great variation in the public health impact [3].

The latent renewable energy potential for a region, stripped of other concerns, is driven by geography. Some regions, due to climate, elevation, and proximity to water, may be relatively rich in a particular potential renewable energy resource and deficient in others. Some regions have witnessed an emphasis of interest in solar photovoltaic applications due to climate, while others with conducive topography may have a stronger suit in land-based wind- and hydro-generated electricity, and still others with proximity to oceans and bays have entertained offshore wind- and tidal-generated electricity. Secondary to geography, the focus on type of renewable energy source may be conditioned by current land use (e.g., competition with agriculture, urbanization, state and national parks and recreation areas, etc.), public acceptance of altered viewscapes (e.g., horizon containing wind farm turbines, transmission and distribution lines, etc.), and wildlife protection/conservation efforts (e.g., facility encroachment on protected or fragile ecosystems). Within Southeastern Virginia there has been, for example, considerable discussion of the potential for offshore wind initiatives within the context of the impact of altered viewscapes on tourism and property values. The intensity and direction of these secondary drivers on the penetration of renewable energy sources into the marketplace has been, and will continue to be, a product of balance struck within our political system. The goal here is not to predict which combination of energy sources will be employed to serve Southeastern Virginia but, rather, to simulate the public health impact of a variety of likely energy portfolios.

4. RESEARCH APPROACH

This research proposes a Modeling and Simulation (M&S) approach to investigate and better understand the intersection between likely near-future electric energy systems and public health within Southeastern Virginia. The optimal mix of parameters within an electric energy production, storage, and distribution system designed to maximize efficiency alone may be quite different from a mix that balances efficiency with public health. The three basic steps that gird our approach are as follows: 1) identify electric energy generation sources that have the potential to be included within a near-future electric energy portfolio, 2) define the public health costs associated with each component of the electric energy sources, and 3) model the interactions among the source-components and their associated public health costs to produce a best-feasible energy portfolio solution.

Model Parameterization: Electric Energy Portfolio

Step one includes specification of electric energy generation sources and their potential to be deployed within the evolving portfolio. This entails considerations of two broad categories, the first of which is the variety of potential energy sources such as nuclear, coal, oil, natural gas, solar photovoltaic, wind, biomass, municipal solid waste landfill gas, hydro, tidal, and geothermal. The second consideration is the magnitude or centrality of the energy production within any of these energy sources. A distributed-laden system may contain a decentralized network of energy production sites, as may be employed in a residential distributed photovoltaic (PV) energy arrangement or may be spatially concentrated to a utility scale as found in concentrated solar power (CSP). Conceivable, a national clean energy smart grid may direct energy long distances over high voltage lines from a renewable-rich region to the Hampton Roads urban area.

Model Parameterization: Public Health

Step two includes identification of public health issues associated with the type and placement of the components (inclusive of infrastructure such as transmission and distribution lines) that comprise the system. The design and construction of traditional electric energy systems has been heavily influenced by the desire to generate reliable, cost-efficient electricity, often at a distance from urbanized areas. The emphasis on cost efficiency led to reliance on energy-rich – and emission-laden – fossil fuels and a more centralized system of production. The public health impact of such an approach has been well documented. For example, for several decades it has been acknowledged that fine particle emissions generated from the combustion of fossil fuels in the generation of electricity contributes adversely to the health of our citizens [12]. But other forms of energy production, including renewable energy sources, while attractive relative to fossil fuels, may have definable public health impacts as well [13][14]. For example, the physical location of a biomass fuel production site, concentrated solar power site, or the more distributed placement of PV and the associated storage and transmission infrastructure (see Ferrey [15]) may condition property values and the built environment. The built environment may condition retail options, recreational choices, and availability of medical venues, all of which may be connected to community health. Collectively, the placement of an electricity network will surely impact the qualitative environment, viewscapes, green space, public space, property values, etc. [2][16]. Further, these changes in the landscape may have disparate impacts upon traditionally underserved populations. Unlike standard assessments of proposed portfolio mixes, our approach captures and quantifies the qualitative public health concepts that traditionally have not been considered in an evaluation of the optimal energy source mix within a portfolio.

Capturing Complexity

The third step in our proposed approach, the methods to be employed to sort through the complexities inherent in a system of competing interactions and feedback loops, relies on Systems Dynamics (SD). SD has been extensively used as a reliable approach that captures and analyzes the dynamic, complex interaction among dissimilar subsystems. The modeling strategy includes representing the structure of subsystems in terms of flows and stocks. As a result, components can be integrated as building blocks that articulate the interactions and dynamics among the involved subsystems. The SD framework acknowledges the complex interactions among many feedback loops, considers linear and nonlinear cause-and-effect, and considers the potential impact of effects on causes. SD transforms these components of causality into structured difference and differential equations. SD allows the revelation and quantification of 'unseen' dynamics that might have a relevant effect on public health, the central object of our interest.

Applications of System Dynamics

SD has been used as a framework to study renewable energy within broader systems. By integrating dissimilar systems, SD has been used to analyze the effects of emissions on a region [17], to analyze energy use at the national level [18], to determine the best strategy to meet future energy demand while reducing CO_2 emissions [19], to analyze highly seasonal and exponential electricity demand on policy making [17], to model changes of energy intensity in residential sectors [20], and to optimize existing methods of power production through government subsidies [21].

In addition, researchers have successfully included in SD models searching mechanisms that allow plausible combinations of parameters that optimize a set of goals [22][23][24][25][26]. Although the number of these simulation-based optimization studies is rather limited, Eksin [7] uses parametric search via genetic algorithm to meet multiple goals in the behavior of dynamic systems and Barlas [27] develop a supporting tool for pattern-based parameter searches that can be utilized in model identification, validation and policy analysis stages.

5. ELECTRIC ENERGY PORTFOLIO

The state-wide portfolio of electricity generated sources includes coal-fired (38.7 percent), natural gas-fired (18.8 percent), petroleum-fired (1 percent), nuclear (35.9 percent), hydroelectric inclusive of pumped storage hydro (3.2 percent), and other renewables including geothermal, photovoltaic, wind, municipal solid waste biogenic, and biomass (2.2 percent). Exclusive of pumped storage hydroelectric, in total renewables account for just 3.8 percent of the electricity generation. Currently, there is negligible solar and wind contribution to the power grid serving the greater Hampton Roads area. While the region's potential for engaging renewable energy sources has not yet been realized, there are technological, geological, atmospheric, cultural, and social parameters that place practical limits on both the state and region's capacity to evolve its energy portfolio towards renewable energy generation. Despite the technological advances in the design and construction of wind, solar and geothermal systems within the past 40 years, the initial capital cost to purchase and install many of these systems make them, in the short term, less economically attractive relative traditional sources of power such as coal and natural gas [28]. Government regulation can help spur the development, adoption, and implementation of clean coal technologies, but their emissions are still greater than other generating technologies [29]. Despite these very real constraints, the size of the region's appetite for electric consumption has grown by roughly 3 percent per year. The Virginia Energy Plan, adopted by the state in the Fall, 2007 and extending through 2016, seeks to slow this rate of growth (especially through energy conservation) as well as increase electricity production to meet the expanding demand.

The Hampton Roads region is home to three coal-fired plants, two natural gas plants, and one of the state's two nuclear power plants (the plant, in Surry County, contains two reactors). Aside from nuclear, coal is the state's primary resource for electricity generation. While the majority of Virginia's coal resources are located in the Central Appalachian Basin (Dickenson, Wise, and Buchanan Counties), roughly one-third of the coal consumed for electric power generation comes from the nearby coal fields in West Virginia and Kentucky. In addition to this 'imported' coal, in total nearly one-fifth of the electricity consumed within Virginia comes from neighboring states.

Load Production

Electric power generation must meet the minimal, continuous demand (base-load production) and also meet the seasonal and daily cyclical demand above this baseline (peak-load production). Traditional coal and nuclear plants may not easily be turned on and off to meet either a surge or a decrease in demand. This necessarily means that peak demand must be satisfied either with stored electrical energy or power plants that may be brought on- and off-line rather quickly (e.g., gas fired plants or hydro pumped storage). The production of peak-load electricity generally is more expensive per megawatt hour relative base-load production.

Wind

The efficient generation of electricity from wind requires sustainable winds. With the exception of ridgelines in the mountainous western third of the state (encompassing state and national parks), land-based wind generation in Hampton Roads is unlikely due to a wind power class of 3 or less (< 15.7 mph). Thus, an increase in generation of land-based wind power would necessarily not be within, or immediately proximate to, the Hampton Roads region. Two new, large wind turbine projects may come online in Tazewell and Wise Counties, both located close to the Kentucky and Tennessee borders and far from transmission to the Hampton Roads region. The greatest demonstrated potential for wind power close to the region is offshore, much of it in and around the Chesapeake Bay and in close proximity to the Hampton Roads population centers. Beyond 20 miles offshore, substantial territory exhibits class 5 and 6 winds (>16.7 mph) [30]. Questions about the impact on quality of life and tourism stemming from diminished viewscapes as well as potential interference with shipping and naval testing pose constraints to offshore wind power generation.

Photovoltaic

The region's solar photovoltaic potential is conditioned by intermittency related to weather conditions [7]. Concentrated solar power is land-intensive, competes with agriculture, urban development, recreation, and conservation. Many of the urban areas offer potential commercial and industrial rooftop and parking garages suitable for decentralized, small scale generation. The more proximate nature of a distributed system allows for fewer miles of transmission (thus potentially reducing network congestion) and avoiding the need for capital investment in transmission infrastructure [30]. Nevertheless, the cost of solar photovoltaic adoption is intially expensive relative traditional sources of generation. Photovoltaic devices also have brought opportunities to homeowners to engage renewable energy at the point of consumption. The adoption of smart meeting (bi-directional metering) and digital, real-time communications among electricity suppliers and consumers has the potential to add capacity to the existing network, although directional transmission may frustrate this potential. While the running cost of these systems is quite attractive for the homeowner, the initial capital investment in the technology may be daunting for modest income households.

Landfill Gas

The state currently has 22 functioning landfill gas sites and 11 landfill sites that are candidates for gas production. Within the Hampton Roads region, there are four operational landfill gas sites, two of which generate electricity. The region has only two other landfill sites that are identified as untapped candidates [31]. Landfill gas is one of the lowest cost renewables, however the potential for substantial new electricity generation from this source is quite limited.

Hydro Electric

Electric demand within the region also may be satisfied from hydro electric power generation. Hydro electric power generation is considered a renewable resource and may contribute to both base- and peak-load generation. Although it does not meet the State's legal definition of renewable energy, pumped storage hydroelectric is useful to meet peak-load generation. During times opposite peak-load production, energy from nuclear or coal plants is used to pump water to an upper reservoir in preparation for release during peak-load production. Within transmission distance, potential hydroelectric sites have been identified that could contribute to renewable electricity generation for the region [32]. However, the expansion of existing hydro electric plants or the construction of new plants may face strenuous environmental hurdles.

