

The use of social network analysis software to analyze communication patterns and interaction in online collaborative environments

Amy MAZUR

Department of Teacher Preparation and Special Education, The George Washington University
Washington, DC 20052, USA

and

Chad DORAN

Information Group, Johns Hopkins University Applied Physics Laboratory
Laurel, MD 20723, USA

and

Patricia RICE DORAN

Department of Teacher Preparation and Special Education, The George Washington University
Washington, DC 20052, USA

Abstract

Social network analysis software such as NodeXL has been used to describe participation and interaction in numerous social networks but has not yet been widely used to examine dynamics in online classes, where participation is frequently required rather than optional and participation patterns may be impacted by instructor requirements as well as participants' intrinsic engagement with the subject matter. Such analysis can be valuable in programs focused on teaching collaborative and communicative skills, including teacher preparation programs, to provide information about instructional practices likely to facilitate student interaction and collaboration across diverse student populations. This exploratory study used NodeXL to visualize students' participation in an online course, with the goal of identifying (1) ways in which NodeXL could be used to describe patterns in participant interaction within an instructional setting and (2) identifying specific patterns in participant interaction among students in this particular course. In this sample, general education teachers demonstrated higher measures of connection and interaction with other participants than did those from specialist (ESOL or special education) backgrounds, and tended to interact more frequently with all participants than the majority of participants from specialist backgrounds. We recommend further research to delineate specific applications of NodeXL within an instructional context, particularly to identify potential patterns in student participation based on variables such as gender, background, cultural and linguistic heritage, prior training and education, and prior experience so that instructors can ensure their practice helps to facilitate student interaction in light of each of these potential variables.

Keywords: social network analysis, online collaboration, teacher education, online learning, visualization

Overview

Without face-to-face interaction, learning in online settings occurs instead through a variety of interactions and

experiences: individual readings and activities; asynchronous group discussions; synchronous chats; simulation activities in SecondLife or other virtual reality settings; and social networking [1]. In each of these varied forums, instructors, students and researchers may find themselves identifying new ways to measure student participation, interaction, connectedness with the group, and ultimate mastery of course outcomes [2] [3]. However, research has indicated that it is an ongoing challenge to foster collaboration and interaction among students in online courses [4] [5], as students and instructors must each accept new paradigms for online interaction while also interacting without the benefit of nonverbal, paraverbal and other social cues present in face-to-face environments [6]. For this reason, analysis of online learning must pay particular attention to issues of student engagement and interaction, as students' interactions with others help to determine not only the mastery of educational outcomes but the quality of the student's learning experience [7]. This work is particularly relevant to fields focused on development of interpersonal skills, such as teacher education, as participants in online courses must master collaborative, communicative and interactive skills as part of their efforts to meet professional standards for educational certification [8, 9, 10]. Therefore, collaboration and interaction are not just a means of instruction, but the object and goal of instruction as well.

In considering the role of student interaction in online learning, Slagter and von Tryon address the importance of collaboration and student interaction in constructing a theoretical framework for considering engagement and student "connectedness" in online learning [11]. Built on an overarching principle of constructivism [12], such a framework alludes to the importance student-constructed questions, responses, dialogues, and activities. This research is grounded in a similar belief that student interaction is a central indicator for learning within a constructivist paradigm, particularly one focused on a developmental context for interpersonal and social skills [13]. We refer also to another important construct in theorizing online interaction, social cognition theory, which

holds that an individual's learning experience will be shaped at least in part by the way in which he or she is able to process interpersonal interactions and communication [14]. As students encounter a new set of expectations in online learning, their own learning, as well as their perceptions of the online experience, will continue to be shaped by interactions with others, which therefore assume a role in distance learning which is as important, if not more so, as in face-to-face coursework.

