

Evaluación de métricas de QoS/QoE aplicadas a sistemas de iDTV

Diego Jose Luis BOTIA VALDERRAMA
Grupo de Investigación GITA - ARTICA
Facultad de Ingeniería. Universidad de Antioquia
Medellín, Antioquia, Colombia
Email: diego.botia@gmail.com

y

Natalia GAVIRIA GOMEZ
Grupo de Investigación GITA – ARTICA
Facultad de Ingeniería. Universidad de Antioquia
Medellín, Antioquia, Colombia

RESUMEN

En este artículo se presenta inicialmente un estudio global de las principales métricas de *QoS* (Calidad de Servicio) y *QoE* (Calidad de la Experiencia) aplicados a sistemas de *iDTV* (TV Digital Interactiva). Se muestran sus principales características y como pueden correlacionarse entre sí para obtener altos niveles de calidad en la entrega de video hacia el televidente, por medio de una cuidadosa selección de métricas y de esquemas de *QoS*. Se presenta un *testbed* (entorno operativo de simulación) que a través de tres escenarios de simulación como una *QoS* (tipo *Besteffort*), una red con *QoS* (Tipo *Diffserv*) y un escenario utilizando a técnica *MDC* (*Multiple Description Coding*), permiten comparar las métricas de calidad objetivas y subjetivas entre las que se destacan las métricas referenciadas (*FR* o Referencia Completa, *NR* o Sin Referencia y *RR* o Referencia Parcial) de acuerdo a si necesitan comparar el video original codificado en algún formato de compresión, contra el video entregado por la red (con distorsión); además se explica el proceso de medición y las metodologías de evaluación propuestas por diversas investigaciones, teniendo en cuenta que es necesario en todo momento garantizar el cumplimiento de las necesidades de *SLA* (*Service Level Agreements*) de las empresas de telecomunicaciones y del televidente, por lo tanto es muy importante la entrega de video sin artefactos que disminuyan su calidad.

Palabras Claves: Métricas Objetivas, Métricas Subjetivas, iTV, IPTV, MDC, QoS, QoE.

1. INTRODUCCIÓN

La televisión digital actualmente está generando una nueva tendencia en la prestación de servicios de telecomunicaciones debido a que permite no sólo enviar video y audio en definición estándar (*SD*) o en alta definición (*HD*), sino que además se empaquetan servicios de datos, que conlleva a la creación de una amplia gama de aplicaciones como el *VoD* (*Video on Demand*), *PVR* (*Personal Video Recorded*), *Streaming*, aplicaciones interactivas, etc; pero que en algunas ocasiones pueden verse afectados en la entrega al usuario final, debido a factores como errores en la red, mala planeación de los recursos de red, deficiente configuración de los servicios y enlaces, o un uso inadecuado de los mecanismos de codificación o decodificación del video y que producen una pobre calidad en la última milla. Por lo tanto un nuevo factor que las empresas prestadoras de servicios (*Telcos*) deben tener en cuenta es el

cumplimiento de las expectativas de sus usuarios respecto al servicio que reciben, muchas veces plasmadas en los *SLAs* (*Service Level Agreements*), para lo cual se debe aplicar algún mecanismo que permita correlacionar métricas de *QoS/QoE*, que permitan alcanzar altos niveles de calidad extremo a extremo, realizando un dimensionamiento y planeamiento adecuado del tráfico.

En particular, los servicios que generan tráfico inelástico, como la voz y el video, imponen ciertas restricciones en cuanto a las métricas de *QoS*, porque el retardo (*Delay*), la pérdida de paquetes (*packet Loss*) y la varianza del retardo (*Jitter*) influyen directamente en la calidad de la imagen o el sonido percibido por el usuario. Para mejorar la calidad de la transmisión de este tipo de tráfico, se han propuesto diversas técnicas, basadas en el principio de reserva de recursos y en el control de admisión [11][12]. Para evaluar el desempeño de estos mecanismos, se debe tener en cuenta la percepción del usuario, que se traduce en un conjunto de métricas, conocidas como calidad de la experiencia (*QoE*), que le permiten evaluar el servicio que percibe y que pueden medirse tanto objetiva (cuantitativamente) como subjetivamente (cualitativamente) y que en el caso de los servicios de video, depende del sistema de visión humana (*SVH*)[13]. De esta forma algunos artefactos, como la visión borrosa (*Blur*), las sacudidas del video (*Jerkiness*) o la Tasa de pérdida de cuadros (*Frame Loss Rate*) son percibidas por el usuario como una mala imagen, y se traducen por tanto en una baja *QoE*.