Biomass

The region is also home to several energy recovery incineration facilities. The region's urban areas produce large amounts of biomass in the form of wood residue; the region is inclusive of forests and agricultural activities that may provide additional biomass; some areas may be conducive to switchgrass, although large scale experimentation has not yet been conducted. Avoiding substantial additional capital investment in new facilities, much of this biomass has the potential to be co-fired within existing coal-fired power stations as is already the practice at power plants located near logging operations and sawmills. The energy density of biomass is generally much less relative fossil fuels and has higher transportation costs, however it is considered carbon-neutral.

6. PUBLIC HEALTH

Evolution towards Centralized Power Generation

The complexion of the energy system is dependent, in part, on advancements in technology and the public's acceptance of the risks and costs associated with adopting the technology. The availability and access to affordable technology conditions public health. In less developed regions, fuels such as wood, charcoal, and coal may be used for cooking and heating within the home, contributing to poor indoor air quality and associated poor health conditions. In more developed regions where the combustion of fuel is centralized and the generated electricity is distributed, the health-related impact stemming from proximate biomass conversion within the home is reduced [33].

The current system within Southeastern Virginia is dominated by the coal- and gas-fired generation of electricity at centralized locations and the high voltage transmission across long distances to stations where it is distributed at lower voltage within a localized system. Relative to pre-centralized energy generation, the use of petrochemical products and the centralized generation of fossil-derived electricity has contributed substantially to better health and increased quality of life [33][34]. In addition, modern industrialized nations have seen gains in energy efficiency (e.g., decreases in the amount of energy to produce a lumen of light, propel an automobile, transport via airplane a person or cargo across the continent, or heat a typical residence). However, these efficiencies have also coincided with the overall growth in both affluence and energy consumption which, in turn, implies increased emissions stemming from combustion of fossil-based fuels [35].

Particulate Emissions

Combustion produced emissions associated with electricity generation is a primary health concern. On average, the state produced 4.3 lbs of sulfur dioxide, 1.7 lbs nitrogen oxide, and 1,254 lbs carbon dioxide per megawatt hour generated [31], although urban areas tend to have higher carbon dioxide emissions relative rural areas [36]. Emissions from coal-fired power plants have long been recognized as contributing to the respiratory conditions and hospital admissions among adults and children [37][38][39] and premature mortality [40]. Urban air pollution stemming from the combustion of fossil-based fuels is characterized by atmospheric gases and particulate matter. Short-term exposure to airborne particulate matter (PM) in our urban environments has been associated with adverse pulmonary health effects and respiratory distress [41][42][43]. The composition and size of this airborne matter, usually in the form of respirable particles in the range of 10μm or less, allows it to enter the respiratory system resulting in any number of symptoms including wheezing, coughing, infection, and impaired lung function. We can place an economic valuation (from a statistical life approach) on the premature mortality, illness, loss from work, and occupational exposure associated with the various combinations of electricity sources within an energy portfolio (e.g., [44][45]). Electricity generation from renewable sources has fewer detrimental health-related outcomes relative to traditional gas- and coal-fired plants due to differences in type and quantity of emissions. While particulate matter in the form of air pollution posses a particular burdensome health concern, the construction and maintenance of the power systems including the transmission and distribution systems impose occupational hazards [45] as well as the potential for accidents and terrorism [28].

7. MODELING COMPLEXITY

Causal Diagrams

Causal diagrams to describe complex causal systems and analyze the health impacts of regulatory policy interventions have been suggested as enabling techniques that may allow researchers to capture the complexities of a broader healthcare system [46][47][48]. Figure 1 (below) presents the high-level organizational logic of the relationships among the electricity generating sources and public health. The Figure illustrates at the most general level the several positive and negative feedback loops comprising the system. Not illustrated within this diagram are further subsystem components that comprise, for example, the various levers exercised through government intervention, the development and adoption of emission control technology, and state dependency on imported energy.

The degree of 'greenness' within a portfolio is a function of the variety, the carbon neutrality, and the distributed nature of the electricity sources. An increase in a portfolio's greenness contributes to a decrease in emissions which, in turn, conditions the quality of community health. The public's awareness of the connection between emissions and adverse respiratory conditions, for example, is a product of the presence of those respiratory conditions; as the quality of respiratory-related community health makes strides, the public's energy-health awareness declines and this, in turn, conditions public support for government interventions. An increase in a portfolio's greenness – and the degree to which the evolution is towards a distributed system – necessarily alters viewscapes and places the infrastructure in competition with land which, in turn, conditions property values and contributes to our assessment of the built environment, the built environment having implications for community health and public support for government intervention. The portfolio's greenness is also related to a decrease in reliability in the short-term and an increase in cost of kwh generation in the short-term, both of which condition public support.

Government attempts to control the direction and rapidity of the evolutional movement in the electric energy portfolio can take the form of subsidies or land trades at the municipal and county levels, price regulation and targets for energy source generation at the state level, and tax credit incentives for the development of emission control technology, traded energy credits, and unfunded clean air mandates at the national level. These interventions and perceptions about the government's long-term commitment condition public-private partnerships and the private sector's willingness to phase-out older technology and commit capital to emerging renewable energy sources.

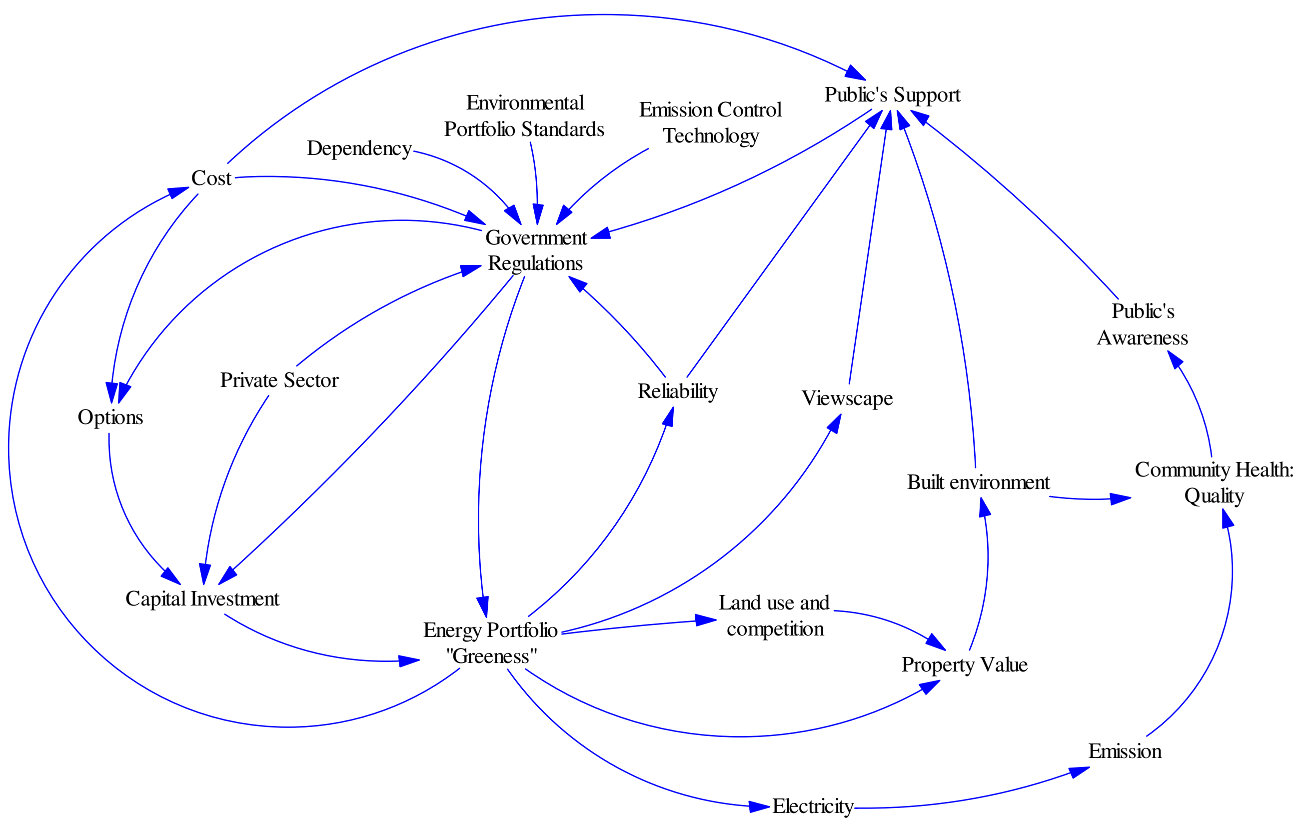


Figure 1: High Level Representation of Feedback between Electric Energy Portfolio and Public Health

8. CONCLUSIONS

The evolution of an electricity portfolio does not take place within a static regional environment; it is expected that the region's demand for electricity will grow over the coming decades, despite conservation and efficiency. Growth in the share of the region's electricity generation stemming from renewables (and less carbon-intense methods) does not necessarily assure a decrease in related disease and chronic conditions. It is conceivable, with the expansion in demand, that even modest growth in renewable-generated electricity may cover only this increased demand, let alone the replacement of electricity generated from retiring coal-fired plants or less-emissive gas-fired plants. The complexity of the system is evident: there is an imperative to meet exiting demand as well as the anticipated growth in demand, but this is within the constraints of the EPA's regulations and emissions caps, maintaining reliable and affordable electricity, and improving the quality of public health. The diffusion across the Hampton Roads region of less carbon intense forms of electricity production may not solely be a matter of technology, but one of cautious acceptance as the socio-political environment weighs the competing cost-benefit claims relating to economics and community health. Proffered is an optimization modeling approach that efficiently explores using positive and negative feedbacks a complex decision space that takes into account the variety of the energy sources, inclusive of the distributive nature of these sources, and public health

9. REFERENCES

[1] Palmer, K., D. Burtraw, and J. Shih. "The Benefits and Costs of Reducing Emissions from the Electricity Sector", **Journal of Environmental Management**, 83, 115-130. 2007.

[2] Akella, A.K., R.P. Saini, and M.P. Sharma. "Social, Economical and Environmental Impacts of Renewable Energy Systems", **Renewable Energy**, 34, 390-396. 2009.

[3] Ratliff, N. and D.H. Smith. "Renewable Energy Electricity State Level Policies in the WRAP Region: What, Why, and Maybe How", **Energy Sources**, 27, 431-444. 2005.

[4] Haines, A. "Sustainable Energy, Economic Growth and Public Health", **Medicine**, **Conflict, and Survival**. 17: 56-62. 2001.

[5] Hirschberg et al. "Health and Environmental Impacts of China's Current and Future Electricity Supply, with Associated External Costs", **International Journal of Global Energy Issues**. 22: 155-179. 2004.

[6] Oikonomou, V. et al. "Energy Saving and Energy Efficiency concepts for Policy Making" **Energy Policy**. 37: 4787-4796. 2009.