In considering student experiences with respect to interaction and collaboration, social network analysis offers a useful way to organize inquiry and analysis. Social network analysis is an increasingly popular means of identifying patterns of interaction among participants in online networks, particularly variables such as degree (the number of members with whom each member interacts), centrality (how important or "central" each member is within the network), and clustering (the ways in which various members form cliques or groups). There are various measures of centrality, including closeness centrality (a measure of the direct or indirect connections between group members), betweenness centrality (a measure of the way in which each member helps to connect other members of the network), and Eigenvector centrality (the degree to which a participant is connected to other active participants) [15, 16, 17, 18]. Prior research has begun to identify social network analysis as an effective means of examining dynamics in the educational arena [19], and particularly as an instructional tool in online learning [20]. Previous research has also examined learning experiences as social networks; Calvino, Fine and Marcello examine social network analysis as a means of analyzing the quality of discourse in the learning setting using an add-on software to the Moodle learning suite [21].

Shea et al considered the dynamics of a social network in university coursework, finding that interaction among participants can be described in terms of qualitative and quantitative measures such as frequency of participation, quality of reflection evident in postings, and depth of analysis [22]. There are numerous means of conducting social network analysis of online course discussions, but the free program NodeXL, produced by Microsoft and compatible with Excel, is one of the most recent. Additionally, it offers significant benefits for researchers in fields other than computer science, as the program can easily calculate measures of participant interaction such as centrality within a network, degree of interaction with other participants, and closeness to other participants [18].

Despite its accessibility and potential as a tool for analysis of online courses, little research has been done, to date, to explore the potential uses of NodeXL in evaluating online coursework, particularly in collaborative disciplines such as teacher preparation, where learning environments must encourage participant interaction, collaboration, and involvement. This article describes an exploratory analysis of participant interactions, using NodeXL, in order to describe collaboration among participants in an asynchronous discussion forum offered through the Blackboard learning system.

Purpose

The purpose of this exploratory study was to explore ways in which NodeXL, a relatively new tool for social network analysis, could be used as an element in evaluation of online learning experiences, particularly in the field of teacher education, where the development of professional communicative and collaborative skills is frequently prioritized. The goal of the study was to identify ways in which this tool could be used to visualize and describe relationships among participants in an online learning forum which was geared toward developing multicultural competence and communication skills in inservice teachers.

Methods

19 participants were enrolled in a continuing education course, offered by a midsize research university in partnership with a local school system, focused on providing teachers with the skills to address issues of cultural and linguistic diversity in the classroom, issues of increasing importance due to changing demographics in the school system. Participants did not receive university credit but did receive relicensure credit from their school system upon successful completion of the university course. Nine participants reported their background as being in general education; seven participants reported their background as special education; two participants reported their background as being English as a Second Language (ESL). Four participants reported their current teaching assignment as being focused primarily or entirely in high school; five reported their background as middle school; eight reported their background as elementary school. The participants completed online activities, using the Blackboard learning system, which included discussion questions for each of six online modules, as well as evaluations of each of six learning modules after each module was completed. Participants interacted with the course instructor/ facilitator, as well as with the program director overseeing the project, throughout the program. Participants were required to post a total of three times in each online discussion forum, once as a primary response to the question and twice in response to other participants. The first online forum was an introductory forum where participants shared information about backgrounds and expectations for the course; subsequent forums focused on content presented during the continuing education modules.

For this pilot study, data was examined from a content-related forum in which participants were asked to read a series of articles on second language acquisition and post a response of at least 200 words summarizing their understanding of the second language acquisition process and sharing strategies which might be effective in meeting the needs of second language learners and students from culturally diverse backgrounds. The program director did not participate in this forum, but the instructor was an active participant. A social network analysis of the forum was conducted using NodeXL in order to determine whether patterns existed in participant interaction based upon professional background or current grade-level teaching assignment. A visualization of the forum was

generated in order to provide a graphic illustration of participant patterns. The visualization represented each participant as a “node” on the graph, with directional lines indicating interaction among participants. In addition, the analysis identified measures of degree (number of other participants with whom each participant interacts), centrality (the relative importance of each participant in the network or community), and betweenness (the degree to which each participant serves as a “link” or “bridge” among other participants) for each participant. [16, 17, 18, 19]. We reviewed each of these measures for individual participants and also reviewed average measures of degree, centrality and betweenness for groups of participants based upon professional background and current grade-level teaching assignment in order to determine whether differences existed among participation in the class for teachers from differing backgrounds. In addition, qualitative data was examined including participants’ evaluation of their learning experiences in the module.