La primera aproximación a la relación de la *QoS/QoE* es dada por Nieblas et.al[29] y Nordström [31] en el cual se proponen dos dominios globales, donde el dominio de la *QoE* está asociada a la capa del servicio que es más cercana al usuario y donde se presenta el software y las aplicaciones finales y las principales métricas a medir aquí son la Calidad de todo el video, Tiempo de arranque del sistema, Respuesta del tiempo Navegación, Channel Zapping Time y Tiempo de Respuesta al Control del *VoD*; por otro lado el dominio de la *QoS* se asocia a la capa de Aplicación, y está relacionada con el tipo de video que se maneja para alcanzar los niveles esperados de *QoE*, acá las principales métricas a medir son el Retardo del Middleware (Ej. *MHP*- OpenIPTV, etc), Retardo del Codificador, Tiempo de arranque del STB, Tiempo de procesamiento de comandos del STB, y Tiempo de procesamiento de los Servidores *VoD*. La otra capa de este dominio es la capa de red / transporte que se relaciona con el performance de la red que va de acuerdo a las

métricas que se pueden gestionar, entre las que están *Packet Loss Ratio (PLR)*, *Bit error Rate (BER)*, *Packet Latency*, *Packet Jitter* y Problemas de asignación de ancho de banda. Existen trabajos dónde se proponen nuevas métricas para sistemas de IPTV, como el propuesto por Romaniak [30], el cual hace una primera aproximación para la optimización de la calidad de flujos de video en términos del nivel de satisfacción del usuario. Indica 3 ejemplos de métricas de video como el *PEVQ* (Perceptual Evaluation of Video Quality), *VQM* (Video Quality Metric) y el *MPQM* (Moving Pictures Quality Metric); los cuales tienen en cuenta factores como el tipo de compresión, el análisis de bloques de video, nivel de ruido entre otras.

El artículo está organizado de la siguiente manera, en la sección 2 se presentarán las principales métricas de *QoS* y *QoE* para plataformas de *iDTV*, así como la clasificación de las métricas objetivas y subjetivas (*Full Reference o de referencia completa*, *No Reference o sin referencia*, y *Reduced Reference o Referencia Parcial*) y se explica la técnica *MDC* (*Multiple Description Coding*). En la sección 3 se presenta el *testbed* y el análisis de los resultados obtenidos en las simulaciones, y finalmente se presentarán las conclusiones y la bibliografía.

2. CONCEPTOS

Métricas de QoS y QoE para Plataformas de TV Digital

Calidad de Servicio (QoS): Se define como la medición de las características de calidad de una red para que ésta funcione de una manera eficiente, donde sus requerimientos varían dependiendo de la aplicación. Según [14], la calidad de servicio es la capacidad que tiene una red de proveer de la mejor manera los servicios, para un tipo de tráfico seleccionado sobre varias tecnologías. El objetivo principal de analizar y proveer *QoS* es poder controlar parámetros críticos como ancho de banda, *jitter*, latencia, o pérdidas de paquetes, para garantizar una buena transmisión de la información y poder dar prioridad a aquellos flujos que la requieren. En [14], se proponen 3 elementos básicos para una arquitectura de red con *QoS*. El primero es el uso de técnicas de identificación y marcado para coordinar *QoS* extremo a extremo en la red. El segundo es la *QoS* en los diferentes dispositivos de la red, donde se aplican conceptos como encolamiento (*queuing*), planificación (*scheduling*), y *traffic-shaping* (Modelado de tráfico). Por último, están las políticas de *QoS*, su manejo y las funciones de contabilidad para controlar y gestionar el tráfico extremo a extremo.