[7] Eksin, C. "Genetic Algorithms for Multi-Objective Optimization in Dynamic Systems" **The 2008 International Conference of the System Dynamics Society**. 2008.

[8] Mccartney, G., P. Hanlon, and F. Romanes "Climate Change and Rising Energy costs will Change Everything: A new Mindset and Action Plan for 21st Century Public Health". **Public Health**, 122, 658-663. 2008.

[9] Virginia General Assembly, H.B. 3068/S.B. 1416 (two identical bills approved by the GA April 4, 2007 and signed into law by Governor Kaine on July 1, 2007).

[10] Johansson, T.B. "**Renewable Fuels and Electricity for a Growing World Economy: Defining and Achieving the Potential**", Renewable Energy: Sources for Fuels and Electricity. Island Press. 1993.

[11] Romm, J. "How Energy Policies Affect public Health". **Public Health Reports**. 3(5), 390-399. 1996.

[12] Physicians for Social Responsibility. "Coal-Fired power Plants: Understanding the Health Costs of a Dirty Energy Source". 2009.

[13] Krewitt, W. "Comparison of Health and Environmental Impact from Electricity Generation Systems", **International Journal of Risk Assessment and Management**, 3. 2002

[14] Rabl, A. "Health and Environmental Impacts of Energy Systems" **International Journal of Global Energy Issues**, 18, 113-150. 2003.

[15] Ferrey, S. "Converting Brownfield Environmental Negatives into Energy Positives." *Environmental Affairs Law Review*, 43, 417-478. 2007.

[16] General Accounting Office. "Electricity Supply: consideration of Environmental Costs in Selecting Fuel Sources", GAO Report. 1995.

[17] Gravouniotis, P. and B. Ausilio. "Energy Equipment Diffusion & Touristic Competitiveness: Building of an SD Model for the Greek Islands" **International Conference of the System Dynamics Society.** 2008.

[18] Ballardin, Giorgio, Di Giulio, Enzo, and Migliavacca, Stefania. "Renewable energy in Italy: scenarios to 2030" **International Conference of the System Dynamics Society**. 2008.

[19] Choucri, N., D. Goldsmith, and T. Mezher. "Framework for Modeling Technology Policy: Renewable Energy in Abu Dhabi" **International Conference of the System Dynamics** Society. 2008.

[20] Jamshidi, Mohamed M. "An Analysis of Residential Energy Intensity in Iran, A System Dynamics Approach" **International Conference of the System Dynamics Society**. 2008.

[21] Keimanesh, M., A.Y. Salekdeh, H. Vaziri, and S. Barahmand. "Effects of Gas Subsidy on the Behavior of Power Stations in Iran: A new policy to reduce energy intensity in electricity section". 2008.

[22] Graham, A. and C. Ariza. "Dynamic, hard and strategic questions: using optimization to answer a marketing resource allocation question", **System Dynamics Review**, 19, 27-46. 2003.

[23] Graham,A.K. and Morecroft,D.W. "Model-supported case studies for management education" **European Journal of Operational Research**, 151-166. 1992.

[24] Miller, J.H. "Active Nonlinear Tests (ANT) of Complex Simulation Models" **Management Science**, 44, 820-830. 1998.

[25] Dangerfield, B. and C. Roberts. "Optimization as a statistical estimation tool: An example in estimating the AIDS treatment-free incubation period distribution" **System Dynamics Review**, 15, 273-291. 1999.

[26] Coyle, G. "Qualitative and quantitative modelling in system dynamics: some research questions", **System Dynamics Review**, 16, 225-244. 2009.

[27] Yucel, G and Y. Barlas. "Pattern-based System Design / Optimization" **International Conference of the System Dynamics**. 2007.

[28] Virginia Center for Coal and Energy Research, Virginia Polytechnic Institute and State University.. "A Study of Increased Use of Renewable Energy Resources in Virginia". 2005.

[29] Heinz Center. Health Implications of Technological Responses to Climate Change. Report of Workshop Held November 29-30, 2000.

[30] Virginia Coastal Energy Research Consortium (VCERC), http://www.vcerc.org/

[31] Environmental Protection Agency (EPA), Landfill Methane Outreach Program (LMOP) http://www.epa.gov/lmop/ . Retrieved 5/2010.

[32] Idaho National Engineering and Environmental Laboratory (INEEL). Estimation of Economic Parameters of U.S. Hydropower Resources. Document EXT-03-0062. 2003.

[33] Wilkinson P., K.R. Smith, S. Beevers, C. Tonne, and T. Oreszczyn. "Energy, energy efficiency, and the built environment" **Lancet.** Sep 29; 370 (9593):1175-87. 2007.

[34] Frumkin H, J. Hess , S. Vindigni. "Energy and public health: the challenge of peak petroleum", **Public Health Report**. Jan-Feb;124(1):5-19. 2009.

[35] Tatum, J.S. 1994. "Technology and Values: Getting Beyond the 'Device Paradigm" **Impasse. Science, Technology, & Human Values**. Winter: 70-87.

[36] Parshall, L. et al.. "Modeling Energy Consumption and CO_2 Emissions at the Urban Sale: Methodological Challenges and Insights from the United States", **Energy Policy**. 10: 1-18. 2009.

[37] Cifuentes, L. et al. "Hidden Health Benefits of Greenhouse Gas Mitigation". **Science**: 293: 1257-1259. 2001.

[38] Szyszkowicz, M. 2007. "Ambient Air Pollution and Daily Emergency Department visits for Asthma in Edmonton, Canada". **International Journal of Occupational medicine and Environmental Health**. 21: 25-30.

[39] Peel, J. et al. "Ambient Air Pollution and Cardiovascular Emergency Department Visits in Potentially Sensitive Groups" **American Journal of Epidemiology**. 165: 625-633. 2007.

[40] Hermann, R.P. and J.O. Lanier. "Predicting Premature Mortality from New Power Plant Development in Virginia", **Archives of Environmental Health**. 59: 529-535. 2004.

[41] Dockery, D.W. et al. "Change in Pulmonary Function in Children Associated with Air Pollution Episodes", **Journal of Air Pollution Control Association**. 32: 937-942. 1982.

[42] Folinsbee, L.J. "Human Health Effects of Air Pollution. Environmental" **Health Perspectives**. 100: 45-46. 1993.

[43] Ezzati, M. et al. "Selected Major Risk Factor and Global and Regional Burden of Disease", **Lancet** 360: 1347-1360. 2002.

[44] Levy, J.I. et al. "Uncertainty and Variability in Environmental Externalities from Coal-fired Power Plants in the United States", **Risk Analysis.** 29: 1000-1014. 2009.

[45] Bird, M.J. et al. "Occupational Exposure During Routine Activities in Coal-Fired Power Plants", **Journal of Occupational and Environmental Hygiene**. 1: 403-413. 2004.

[46] Joffe,M. and J. Mindell. "Complex Causal Process Diagrams for Analyzing the Health Impacts of Policy Intervention", **American Journal of Public Health**, 96, 473-479. 2006.

[47] Homer, J. and G. Hirsch. "System Dynamics Modeling for Public Health: Background and Opportunities" **American Journal of Public Health**, 96, 452. 2006.

[48] William, M.T., A.C. Derek, M. Bobby, S. Richard and J.L. Scott. "Practical Challenges of Systems Thinking and Modeling in Public Health", **American Journal of Public Health**, 96, 538. 2006

Cost and Emissions impacts of Plug-In Hybrid Electric Vehicles on Ohio Power Grid

Riccardo FAGIANI
Dipartimento di energia, Politecnico di Milano,
Milano, 20133, Italy

Vincenzo MARANO
Center for Automotive Research (CAR), The Ohio State University
Columbus, OH, 43212, USA

Ramteen SIOSHANSI
Integrated Systems Engineering, The Ohio State University
Columbus, OH, 43212, USA

ABSTRACT

This paper analyzes the impacts of plug-in hybrid electric vehicle (PHEV) charging on the Ohio power system to estimate the net cost and emissions impacts of PHEV use. We consider two charging scenarios—one in which the grid operator makes charging decisions and another in which these decisions are made by individual vehicle owners and consider cases with PHEV penetrations of between 1% and 5%. Results show that PHEV use would result in a close to 70% reduction in gasoline consumption compared to conventional vehicles and a 50% reduction in driving costs. The emissions impacts are more mixed, with some pollutants being reduced and other increased, due to use of coal-fired generation in Ohio.

Keywords: Plug-in Hybrid Electric Vehicle (PHEV), Grid, Charging, Emissions, Costs.

INTRODUCTION

Plug-In Hybrid Electric Vehicles (PHEVs) have been promoted as a potential technology that can reduce vehicle's fuel consumption and decrease transportation-related emissions, overall costs, and oil dependence. This is obtained by using electricity as a primary source of energy to run the vehicle, using gasoline as a secondary backup fuel. Several studies have found that by using electricity, PHEVs could emit less CO_2 and certain other pollutants over their entire fuel cycle than conventional vehicles (CVs) and hybrid-electric vehicles (HEVs) [1], [2], [3], [4], [5], [6]. This is because in many regions grid electricity is effectively a cleaner source of transportation fuel than gasoline. Furthermore, pollutants can be better-controlled in power generation plants and there are local benefits from shifting emissions away from population centers to point sources such as power plants. Emissions impacts of PHEVs will depend strongly on the generation mix present in the power system; clearly, a high penetration of cleaner sources of energy such as solar, wind, hydroelectric, or nuclear energy could lead to a great decrease in pollutants emissions while a predominance of coal as primary source of energy is less effective in this sense.

Stephen and Sullivan [3] estimated average PHEV-related CO_2 emissions for the U.S., showing that a PHEV will emit 221 g/km of CO_2, about 51% of conventional vehicle's emissions. Similar results were obtained in a study on the Colorado electric system, which estimates that a PHEV will emit about 60% of the CO_2 normally produced by a conventional vehicle [2]. PHEVs could also help countries reduce oil dependence by decreasing gasoline consumption of vehicles running today [6], [7]. Since electricity is less expensive than gasoline, PHEVs could provide important operational cost savings, making them attractive for future buyers despite the additional initial cost compared to CVs and HEVs [7], [8], and representing a first step toward future transportation electrification.

MODELING ASSUMPTIONS

This paper describes an analysis of the effect of a PHEV fleet on the Ohio power system. This analysis helps determine how a PHEV fleet would affect net emissions and costs and whether PHEVs would be economically attractive for new customers. Two different scenarios are considered: a **controlled charging** scenario, which assumes that the utility co-optimizes the operation of generators and the charging of PHEVs to minimize the total cost of generation and vehicle driving; and an **uncontrolled charging** scenario in which PHEVs are recharged whenever they are connected to the grid (without regard to charging cost), and the utility must serve these loads. This scenario assumes that because electricity is a cheaper source of transportation energy, PHEV drivers will always opt to recharge their vehicles when they are not being driven.