Findings

While the data set examined was not of a significant enough size to allow for generalizable conclusions, data analysis did reveal (1) that NodeXL was a useful tool in visualizing and describing social relationships among online participants and (2) that participants in this particular forum tended to interact in ways that appeared to be correlated with similarity in their professional backgrounds, particularly for those from a generalist, as opposed to specialist, background.

Visualization:

Metrics and visualization examination show that the instructor is the most central individual in the course. While online learning has been viewed as a platform which allows for greater student involvement [13, 21], this particular visualization demonstrates that in some online environments, the instructor remains the most central element, is responsible for initiating much of the communication between students and instructors, and comments upon the majority of online posts. Examination of metrics describing participant relationships also demonstrated this fact. In this visualization, each participant is assigned a “node” or vertex on the graph, with directed lines (edges) between the nodes representing frequency and direction of interactions. Edge width is associated with frequency of interaction among participants. Shape is indicative of participant role; size is indicative of betweenness centrality; shading is indicative of Eigenvector centrality. The visualization is produced first with the instructor in the center and then with the instructor removed, so that relationships can be clearly seen both in light of the instructor’s involvement and independent of the instructor.

Visualizations Key: Circle: special educator. Triangle: ESL teacher. Square: general educator. Sphere: instructor. Size is associated in this graph with betweenness centrality (larger figures are more central), and shading is associated

with Eigenvector centrality (darker figures are more highly connected to other connected members of the class).

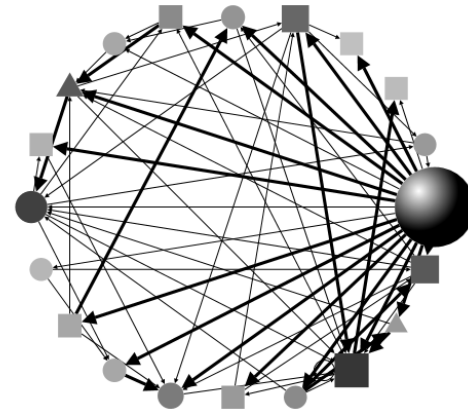


Fig. 1. Visualization of interaction among all participants and instructor, based on disciplinary affiliation. The instructor is the large sphere.

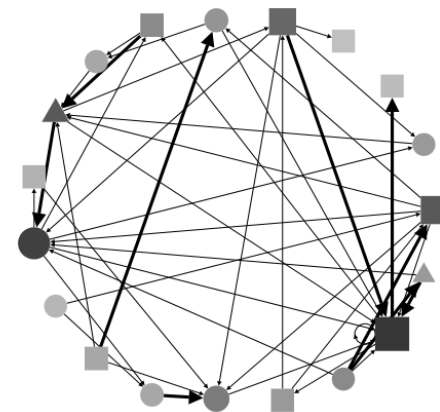


Fig. 2. Visualization of interaction among all participants, without instructor, based on disciplinary affiliation.

Measures of participant interaction:

Degree. The median measure of degree (the number of other participants with whom each participant interacted) was 4.5; the group’s mean was 6.3. The instructor was the highest, with a degree measure of 19, indicating that she interacted with 19 other participants and potentially weighting the group average. Four participants had interactions with only two or three others in the group (including the instructor), indicating that they either did not meet or barely met the minimum standard for participation. Other than the instructor, the two individuals with the highest degree measures (indicating interaction with the highest number of participants) were a general educator (14) and an ESL teacher (11).

Closeness. Closeness centrality is a measure of the distance between each participant and all other participants (i.e., is a

participant connected directly to all other participants, or would information need to travel through a number of other participants in order to reach that individual?). In this network, closeness centrality is impacted in this network by the instructor's extensive involvement with all participants, as the instructor is a key link between a number of participants. For that reason, this exploration did not examine closeness centrality as a measure of participant interaction, as the instructor's extensive interaction with all participants ensured that all participants were connected to some degree by their interaction with her and made it difficult to conduct an examination of participants' closeness relative to one another.

Betweenness. Betweenness centrality, which measures the ways in which each participant is connected to all other participants, is a useful measure in this group of whether, and how, participants have formed connections among themselves, both with the instructor's facilitation and independent of the instructor. Betweenness centrality allows one to quantify how highly connected an individual may be to others in the network, whether through direct or indirect interaction. The mean betweenness score for the group was 0.09; the median betweenness score for the group was: 0.01; the instructor's betweenness score was 1.0. There is a fairly large difference between the instructor's "betweenness" score and the median and mean scores for the group, thereby reinforcing the conclusion that this was an instructor-driven forum and that, independent of the instructor's interaction, participants (with a few exceptions) tended not to form strong connections among themselves.

When betweenness for participants was calculated based on disciplinary affiliation, without the instructor being included in calculations, the following was noted: The betweenness score for special educators was: 0.024. The betweenness score for ESOL teachers was 0.081 (though there were only two representatives in this group). The betweenness score for general educators was 0.046. This data suggests that general educators tended to be more connected to all other participants in the network than did special educators. The two ESOL teachers tended to be most connected to other participants by discussion and interaction.

Eigenvector centrality. Eigenvector centrality is the measure of how much a participant interacts with other "connected" or highly interactive members of the group. In plain language, this might be described as whether a participant interacts with other active group members, or tends to interact with outliers and less active members. Eigenvector centrality can be useful in providing a snapshot of how group members interact, as the Eigenvector scores of different members or subgroups may indicate the degree to which they have formed strong relationships with other influential or active members in the group. In this sample, as with betweenness, Eigenvector centrality varied among those with different disciplinary backgrounds. The mean Eigenvector score for the group was 0.2; the group's median was 0.17. The instructor again had the highest Eigenvector score, with .45. Special educators, as a group,

again had the lowest Eigenvector score, 0.17, indicating a lower average level of centrality than the group as a whole. The two ESL teachers again had the highest average, .232; general educators had a mean of .20.

Discussion

As indicated previously, the significance of this study is limited by the small size of the data set examined. In order to characterize interactions or conduct evaluations of online learning over the course of an entire semester or program, it would be necessary to conduct examinations of multiple discussions, to evaluate multiple courses, and to analyze interactions among a larger number of participants. Such a sample size would also allow for statistical analysis of variance in order to identify which differences had statistical significance.

However, this data provides a demonstration of the ways that NodeXL, a tool which has not yet been extensively used to analyze social interaction and networking in online learning in general or in teacher preparation courses in particular, can offer insight into the structure of group discussions, participant groupings and interactions, and instructor involvement. Based on the findings of this exploratory examination, the following trends may merit further research and investigation:

Instructor involvement. The instructor is deeply involved in a majority of interactions in the course. As Santovec [23] suggests, intensive instructor involvement demonstrates responsiveness to students but may also discourage students from responding to one another. Given the opportunity to provide qualitative feedback about the most helpful element of the first module in allowing them to meet the objective, only two participants identified the discussion forum as most useful. Qualitative feedback from one participant, in particular, identified the discussion as helpful, despite the complexity of multiple discussions and threads, primarily because of the opportunity it afforded to interact with other teachers: "At times, I thought it was a little difficult to keep up with the Blackboard discussion board because there were many different conversations going on at once. Each thread had a different topic being discussed and I had a hard time following/reading all of them. However, I still thought it was useful and helpful to learn what other teachers were thinking and doing in their classrooms. The discussion board is also a great place to post questions to other teachers. I took advantage of the reflective dialogue to find out other options about 'wait time' in the classroom."

Quantitative measures corroborate the perception that there is a wide gap between instructor involvement and importance in the forum and that of participants. In light of other studies, it may be useful to track betweenness and centrality measures for forums with varying degrees of instructor involvement in order to identify whether participants' level of involvement changes as the instructor's responsiveness increases or decreases. A lower level of instructor involvement may also increase the validity of closeness centrality as a measure of student

involvement, as a greater number of interactions will occur among students without direct intervention by the instructor.