PHEV data

For each set of model runs, the PHEV fleet is assumed to consist of a fixed number of vehicles. The total vehicle fleet size is taken from 2007 Ohio vehicle registration information reported by the U.S. Department of Transportation's Federal Highway Administration [9]. Simulations are carried out with three different PHEV penetration levels: 1%, 3% and 5% of the total vehicle fleet. Vehicle driving patterns are based upon empirical driving data collected from a sample of drivers with instrumented vehicles. Figure 1 represents some statistical information about drivers.

Driving data are coupled with the ADVISOR vehicle simulator model to estimate gasoline and electric consumption of an equivalent PHEV. Based on this empirical data different driver profiles were developed. The model assumes that the PHEV fleet is evenly divided into the driving profiles corresponding to the driving pattern data. Driving data were used to determine the hours in which the PHEVs are driven, the total distance traveled in that hour, and those in which they are grid-connected and could recharge their batteries. In doing so, it is assumed that a PHEV must be parked for an entire hour to be considered 'grid-connected' in that hour.

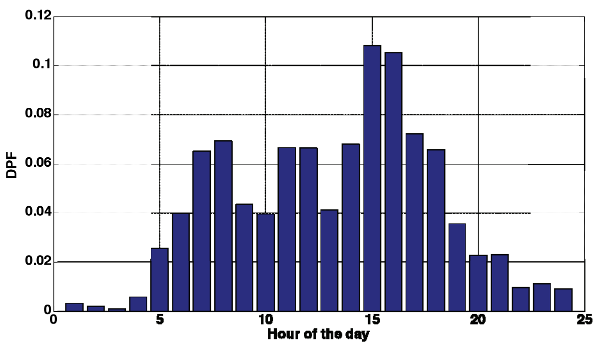


Figure 1: Average distribution of miles driven by drivers during the day

Depending on the state of charge (SOC) of a PHEV's battery, the vehicle will either be driven in charge-depleting (CD) mode, in which case the battery is the primary energy source and the gasoline engine is used only on a supplemental basis (for quick accelerations), or charge-sustaining (CS) mode, in which case the gasoline engine is used to maintain the same average SOC (driving like an HEV). The adopted control strategy is EV (electric vehicle) mode control—the vehicle is driven as an electric vehicle until the lower limit of SOC is reached, then the vehicle operates in CS mode as an HEV. The empirical driving data and ADVISOR model are used to estimate the average gasoline and battery energy usage for each PHEV driving profile in both CD and CS modes. The model also assumes that PHEVs always have sufficient gasoline to operate in either CS or CD mode. It is assumed that the charger has a power capacity of 1.875 kW and charging efficiency of 90%.

Generators data

Generation costs were calculated based on estimated generator heat rates and fuel costs. Heat rates were estimated starting from historical continuous emissions monitoring system (CEMS) data from the U.S. Environmental Protection Agency (EPA) [10]. CEMS data specifies the generation and total heat content of fuel burned by each generator in each hour. These data were used to estimate:

- *startup fuel* burned whenever the generator is brought online from an offline state;
- *spinning no-load fuel* burned whenever the generator is operating, independent of its generation output.;
- *variable fuel* burned depending on the electric output of the generator and was assumed as a step function.

Figure 2 shows the actual CEMS data for one generator in our sample set and the estimated variable fuel cost. CEMS data were also used to determine the minimum power output of each generator (when online), ramping limits, and minimum up and down times when each generator is started up and shutdown. Fuel costs were estimated based on purchase price data reported in form FERC-423, as reported by the U.S. Department of Energy's Energy Information Administration (EIA) [11]. Nuclear generators are assumed to be non-dispatchable and run at full capacity, thus their generating costs are modeled as constant values. Because Ohio is a part of the PJM system, the state may export or import energy from neighboring control areas, depending on the relative cost of energy that can be imported from or exported to the rest of the of the PJM system

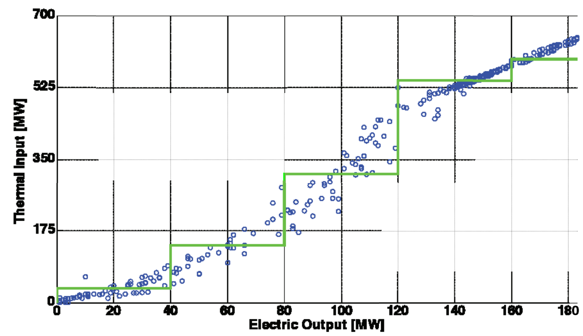


Figure 2: Heat rate data of one of the generators taken from EPA

The price at which energy can be bought and sold will generally vary depending on the volume of transactions, and this is captured in the model by assuming that the price of energy that is bought and sold from the rest of the market is a function of the transacted quantity. Specifically, historical PJM day-ahead market bid [12] data were used to estimate the relationship between price and load. As shown in Figure 3, the energy price function is derived from the bids based on their merit order. The actual historical PJM load is used to shift the market price function horizontally, and the shifted function is used to determine the price of net energy transactions between Ohio and the rest of the PJM market. Market energy transactions are then taken into account when calculating generator's emissions, as explained in the next section.

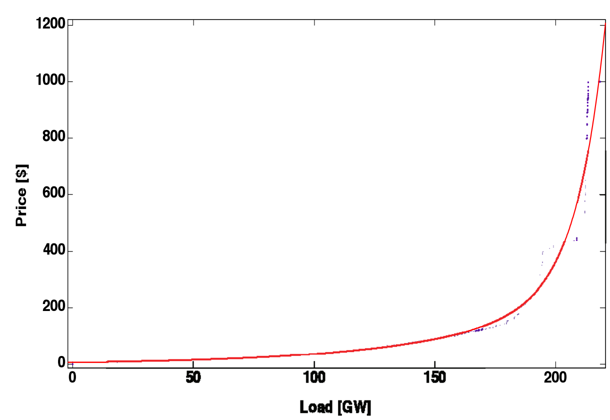


Figure 3: Price of market transactions as a function of transacted energy

Power system model

The analysis is based upon a unit commitment model of the Ohio electric power system, which is formulated as a mixed-integer program using the AMPL and solved using the branch and cut algorithm in cplex 12.1. The model simulates the commitment and dispatch of conventional generators and PHEV charging behaviors. The unit commitment model has a one day planning horizon (typical of day-ahead electricity markets) with an hourly time-step for the commitment and dispatch variables. Each of the 365 days in the sample is simulated independently, except that the commitment and dispatch of each conventional generator and the charge level of each PHEV battery at the beginning of each day is fixed based on the ending values from the previous day's run. Each day's unit commitment is solved in two steps. First a unit commitment model with a two-day planning horizon and a four-hour timestep for the commitment variables (the dispatch variables still have an hourly timestep in this first commitment problem) is used to determine and fix the ending commitment and dispatch of each generator and charge level of each PHEV battery. After these variables are fixed, the one-day problem is solved with hourly timesteps for all of the variables.

PHEV charging decisions are modeled differently in the controlled and uncontrolled charging scenarios. In the controlled charging scenario, the grid operator makes all charging decisions and coordinates these with power system operations. The grid operator's model consists of both the power system and PHEV models, and the grid operator's objective is to minimize the sum of generation costs and the cost of gasoline used by the PHEVs (the cost of electricity used by the PHEV fleet is accounted for in the total generation costs). The controlled charging model also includes a constraint to ensure that each PHEV battery is fully recharged in time for the first vehicle trip of each morning. In the uncontrolled charging scenario, PHEV owners are assumed to make charging decisions on their own without any regard for the impact of vehicle charging on the power system.

Model simulations are used to evaluate the cost and emissions impacts of PHEV use. The analysis considers three different emissions: CO_2, NO_x and SO_2, and considers emissions from two different sources: generator emissions and PHEV tailpipe emissions. Input-based generator emissions rates are estimated using 2007 CEMS data. These emissions rate estimates are combined with the simulated commitment and dispatch for Ohio generators to determine total emissions for the Ohio generator fleet. Energy that is imported into or exported from Ohio will also have an emissions impact—imports will result in greater emissions from generators outside of Ohio, and exports will reduce generator emissions outside of Ohio due to the fact that less energy needs to be supplied by the rest of PJM market. These emissions are estimated based on the hourly marginal fuel mix data (which specifies the mix of generating technologies that are marginal in each hour), reported by PJM, and estimates (based on the CEMS data) of generator emissions rates for each generation technology.

Tailpipe emissions are estimated using emissions regulations and gasoline chemical composition. CO_2 emissions are estimated at 8.87 kg/gallon. For SO_2 emissions, we assume that emissions will exactly comply with EPA's Tier2 requirement of 0.17 g/gallon. Tier2 also requires NO_x emissions to be less than 0.07 g/mile. We assume that CVs and HEVs will be designed to exactly meet these requirements and PHEV emissions are estimated from HEV emissions based on the proportional reduction in gasoline consumption.

RESULTS

PHEV charging scenario will influence the power demand on the grid and the dispatch of generators during the day. As shown in Figure 4, in the controlled charging scenario, the peak load is typically not increased by vehicles charging, since most of the charging is done in the morning (to exploit the availability of lower-cost energy). In the uncontrolled scenario, by contrast, the peak load is increased since PHEVs are plugged in during peak hours, as shown in Figure 5.

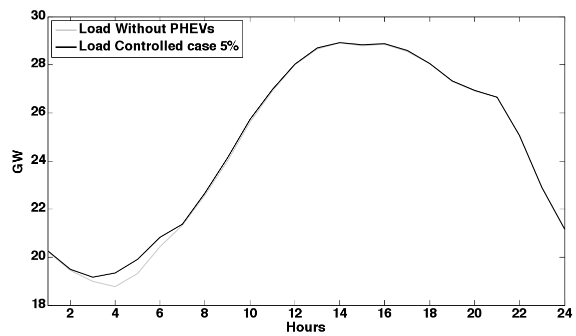


Figure 4: Total load on the grid with a controlled charging strategy on August 6th, summer peak

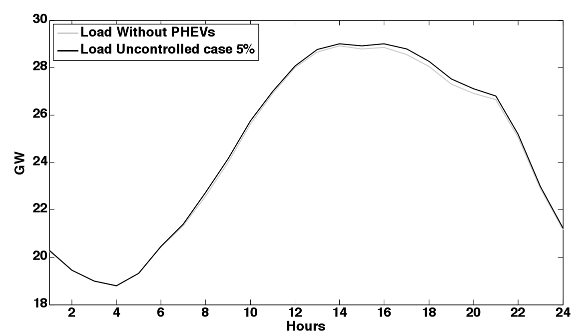


Figure 5: Total load on the grid with an uncontrolled charging strategy on August 6th, summer peak

Figure 6 and Figure 7 show that controlled charging would add a peak of 600 MW on the grid (due to the delaying of vehicle charging overnight), while uncontrolled charging would yield a less-intensive load of between 100 MW and 250 MW distributed during the afternoon. These differences in the PHEV charging load peak could have important implications for distribution-level constraints, since the more-concentrated PHEV charging that would occur with controlled charging may present challenges at the distribution-level for utilities and load-serving entities. A 5% fleet penetration in the uncontrolled charging scenario will add about a 300 MW load in the afternoon, less than 1% of total load; uncontrolled charging could add a more considerable load if the PHEV fleet grows in the future.

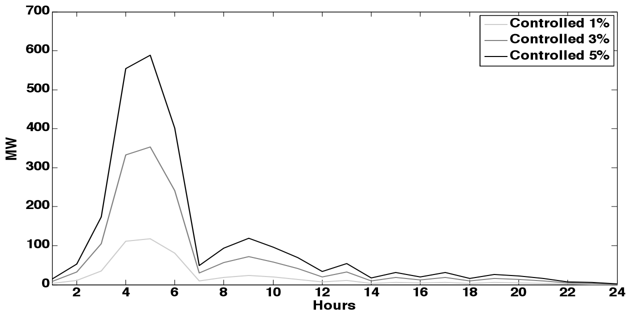


Figure 6: Total charging load on the grid with controlled charging on August 6th

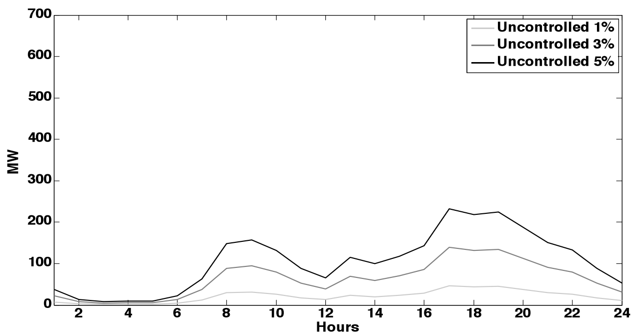


Figure 7: Total charging load on the grid with uncontrolled charging on August 6th

According to the National Research Council, PHEV market penetrations of up to 30% of the total vehicle fleets could be seen by 2050 [13]. Figure 8 shows that a 30% PHEV market penetration charged without any control from the grid operator would increase the Ohio system peak load by about 1.5 GW, shifting it later in the afternoon. This higher load could create problems to grid stability (assuming there is enough capacity to satisfy it).

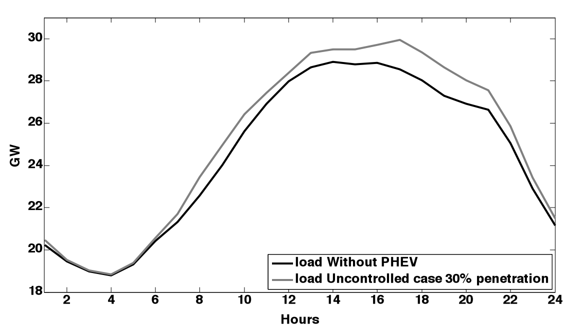


Figure 8: Estimated load on the grid with 30% PHEV fleet penetration

Emissions impact

Results show that the high penetration of coal plants in Ohio leads to in an increase of total SO_2 emissions, a decrease of total CO_2 emissions, and minimal variation of NO_x emissions. Figure 9 shows PHEV use can reduce CO_2 emissions by about 25% on an annual basis in an uncontrolled charging case, while benefits are negligible for the controlled charging scenario.

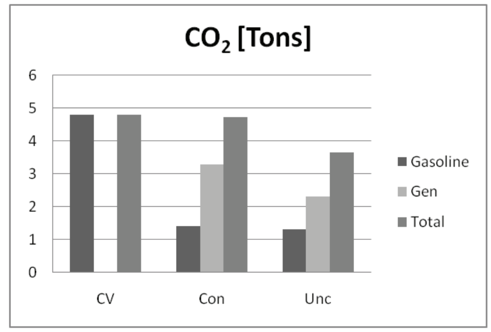


Figure 9: Annual per-vehicle CO2 emissions

Annual NO_x emissions will increase around 5 kg in both cases because emissions connected to power generation are higher than the emissions reductions from reduced gasoline use, as shown in Figure 10. Figure 11 further shows that due to the high penetrations of coal and heavy oil as generation fuels in Ohio, annual per-vehicle SO_2 emissions will increase by between 10 and 12 kg with PHEV use.

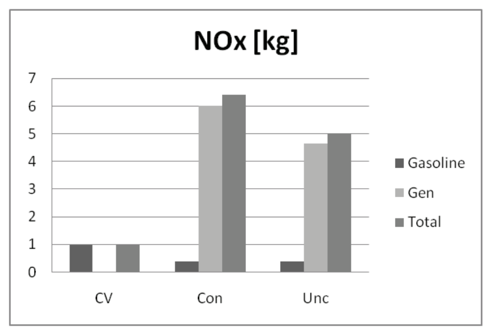


Figure 10: Annual per-vehicle NOx emissions

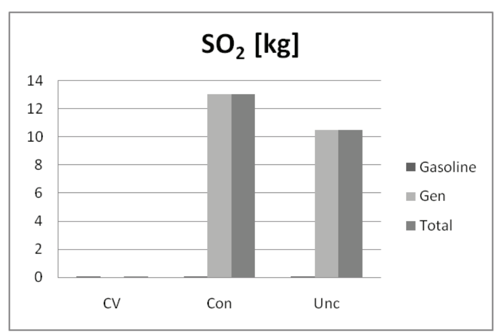


Figure 11: Annual per-vehicle SO2 emissions

Ownership cost

The architecture of a PHEV is differentiated from that of an HEV by its ability to further displace fuel usage by charging off-board electrical energy from the electric utility grid while not being driven. To accommodate the increased dependence on electric power while maintaining an appropriate vehicle weight, the PHEV uses a battery pack with a larger capacity and a smaller internal combustion engine and fuel tank. Also, an inverter-integrated charging plug is needed to connect the enhanced battery pack to a standard electrical socket for recharging purposes. Since all PHEVs produced in the future are expected to offer at least 220 V charging capability, the estimated cost to install an additional outlet in an accessible location to the PHEV is included in this cost analysis.

Currently, conventional vehicles exhibit the least-expensive initial cost of about $21,390, which is not expected to vary significantly in the future (in terms of real cost, excluding inflation). PHEVs will experience a more dramatic cost reduction from current cost estimates of $51,388 to about $27,668. With these cost reductions, PHEVs are expected to have a price premium of approximately $6,200 in the future as compared to conventional vehicles [14]. Our analysis of vehicle operating costs focuses on the cost of fuels—gasoline and electricity—used in driving PHEVs and compares this to the driving costs of CVs. PHEVs will clearly have lower gasoline costs, due to the use of electricity. The average daily trip of drivers considered in the model is about 38 miles and the PHEVs considered have an all-electric range of approximately 22 miles if they operate using the electric vehicle control strategy assumed. Figure 12 shows that gasoline savings can reach more than 70% and one single PHEV will reduce annual gasoline consumption by an average of around 385 gallons compared to a CV.

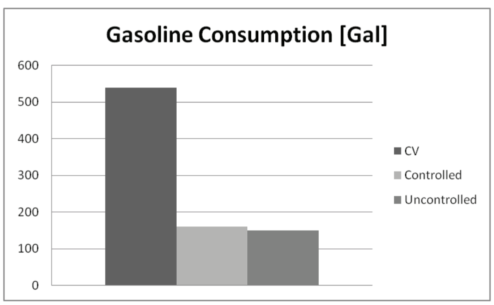


Figure 12: Average annual per-vehicle gasoline consumption (gallons)

Although fuel consumption is reduced through the use of PHEVs, there will be an increase in electricity demand. Retail residential electricity rates averaged about $0.111/kWh in Ohio in 2007, based on values reported by the EIA [14]. While the cost of energy contributes to determining this retail rate, there are non-energy related charges, such as fixed cost recovery, transmission and distribution costs, and metering that are included as well. Our simulations show that adding the PHEV charging loads increases the annual cost of generation by about $0.0032/kWh and $0.0024/kWh in the uncontrolled and controlled cases, respectively. Since utilities would need to recover these additional generation costs, it is plausible to assume that retail rates would increase by this amount above the $0.111/kWh average. It is worth noting, however, that policy makers may opt to levy a different rate on PHEV charging for a number of reasons. One is that it may be desirable to charge a lower rate, as an incentive or subsidy for adoption of PHEV technology. Indeed, in a controlled charging scenario, utilities may opt to give preferential rates in exchange for vehicle owners allowing the utility to control charging. On the other hand, it may also be desirable to charge higher prices in order to recover gasoline taxes, which are typically used for road construction and maintenance, which PHEV owners do not pay due to reduced gasoline consumption. For this reason, we do our cost analysis by parameterizing the cost of electricity. Similarly, although the average retail price of gasoline in Ohio in 2007 was $2.27/gallon [16], gasoline prices soared to much higher levels in 2008. Although prices have dropped from the highs seen then, this is largely attributed to the recent global recession with many expecting prices to rebound as economic activity increases.

It is worth noting that this cost analysis focuses on the uncontrolled charging scenario, since this is the most likely scenario when PHEVs first enter the vehicle fleet. Figure 13 summarizes the effect of gasoline and electricity prices on the relative operating cost savings of PHEVs over CVs.

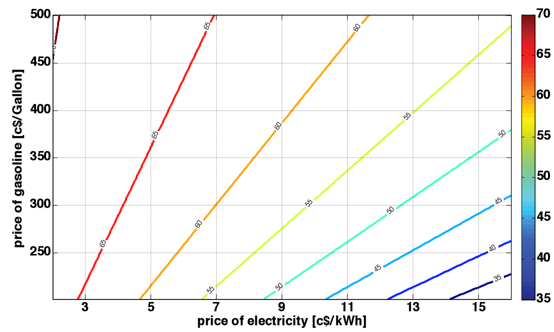


Figure 13: Savings in operation cost [%] of PHEV with respect to CV

Payback time

Payback time is an important factor in determining whether future consumers will decide to buy a PHEV. The basic tradeoff that the consumer faces is the higher upfront capital cost of a PHEV relative to the stream of driving cost savings due to reduced gasoline costs. Figure 14 and Figure 15 summarize the effect of gasoline and electricity prices on the payback time of a PHEV, relative to a CV, with a capital cost difference of $6,000 and $10,000 between the PHEV and CV. As discussed above, this analysis assumes a future scenario in which mass production of PHEVs and their batteries (which is expected to be the major driver of the higher cost of PHEVs) bring the price difference between a PHEV and CV to these levels.