Participant engagement and interaction.

As Gunawardena et al suggest [2], students' engagement and interaction with one another is crucial to the successful implementation of collaborative online learning experiences. Particularly in the area of teacher professional development, development of collaborative skills is important not only in the theoretical framework of online learning but also as a required outcome of teacher education courses. In this particular forum, the data suggests that participants from general education backgrounds and ESL backgrounds tended to be more deeply involved with other participants in the forum than did those from special education backgrounds. This may be due to the structure of the discussion question, which asked participants to summarize a somewhat technical concept in second-language acquisition theory and then to discuss classroom implications. This finding is consistent with literature indicating that prior knowledge base can influence student interest and engagement [24]. Further research may be useful to identify the ways in which, over the course of a semester, different groups of students respond to different types of discussion prompts. In courses where participants come from diverse backgrounds, it may be useful to use heterogeneous small-grouping (splitting participants into smaller heterogeneous groups to force interaction with others) or to examine the phrasing of questions to ensure that the discussion question builds on prior knowledge for all participants.

Conclusion and Recommendations for Further Research

This exploratory data analysis suggests that student involvement, as measured by several quantitative indicators, may be influenced by prior knowledge of subject matter. The data analysis also indicates that social network analysis tools such as NodeXL may be valuable elements in program analysis, particularly in teacher education, as they provide graphical evidence of student involvement, provide quantitative measures of students' interaction with one another and each student's relative importance to the course "network," and provide a means to view each student's interaction with the instructors and with other students of similar and diverse backgrounds.

Given these findings, it is recommended that further research be conducted to refine a method for using NodeXL to analyze interaction in online courses; to determine whether participants in online courses interact differently with those of similar or dissimilar backgrounds and roles, based on their own backgrounds or roles; to determine whether students' relative "closeness" within a network is impacted by the degree of instructor involvement; to determine whether differences in participation levels among various groups of students have statistical significance and whether changes in instructional practice lead to significant differences in participation; and to determine, through qualitative as well as quantitative

research, the extent to which varying types of discussion prompts may impact student engagement in online discussions. As social network analysis has the power to quantify dynamics of participation according to various factors, this method of analysis also can facilitate the important goal of culturally and linguistically responsive practice [25], as students' interaction and participation can be analyzed according to cultural and linguistic background to identify questions, discussion practices, instructional activities that facilitate optimal involvement for all learners, however that term may be defined within the context of a specific course, program or discipline.

Within the context of teacher preparation, as previously stated, students' quantity and quality of participation, and ability to interact with those of varying backgrounds, can be seen to have direct implications for the development of needed professional skills. Given the increasingly significant role of online coursework in higher education and in teacher preparation, we believe that social network analysis, including both visualization and quantitative measures, has an important role to play in helping instructors, course developers, and program administrators to identify factors and dynamics which can impact student involvement and mastery of course and program objectives.

References

- [1] S. Cleaver, "Beyond Blackboard and into Virtual Communities," *Issues in Higher Education*, 25:18, 2008. Retrieved from ERIC database.
- [2] C. Gunawardena, L. Ortegado-Layne, K. Carabaja, C. Frechette, K. Lindemann, K. & B. Jennings, "New Model, New Strategies: Instructional Design for Building Online Wisdom Communities, Distance Education, 27 (2), p. 217-232.
- [3] M. Wang, Designing Online Courses that Effectively Engage Learners from Diverse Cultural Backgrounds. *British Journal of Educational Technology*, 38 (2), 2007, pp. 294-311.
- [4] M. J.W. Thomas, Learning within Incoherent Structures: The Space of Online Discussion Forums, *Journal of Computer Assisted Learning* 18, 2002, pp. 351-366.
- [5] D. Curtis and M. Lawson, "Exploring Online Collaborative Learning," *JALN* 5:1, February 2001, 21-34.
- [6] S. Wan and S. Lin, The application of social cognitive theory to web-based learning through NetPorts, *British Journal of Educational Technology*, 38:4, 2007, p. 600-612.
- [7] Y. Gulbahar, R.O. Madran, Communication and Collaboration, Satisfaction, Equity, and Autonomy in Blended Learning Environments: A Case from Turkey, *International Review of Research in Open and Distance Learning*, 10:2, Apr 2009.