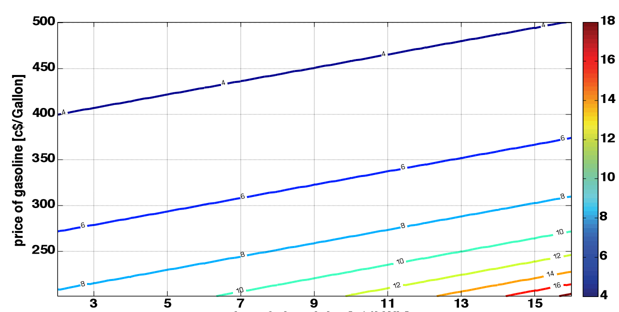


Figure 14: Payback time of a PHEV as a function of gasoline and electricity prices with PHEV costing an additional $6000 compared to a CV

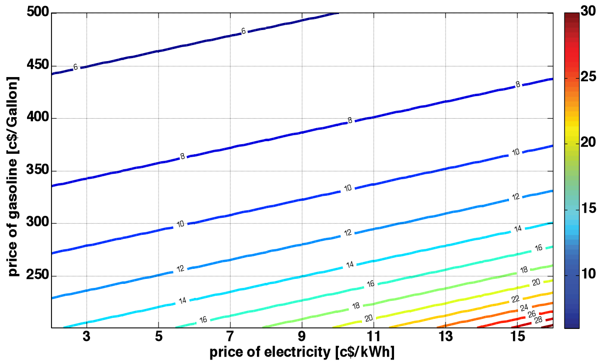


Figure 15: Payback time of a PHEV as a function of gasoline and electricity prices with PHEV costing an additional $10000 compared to a CV

CONCLUSIONS

In this study the Ohio power system has been modeled as a mixed integer program to analyze the impact that a PHEV fleet would have on it. Two different charging scenarios, a controlled charging and an uncontrolled one, have been modeled to estimate effects of the timing of PHEV charging. This analysis has focused on how PHEVs would affect net emissions and costs, and what impacts they would have on the grid load profile.

Results show that in the near future, the uncontrolled charging scenario is preferable—obtaining better gasoline saving and environmental impact. Furthermore, this scenario does not require complex communication and control systems between the power grid operator and vehicles to control the charging process. If in the future PHEVs reach a bigger market penetration, it will be necessary to verify if power grid capacity is sufficient to satisfy higher peak loads during afternoons and if transportation and distribution infrastructures will be able to transport energy from power plants to where vehicles will be plugged-in. Results also show how PHEVs could lead to substantial reductions in gasoline consumption of close to 70% relative to CVs, providing a initial step towards transportation electrification. Since electricity is less expensive than gasoline, the additional investment required to buy a PHEV could be repaid by great savings in operation costs. Nevertheless, federal and state incentives could be necessary to help PHEVs gain a foothold in the market due to their long payback-time.

Looking at environmental impact it is clear that even with a high penetration of coal in the Ohio power system, it could be possible to decrease greenhouse gasses (CO_2 emitted by a single vehicle in one year decreases by about 25%), as shown in other studies. Using cleaner sources of energy for electricity production will improve this benefit (for example, an analysis of PHEVs in Colorado showed a 40% reduction in per-vehicle emissions of CO_2 [2]) and avoid increases in SO_2 and NO_x emissions.

Future studies will have to focuses on smaller scale cases to verify what influence PHEVs will have on local infrastructures and what benefits could be achieved by vehicle to grid and other services. These types of services could also further increase the economic justification of a potential PHEV purchase. Other important aspects to be analyzed are how PHEV use can help support renewable integration and the effects of increased electricity demand due to PHEV charging on electricity prices.

REFERENCES

[1] Electric Power Research Institute, Environmental Assesment of Plug-In Hybrid Electric Vehicles, Volume 1: Nationwide Greenhouse Gas Emissions; EPRI: Palo Alto, CA, 1015325, 2007.

[2] K. Parks; P. Denholm; T. Markel, Costs and Emissions Associated with Plug-In Hybrid Electric Vehicle Charging in the Xcel Energy Colorado Service Territory; Technical Report **NREL/TP-640-41410, 2007**.

[3] C.H. Stephan; J. Sullivan, Environmental and Energy Implications of Plug-In Hybrid-Electric Vehicles. **Environ. Sci. Technol. 2008, 42, 1185–1190**

[4] R. Sioshansi, P.Denholm. Emissions Impacts and Benefits of Plug-In Hybrid Electric Vehicles and Vehicle to Grid Services**. Environmental Science and Technology, Vol. 43, No 4, pp 1199-1208, February, 2009.**

[5] C. Samaras; K. Meisterling, Life Cycle Assessment of Greenhouse Gas Emissions from Plug-in Hybrid Vehicles: Implications for Policy. **Environmental Science Technology 2008, Vol. 42, pp 3170–3176**.

[6] V. Marano; G. Rizzoni, "Energy and Economic Evaluation of PHEVs and their Interaction with Renewable Energy Sources and the Power Grid," **2008 IEEE International Conference on Vehicular Electronics. and Safety**, Columbus, OH, US, September 2008

[7] International Energy Agency (IEA). World Energy Outlook, 2006.

[8] P. Denholm; S.E. Letendre, Grid Services From Plug-In Hybrid Electric Vehicles: A Key to Economic Viability? **Electrical Energy Storage; Applications and Technology (EESAT) Conference,** 2007

[9] U.S. Department of Transportation's Federal Highway Administration www.fhwa.dot.gov.

[10] Environmental Protection Agency (EPA), www.epa.gov.

[11] Energy Information Administration (EIA), Monthly Cost and Quality of Fuels for Electric Plants Data.

[12] PJM, www.pjm.com.

[13] National Research Council, Transition to alternative transportation technologies: Plug-in Hybrid Electric Vehicles

[14] K. R. Genung, C. Lawrence, T. Markel, C. S. Hinds, R. McGill, R. E. Ziegler, D. E. Smith., D. Brooks. S. W. Hadley, R. L. Smith, D.L. Greene, H. Wiegman. and V. Marano, *Plug-in Hybrid Electric Vehicle, Value Proposition Study, phase 1, task 3: Technical requirements and procedure for evaluation of one scenario*, **tech. rep., ORNL under contract No. DE-AC05-0OR22725, January 2009.Final Report**.

[15] Energy Information Administration (EIA), Average Retail Price of Electricity to Ultimate Customers by End-Use Sector, by State.

[16] Energy Information Administration (EIA), Retail Gasoline Historical Prices.

Strategic Planning for Energy: Too Little Too Late?

Position paper

Christopher Townley and Joe Howe*

Abstract
Strategic planning for energy is an area that has required work for at least thirty years now. This follows the oil shocks during the 1970's which resulted in economic recession and inflation, and served to demonstrate how vulnerable 'Western' or '1st world' lifestyles were, and still are, to disruptions in a consistent supply of this resource. The oil shocks provided a flavour of what is to come in the event of a permanent decline in non-renewable resources. 'Too little too late' refers to the where and how much of the monetary revenue, which has been generated in both the private and the public sector from the exploitation of our non-renewable resource base to date, has been invested in preparing for when we no longer have it. Technical and economic tools, measures and initiatives have been researched, created, assessed, and some of them have been implemented. However, the challenges with respect to non-renewable resources depletion and unsustainable patterns of production and consumption persist, which is an indication that strategic planning for energy has not been optimal. This paper aims at improving strategic planning for energy by suggesting that it should include a psychological component in order to complement the technical, economical and geopolitical elements, which are insufficient to understand and integrate the behavioural dimension of the changes required to reach a sustainable energy future. Little attention has been given to the psychological elements so far, which makes research on this subject not only original but also a useful complement to the existing literature.

Keywords: Strategic planning for energy; Energy security; Sustainability; Production and consumption patterns; Natural resources depletion.

Strategic planning for energy: What is at stake?

Strategic planning is 'the process of determining long-term goals and then identifying the best approach for achieving those goals' (InvestorWords, 2010). During strategic planning, the political, technical, economic and social elements of an activity are supposed to be carefully considered and to reflect driving forces in the environment, for example increasing competition and changing demographics (FreeManagementLibrary, 2010). For energy, strategic planning implies long-term movements and safeguards for involved stakeholders and energy systems. In other words, strategic planning for energy should determine long-term goals for our energy future as well as the most relevant approach to reach them, given long-term energy projections, patterns of production and consumption, limited resources, growing population, and a fragile environment (Edmonds et al, 1985). Energy planning should also be 'consistent with overall goals of national sustainable development' (Sokolov and Beatty, 2009).

The oil shocks during the 1970's, which resulted in economic depression and inflation, demonstrated how vulnerable 'Western' or '1st world' lifestyles were (and still are) to disruptions in a consistent supply of, and stable market prices for, natural resources (in this case, oil). Changes in the price of oil rippled through the economy and increased the cost of industrial and agricultural processes and production, ultimately reflecting in increased end prices of commodities. This resulted in a reduction in demand for consumer products and the preceding industrial output, which created rising unemployment, increases in personal debt, and reductions or constraints in consumer choice with respect to spending habits. In terms of strategic planning, this would have been an ideal opportunity to define new goals, to develop new approaches, and to implement restructuring measures, including:

- Educational and communication reforms, i.e. preparing children for a future with limited fossil fuel resources and creating public awareness about what is coming and how it will influence decision with regard to ways of living;
- Fiscal reform, i.e. implementing new taxes to internalise the external costs associated with burning fossil fuels, which in turn make investment in research and development (R&D) for renewable energy provision and their installation more economically favourable;
- Infrastructural reforms, i.e. improving existing housing stock and the design of any new build homes for energy conservation purposes.

Measures like these would have helped isolating ourselves from interruptions in supply and mitigating ourselves against the affects of a depleting non-renewable resource stock. They would have begun transforming our institutions and engaging us individually and collectively in society, at a time when the wound was fresh and the motive to take action at its greatest. However, these measures were not (or not efficiently) implemented, although warning signals and rational for the affect

* Centre for Sustainable Development, University of Central Lancashire, UK. Contact information: CTownley@uclan.ac.uk

of long-term dependence on fossil fuels were in abundance. For example, Jimmy Carter, U.S. president at the time, was vocal and conscientious of this, stressing 'the need to balance our demand for energy with our rapidly shrinking resources' (Carter, 1977)[1]. Similarly, Donella Meadows and the Club of Rome (1972) built on earlier work by Thomas Malthus and explained the effects of a rising population and limits to growth with respect to pollution, waste, resource depletion and industrialization. Furthermore, they explained how this ultimately undermines the make-up of the 'Western' lifestyles and significantly compromises future sustainability (Meadows et al., 1972).

Today, population and economic growth, affluence, as well as limited resources and a fragile environment are all integrated dynamics. Growth of the population and rising living standards cannot continue without the wealth of natural resources we have enjoyed to date. These challenges were presented to us during the 1970's and continue today, but they are more complex to solve as more people will be affected, and we have less time and resources available to organise ourselves, draw up a sound plan of work and take action in reducing our energy consumption, and providing ourselves with alternatives. The world population has increased to over six billion and, if left unrestrained, it is projected to exceed approximately 9 billion by 2080 (UN, 2004). At the same time as the population has grown, human development (life expectancy, education, standards of living) has increased, particularly in the 'South' (UN, 2009). However, the production and extraction rates of oil and natural gas, coal, and primary mineral ores are all exponential and cannot be sustained. We are currently presented with the problem of a declining natural resource base coupled with lack of investor confidence into alternatives, and economic market recession, which will only deteriorate if patterns of unsustainable production and consumption remain the same (Freidman, 2008).