- [8] Teachers of English to Speakers of Other Languages, TESOL/ NCATE Program Standards for Accreditation of Initial Programs in P-12 ESL Teacher Education, New York, NY: 2009.
- [9] Council for Exceptional Children, What Every Special Educator Must Know, Arlington, VA: Council for Exceptional Children, 2002.
- [10] National Council on Accreditation for Teacher Education, Professional Development School Standards: Core Concepts, 2007. Retrieved Mar. 30, 2007 from <http://www.ncate.org/public/pdsConcepts.asp?ch=138>
- [11] P. J. Slagter van Tryon, M. J. Bishop, Theoretical Foundations for Enhancing Social Connectedness in Online Learning Environments, Distance Education, vol. 30, no.3, 2009, p. 291-315.
- [12] J. Lave, & E. Wenger, Situated learning: Legitimate peripheral participation, Cambridge: Cambridge University Press, 1991.
- [13] D. Curtis and M. Lawson, "Exploring Online Collaborative Learning," JALN vol. 5, no.1, February 2001, pp. 21-34.
- [14] A. Bandura, Social foundations of thought and action: a social cognitive theory, Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [15] D. Hansen, D. Rotman, E. Bonsignore, N. Milic-Frayling, E. Rodrigues, M. Smith, B. Shneiderman, "Do You Know the Way to SNA?: A Process Model for Analyzing and Visualizing Social Media Data", in University of Maryland Tech Report: HCIL-2009-17
- [16] P. Marsden, "Recent developments in network measurement," in J. Carrington, J., and S. Wasserman, (eds.), Models and Methods in Social Network Analysis, Cambridge, UK: Cambridge University Press, 2005, pp. 8-30.
- [17] L. C. Freeman, "Centrality in Networks: I. Conceptual Clarification", Social Networks vol. 1, 1979, pp. 215-239.
- [18] E.M. Bonsignore, C. Dunne, D. Rotman, M. Smith, T. Capone, D.L. Hansen, B. Shneiderman, "First steps to NetViz Nirvana: evaluating social network analysis with NodeXL", SIN '09: Proc. international symposium on Social Intelligence and Networking. IEEE Computer Society Press.
- [19] W. Penuel, M. Riel, A. Joshi, L. Pearlman, M.K. Chong, K. Frank, "The Alignment of the Informal and Formal Organizational Supports for Reform: Implications for Improving Teaching in Schools, Educational Administration Quarterly, Feb2010, vol. 46, no. 1, p.57-95.
- [20] M. Smith, B. Schneiderman, N. Milic-Frayling, E. Mendes-Rodrigues, V. Barash, C. Dunne, T. Cap, A. Perer, and E. Gleave, "Analyzing (social media) networks with NodeXL", Proceedings of the 4th Annual Conference on Communities & Technologies, University Park, PA, 2009.
- [21] A. Calvani, A. Fini, M. Marcello and M. Ranieri, "Visualizing and Monitoring Effective Interactions in Online Collaborative Groups", British Journal of Educational Technology vol. 41, no. 2, 2010, pp. 213-226.
- [22] P. Shea, S. Hayes, J. Vickers, M. Gozza Cohen, S. Uzuner, R. Mehta, et al, "A Reexamination of the Community of Inquiry Framework: Social Network and Content Analysis", Internet & Higher Education [serial on the Internet], vol. 13, no. ½, Jan. 2010, pp. 10-21.
- [23] M. Santovec, Strategies to ensure quality, Distance Education Report, vol. 8, no. 22, 2004, pp. 1-7.
- [24] S. Tobias, "Interest, prior knowledge, and learning," Review of Educational Research, vol. 64, no. 1, 1994, pp. 37-54.
- [25] G. Gay, and T. C. Howard, "Multicultural Education for the 21st Century", The Teacher Educator, Vol. 36, No. 1, 2001, pp. 1-16.