It is clear that the challenges with increased commodity price rises, inflation, and their effects are beginning to appear again today. It is occurring much more slowly this time, because it is not due to deliberate interruptions in supply like in the 1970's but to a permanent decline in the quality and global production, not only of oil, but also natural gas, coal, as well as many other mineral ores (IEA, 2008; Vernon, 2007). In addition, geopolitical challenges surrounding the distribution of remaining non-renewable resources persist and consume significant public resources.

In sum, strategic planning for energy appears to have been inefficient in preparing our energy future to the coming challenges. Consequently, the aim of this position paper is to discuss the way to improve strategic planning for energy, by suggesting that it includes a psychological component that complements the technical, economical and geopolitical elements, which are insufficient to understand and integrate the behavioural dimension of the changes required to reach a sustainable energy future.

The balance between technological measures and behavioural changes

The challenges with respect to strategic planning for energy involve practical, technical, political, social and economic forces, which are intimately linked. They were presented to us by the book *The limits to growth* (Meadows et al., 1972) and, during the 1970's oil crisis, the 'Western' world was given a flavour of what can happen when resources are constrained. Since then, some initiatives have been taken in order to prepare for our energy future, taking into account the prospect of limited non-renewable energy supplies as well as climate change. Most of them have to do with sustainable technologies.

Denmark is now a world leader in energy conservation and combined heat and power systems (CHP). Today, 46% of the country's heat and 23% of its electricity demand is supplied through distributed grids using natural gas and/or biomass (straw) as the raw material in CHP plants (Lund, 2009). This model will be useful applying elsewhere; it can be tailored to meet electricity and heat demand, and even provide export revenue, as is the case with supply to Germany (Hammons, 2008). In Sweden, 18% of the energy is produced using CHP and biomass gasification, which not only provides electricity and heat, but also biofuel (Börjesson, 2010). Over forty countries currently employ CHP, which can also be employed and developed in the 'South' (IEA, 2008). At present, it suffers high investment costs and is not seen as economically attractive (Keppo, 2006). This is due to perverse market distortions and the absence of taxes, which need applying to fossil fuels to account for environmental degradation and health costs associated with their use (Knutsson, 2006), and which could be invested in renewable energy.

The USA and China have recently taken the lead over Germany and Denmark in the production and implementation of wind power; lifecycle analysis of this technology demonstrates an energy use pay back of less than 2% of the life of the turbine or about 5 months (Mostafaeipour, 2010). Wind turbine industry has grown exponentially since the 1990's, assisted by the use of feed in tariffs, which are an economic tool introduced by Germany and go some way in balancing the under-priced fossil fuel alternatives (Sijm, 2002; Mitchell, 2006).

Solar and photovoltaic systems have grown in popularity in the USA, Germany and Japan (Makrides, 2010). The lifecycle balance, or payback on energy use and embedded C02 emissions, is within approximately four years. This is expected to decrease as C02 emitted during production of units is expected to decrease over by 40% in the next five years and may be particularly good for the 'South' due to favourable climatic conditions (Varun, 2009).

Geothermal is a resource largely used by China, Sweden, the USA, Turkey and Iceland (Lund, 2007). There are over 72 countries utilizing the resource for producing electricity and space heating or cooling and it has experienced annual growth in more recent years, as countries have conducted the necessary groundwork to quantify their indigenous resources in preparation for rising fossil fuel prices and improving economic viability (Lund, 2000). Large-scale electricity generating projects suffer high capital costs for drilling and exploration. However, the economics will improve, as fossil fuel prices rise. In addition, geothermal can be coupled with CHP to improve efficiencies and economics (Lund, 2007).

Hydropower is employed all over the world and will probably outlast most currently known energy sources (Sternberg, 2010). Small systems can be connected to a distributed grid for domestic or community scale supply of which Turkey serves as a good example of a country presently liberalising its energy markets and attempting to increase the share of renewable, in particular hydropower and the use of distributed systems (Bakis, 2006). Once costs for transmission and environmental externalities are included in economic analysis, capital investment is offset by its long life and low running costs, and it becomes economically competitive (Lise, 2009).

For transportation, a combination of biofuel and electric power is viewed as a way forward due to the positive lifecycle results, the wide range of feedstocks and processes, and its compatibility with existing infrastructure (see Demirbas, 1009; Agarwal, 2006; NREL, 1998; Lammert, 2009)[2].

Despite investments in renewable energy, fossil fuels still have an 81% share in global energy supply, followed by 13% in renewable energy and 6% in nuclear (IEA, 2008). In fact, non-renewable resources have been too inexpensive for too long a time and the privately or publicly generated revenue has been ill invested and/or what has been invested has not made a significant real world difference. As long as renewable energy is more expensive than fossil fuel, the population will choose the cheapest option (Sovacool, 2009; Friedman, 2008), as the whole life cycle is not easily taken into consideration by public (let alone private) economic actors. Changing economic actors' behaviour will require incentive measures, such as taxing 'offending substances so that their prices are high enough to prevent demand from exceeding a politically decided level' and personal carbon caps, which 'forbid consumption above this level' (Alcott, 2009). Taxing non-renewable energy is the most direct approach; it is 'simpler, more transparent, easier to calculate and cut across the whole economy' (Friedman, 2008). It reflects in price signals, which people will recognise and respond to. Indeed, pricing signals are very influential to consumers' behaviour and their choices. However, this will affect people's way of life and understanding their attitude towards this is likely to be the key towards efficient strategic planning.

In addition to technical solutions and incentive measures, complementary conservation initiatives have been devised and promoted, which further require changes in our way of life. For example, a change to a diet containing less meat will lead to an immediate reduction in energy use in the agricultural sector. Indeed, energy is currently used to produce fertilizers for growing animal feed crops (corn, grains and soybeans), operating machinery, transportation, processing and refrigeration of which these are only several of the stages during intensive agricultural and cattle operations (Vuuren et al, 2009; Marlow et al., 2009; McMichael et al, 2007)[3]. Reductions in private car ownership and the frequency of which motor cars are used, as well as car-pooling, will also immediately reduce energy consumption (Gasparatos, 2009). Local employment, recreation, purchasing habits and less urban sprawl are other ways to reduce energy consumption (Gasparatos, 2009; Holm, 2009).

The trend of heavy reliance on energy derived from fossil fuels cannot continue due to physical/geological limits on reserves, as well as to geopolitical and environmental constraints on their production or use (Hallock, 2004). Eventually, production and consumption (and growth) will not be possible once we have exhausted this natural resource base and reached capacity for polluting our environment (Krysiak, 2006); this can be expected to deteriorate (Friedman, 2008; Conway, 2010). Implementing measures to limit our dependence on fossil fuel energy will require more than technical solutions. Taxing non-renewable energy and promoting behavioural conservation measures will engender a change in people's way of life. Our alternatives will probably not provide us with the excesses, which fossil fuels have to date, which is another reason why changing our behaviour is important.

The psychological component of strategic planning for energy

Measures to improve how efficiently we use resources have been implemented on an ongoing basis, and are usually put into context with reducing C02 emissions, energy security and meeting government targets (Barret et al, 2008). However, these measures have not succeeded in reducing energy consumption and the motives do not appear to have had much of an effect. In some cases, they even lead to increases (Holm, 2009; Brannlund et al., 2007). Improvements in efficiency have suffered from a 'rebound effect', and have been undermined by growth in population, increases in affluence or living standards and globalized trade. This trend has been evident since 1980 (Khazzoom, 1980). It shows no sign of changing and it will likely continue if left unchecked (Brannlund et al, 2007).

Energy conservation, renewable expansion and changes in production and consumption patterns have been prescribed in the literature and public policy discourse for a long while and it is prescribed in studies for future forecasts (Dti, 2007). However, little has changed. In fact, very few people have immediate and pressing motives to undertake one or many of the initiatives aimed at reducing energy use, and therefore change their behaviour, habits, routines and way of life. Linden et al. (2006) state that policies to improve energy efficiency at home are designed top down, and 'although they can achieve immediate results, they seldom inspire long term reconsideration of behaviour and routines'. However, it is essential that we achieve this, so as to avoid undermining progress and successfully obtain buy-in from the public. In other words, implementing reforms towards more sustainable energy future requires prompting behavioural changes. Therefore, in addition to considering whether we have the time, money, and natural resources for investing in the development, production, installation and construction of the necessary infrastructure and conservation initiatives for a different type of energy system, we must consider the psychological impact and what is required of people who will actually be obliged to live in these new systems.

In addition, research on this subject has shown the limitations of using traditional approaches to understand and unravel the complex interactions between the social reality and natural phenomena. We think that the explanation of the failure to solve the global energy crisis and to move towards more sustainable use of energy is also to be searched in the way research programmes are conceived and, in turn, serve as a basis for action and strategic planning. Critical and constructivist approaches have brought new elements to the understanding of the interactions between political, economic and social systems in relation to the environment. Political Ecology has particularly shown its usefulness in highlighting the politics behind environmental decision-making and in drawing attention to inequality issues (Bryant & Bailey). We think that psychological elements have to be integrated as well, to understand both the way the human/natural system is conceived and the adjustments people are ready to make. More works need to be done to complement research on these aspects (see notably Devine-Wright, 2006; Owens & Driffill, 2008).

What adjustment will the population need to make? How will it change people's way of life? Will they receive help or support? What happens if they do not change? How long will it take them? These questions need exploring. Attempting to transition from a lifestyle, which the past two to three generations of people have become familiar with and used to, to one, which they have never experienced before and probably find difficult to imagine, is challenging. Unless it is done in a coordinated fashion by well-informed people with at least loosely planned schedule of work, conflict, resentment, fear and apprehension of the unknown may all serve to undermine any attempts at making a feasible plan and range of solutions. Change does not always happen easily and may not be very comfortable to begin with due to deeply engrained habits, and it is very important that any changes required be set in unambiguous context (Marechal, 2010). The context in which change is set will influence people's attitudes, and any changes will affect different members of society in different ways, potentially marginalising poorer or weaker parts of society (Barrett et al, 2008). The burden may disproportionally fall upon certain groups (Carlsson-Kanyama, 2007). All of these elements will be important in determining individual and collective behaviours, adjustment of habits and approaches to change, and therefore, the result of the transition away from a fossil based energy system.

This issue becomes ever more pressing when we consider that, in addition to changing present ways of life in the 'North', where the belief that 'more is better than less' is prominent, many people in the 'South' (where population is rising the quickest) aspire to live like this, while the development regime is modelled on the unsustainable practices of the North. Their way of life will also be required to change, which may be more difficult now that they have developed or are developing a taste for this way of life consumerism.

In conclusion, our position paper has shown that behavioural changes are required along with technical solutions if we want to prepare for a future with limited fossil fuel energy. Implementing behavioural changes reflects back into strategic planning for energy (notably R&D) and our institutions (for example education). Thus, to improve strategic planning for energy, we propose to include a psychological component along with the economic, technical and geopolitical ones. By doing so, strategic planners will be able to take people's aspiration, needs, and fear into account while drawing strategies and policies for our energy future. There is here, potential for a fruitful and exciting seam of research and we encourage researchers on energy issues, from any disciplines, to integrate psychological and behavioural elements into their framework of analysis.

References:

Agarwal, 2007. Biofuels (alcohols and biodiesel) applications as fuels for internal combustion engines. Progress in Energy and Combustion Science, Vol 33, Iss 3, June 2007, Pages 233-271.

Alcott, 2009. Impact caps: why population, affluence and technology strategies shouldábe abandoned. Journal of cleaner production, Vol 18, Iss 6.

Aune, 2007. Energy comes home. Energy policy, Vol35, Iss 11.

Bakis, R., 2007. Electricity production opportunities from multipurpose dams (case study). Renewable Energy, Vol 32, Iss 10.

Barret et al, 2010. How to support growth with less energy. Energy Policy Vol. 36.

Blanco, H, 2010. Hot, congested, crowded and diverse: Emerging research agendas in planning. Progress in planning, Vol 71.

Borjesson,M, 2010. Biomass gasification in cost-optimized district heating systems-A regional modelling analysis. Energy Policy, Vol 38, Iss 1.

Brannlund, R, 2007. Increased energy efficiency and the rebound effect: Effects on consumption and emissions. Energy Economics, Vol 29.

Bryant, Raymond L. and Sinead Bailey. 1997. Third World Political Ecology. London: Routledge.

Carlsson-Kanyama, A.-L, 2007. Energy efficiency in residences - Challenges for women and men in the North Energy Policy, Vol 35.

Carter, Jimmy, 1977. Primary Sources: The President's Proposed Energy Policy. Retrieved at http://www.pbs.org/wgbh/amex/carter/filmmore/ps_energy.html, May, 2010.

Conway,Edmund, 2010. City fears of 'Great Depression Mark II'. Retrieved at http://uk.finance.yahoo.com/news/city-fears-of-great-depression-mark-ii-tele-12d592c21529.html, May, 2010.

Demirbas, M. F, 2009. Biorefineries for biofuel upgrading: A critical review. Applied energy, Vol 86.

Devine-Wright, P., 2006. Energy citizenship: psychological aspects of evolution in sustainable energy technologies. In Murphy , J. (ed.), Governance of Sustainable Technologies. London: Earthscan.

Dti, 2007. Meeting the Energy Challenge. A White Paper on Energy. Published by TSO, 2007.

Edmonds, Jae and John M. Reilly, 1985. Global energy: Assessing the future, Oxford University Press, New York, 1985.

FAO, 2006. Livestocks long shadow. Retrieved at ftp://ftp.fao.org/docrep/fao/010/a0701e/a0701e00.pdf, May 2010.

FreeManagementLibrary 2010. Strategic Planning. Retrieved at http://managementhelp.org/plan_dec/str_plan/str_plan.htm, May, 2010.

Friedman, Thomas, L, 2008. Hot flat and crowded. Published by Farrar, Straus and Giroux; 1 edition.

Gasparatos,A, 2009. A longitudinal analysis of the UK transport sector, 1970-2010. Energy policy, Vol 37, Iss 2.

Hallock, 2004. Forecasting the limits to the availability and diversity of global conventional oil supply. Energy Vol. 29.

Hammons, T, J, 2008. Integrating renewable energy sources into European grids. Electrical power and energy systems, Vol 30, Iss 8, Pages: 462-475.

Holm, S,O, 2009. Increased ecoefficiency and gross rebound effect: Evidence from USA and six European countries 1960-2002. Ecological economics, Vol 68, Iss 3.

http://awellfedworld.org/sites/awellfedworld.org/files/pdf/UNGlobalWarmingReport.pdf, May, 2010.

http://hdr.undp.org/en/media/HDR_2009_EN_Complete.pdf, May 2010.

http://www.un.org/esa/population/publications/longrange2/WorldPop2300final.pdf, May 2010.

Hubbert, M, King, 1956. Nuclear Energy and the Fossil Fuels. Drilling and production practices, American petroleum institute.

Hurst, 2010. Canadian Cement Plant Becomes First to Capture CO2 in Algae. Retrieved at http://earthandindustry.com/2010/03/canadian-cement-plant-becomes-first-to-capture-co2-in-algae/, May 2010.

IEA, 2008. Combined heat and power; Evaluating the benefits of greater global investment. Retrieved at http://www.localpower.org/documents/reporto_iea_chpwademodel.pdf, May 2010.

IEA, 2008. World Energy Outlook. Retrieved at http://www.worldenergyoutlook.org/docs/weo2008/WEO2008_es_english.pdf, May, 2010.

InvestorWords, 2010. Strategic Planning. Retrieved at http://www.investorwords.com/4774/strategic_planning.html, May 2010.

Keppo, 2006. Economic appraisal of small biofuel fired CHP plants. Energy conservation and management, Vol 48, Iss 4.

Khazzoom, D. J. (1980). “Economic implications of mandated efficiency standards for household appliances.” Energy Journal, 1(4): 21-39.Knuttson, David, 2006. Combined heat and power in the Swedish district heating sector--impact of green certificates and CO2 trading on new investments. Energy Policy, Vol 34, Iss 18.

Krysiak,F.C, 2006. Entropy, limits to growth, and the prospects for weak sustainability. Ecological economics, Vol 58, Iss 1.

Lammert, 2009. Twelve-Month Evaluation of UPS Diesel Hybrid Electric Delivery Vans. U.S. Department of Energy, National renewable energy laboratory.

Li Xianguo, 2005, Diversification and localization of energy systems for sustainable development and energy security, Energy Policy, Vol.33, pp.2237-2243.

Linden et al. (2006). Efficient and inefficient aspects of residential energy behaviour: What are the policy instruments for change? Volume 34, Issue 14.

Lise, W, 2009. Towards a higher share of distributed generation in Turkey. Energy Policy, Vol 37, Iss 11.

Lund, H, 2007. Renewable energy strategies for sustainable development. Energy, Vol 32, Issue 6.

Lund, H, 2010. The role of district heating in future renewable energy systems. Energy, Vol. 35, Issue 3.
Lund, John, 2007. Characteristics, development and utilization of geothermal resources. Retrieved at http://geoheat.oit.edu/bulletin/bull28-2/art1.pdf, May 2010.
Lund 2009. Energy system analysis of 100% renewable energy systems—The case of Denmark in years 2030 and 2050. Energy, Volume 34, Issue 5.
Makrides, 2010. Potential of photovoltaic systems in countries with high solar irradiation. Renewable & Sustainable Energy Reviews, Vol 14, Iss 2.
Marlow et al, 2009. Diet and the environment: does what you eat matter? vegetarian diet and the environment, 2009.
Marechal, K, 2010. Not irrational but habitual: The importance of "behavioural lock-in" in energy consumption. Ecological Economics. Vol 69, Iss 5.
McMichael et al, 2007. Food, livestock production, energy, climate change, and health. Lancet, Vol 370.
McNamara, C, 2010. Strategic Plannning. Retrieved at http://managementhelp.org/plan_dec/str_plan/str_plan.htm, May 2011.
Meadows et al, 1972. The limits to growth. Published by Universe books.
Mitchell, C, 2006. Effectiveness through risk reduction: a comparison of the renewable obligation in England and Wales and the feed-in system in Germany. Energy Policy, Vol 34, Iss 3.
Mostafaeipour,A, 2010. Productivity and development issues of global wind turbine industry. Renewable & Sustainable Energy Reviews, Vol 14, Iss 3.
Noon, D. Price Signals and Consumers' Energy Conservation Behavior: The Case of Responsiveness in Residential HVAC Technologies from 1992-2004. Retrieved at http://www.prism.gatech.edu/~dn56/HVAC.JCPA.pdf, May 2010.
Owens, Susan and Driffill, Louise, 2008. How to change attitudes and behaviours in the context of energy. Energy Policy, 36, pp. 4412-4418.
Pimentel, D, 2003. Sustainability of meat-based and plant-based diets and the environment. American Journal of Clinical Nutrition, Vol. 78, Iss 3.
Sheehan et al, 1998. Life Cycle Inventory of Biodiesel and Petroleum Diesel. U.S. Department of Energy, National renewable energy laboratory.
Sijm, J.P.M, 2002. The Performance of Feed-in Tariffs to Promote Renewable Electricity in European Countries. Retrieved at http://www.ecn.nl/docs/library/report/2002/c02083.pdf, May 2010.
Sokolov & Beatty 2009. Sustainable Nuclear Energy. Retrieved at http://www.iaea.org/Publications/Magazines/Bulletin/Bull511/51104023942.pdf, May 2010.
Sovacool, B, 2010. Rejecting renewables: The socio-technical impediments to renewable electricity in the United States. Energy Policy, Vol 37, Iss, 11.
Sternberg 2010. Hydropower's future, the environment, and global electricity systems. Renewable and Sustainable Energy Reviews Vol. 14, Iss 2.
Thanh,Le Tu, 2010. A two-step continuous ultrasound assisted production of biodiesel fuel from waste cooking oils: a practical and economical approach to produce high quality biodiesel fuel. Bioresource Technology, Vol 101, Iss 14.
United Nations, 2004. World population to 2300. Retrieved at
United Nations, 2009. Overcoming barriers: Himan mobility and development. Retrieved at
Varun, 2009. Energy, economics and environmental impacts of renewable energy systems. Renewable & Sustainable Energy Reviews. Vol 13, Iss 9.
Varun, 2009. LCA of renewable energy for electricity generation systems-A review. Renewable & Sustainable Energy Reviews. Vol 13, Iss 5.
Vernon, 2007. Peak Minerals. Retrieved at http://www.theoildrum.com/node/3086, May, 2010.
Vuuren et al, 2009. Future bio-energy potential under various natural constraints. Energy Policy Vol 37.
Yuri Sokolov and Randy Beatty, Sustainable nuclear Energy, IAEA bulletin 51 1 September 2009 PP.39-42.

Endnotes

[1] In 1956, M. King Hubbert, a geoscientist working at the Shell oil corporation, predicted that the US oil production peak would be in 1970 followed by a global oil production peak in 1995 (Hubbert, 1956). His first estimate was remarkably accurate and the second was only delayed by the oil crisis and Gulf war until 2005.

[2] The use of these renewable energy sources may not be without negative impacts on the environment, though (Li, 2005).

[3] In fact, the intensity with which energy has been used has coincided with a rise in standards of living, which generally have diets containing more meat, and increases in population, and it is a trend that cannot continue. Alternative diets can provide the necessary daily calorific intake (Pimentel, 2010), and a transition from centralized monoculture practices and globalized trade, to decentralized operations and the re-integration of crop and livestock activities (FAO, 2006) will reduce overall energy consumption and the impact agriculture has in other areas.

AUTHORS INDEX
Volume I