La ITU [1] ha propuesto las recomendaciones Y.1540 y Y.1541 para definir los parámetros ideales de *QoS*, los cuales se muestran en la tabla 1. En esta se encuentran los parámetros de red mencionados antes y las clases de servicio, en donde la clase 5 corresponde a tráfico de tipo Besteffort (sin garantías), la clase 0 y 1 se usa para aplicaciones sensibles al retardo y su variación (*jitter*), en donde se recomienda hacer separación de colas mediante servicios diferenciados (*Diffserv*). Según Domenico [2] esta clase de arquitectura es muy usada actualmente, porque puede manejar diferentes requerimientos de las aplicaciones y así asignarles prioridades a los tráficos. Para el funcionamiento de esta arquitectura no se necesita ningún protocolo de reserva de recursos como *RSVP* (*Protocolo de Reserva de Recursos*) ya que no es necesario efectuar ninguna clase de petición a los nodos que conforman la red. Al no reservar recursos en la red para el envío de los diferentes paquetes, no se tiene una garantía de *QoS* como se tenía en *IntServ* (Servicios Integrados por reserva de recursos) por lo que el tráfico de baja prioridad puede verse afectado si la red se

sobrecarga con tráfico de alta prioridad. Para evitar esta situación se le asigna a cada categoría un *SLA*. El *SLA* es negociado con la *Telco* (*Empresa de Telecomunicaciones*) previamente y generalmente posee carácter estático. Dentro de la red los *routers* de borde son los responsables del control de admisión o *Policy Control* y así podrán colocar el valor correspondiente en los paquetes salientes para que sean tratados según su categoría dentro de la red del ISP. Los paquetes son así manejados en cada router mediante un condicionador de tráfico definido en el *PHB* (*Per-Hop Behavior*) que se indica según el campo *DSCP* (*Differentiated Service Code Point*) de los paquetes IP. Los routers Core se usan para hacer el tratamiento de la congestión del tráfico.

TABLA 1.
RECOMENDACIÓN ITU Y.1540/ Y.1541

Parámetro ITU Y.1540	Clase de QoS ITU Y.1541					Clase 5
	Clase 0	Clase 1	Clase 2	Clase 3	Clase 4	
IPTD	100ms	400ms	100ms	400ms	1s	N
IP Transfer Delay						A
IPDV	50ms	50ms	NA	NA	NA	N
IP Delay Variation						A
IPLR	$1 \cdot 10^{-3}$	$1 \cdot 10^{-3}$	$1 \cdot 10^{-3}$	$1 \cdot 10^{-3}$	$1 \cdot 10^{-3}$	N
IP Loss Ratio						A
IPER			$1 \cdot 10^{-4}$			N
IP Error Ratio						A

El problema de la arquitectura *DiffServ* es el procesamiento que se debe de hacer a los diferentes tipos de tráfico ya que se tienen prioridades diferentes, por ejemplo, que hacer si un paquete de menor prioridad, es procesado cuando uno con mayor prioridad llega a la cola del *router*. Para esto en *DiffServ* se definen tres tipos de servicio:

- **Servicio ‘Expedited Forwarding’ o ‘Premium’:** Este servicio es el de mayor calidad. Debe ofrecer un servicio equivalente a una línea dedicada virtual, o a un circuito *ATM CBR* (*Constant Bit Rate*) o *VBR-rt* (*Variable Bit Rate – Real Time*).
- **Servicio ‘Assured Forwarding’:** Este servicio asegura un trato preferente, pero no garantiza caudales, retardos, etc. El proveedor puede aplicar *traffic policing* al usuario, y si el usuario excede lo pactado el proveedor puede descartar datagramas, o bien aumentar la precedencia de descarte.
- **Servicio Best Effort:** En este servicio no se ofrece ningún tipo de garantías.

Calidad de la Experiencia (QoE): El concepto de calidad de la experiencia (*QoE*) surge de la necesidad de determinar el grado de satisfacción de los usuarios con el servicio. Al igual que la *QoS*; la *QoE* ha sido definida de diferentes formas por varios autores. Dentro de estas definiciones, tiene especial relevancia la de Li-yuan [15], quien indica que la *QoE* implica dos aspectos: el primero es el monitoreo de la experiencia del usuario en línea y el segundo es el control del servicio para asegurar que la *QoS* pueda conocer ampliamente los requerimientos del usuario. Por otro lado, según López et.al [16], la *QoE* es una extensión de la *QoS* en el sentido de que la primera proporciona información acerca de la entrega de servicios desde el punto de vista del usuario final. Según Kilkk

[17] y Winkler et.al [18], en la *QoE* se hacen mediciones subjetivas y objetivas al video transmitido sobre la red, que dependen de las expectativas del cliente y por lo tanto se tienen en cuenta las sensaciones, las percepciones y la opinión de los usuarios. Según estos autores, entre los principales factores que se deben tener en cuenta para medir y evaluar la *QoE* en el caso de *iDTV* están: los intereses individuales del televidente (por ejemplo, si observa un programa de televisión favorito y por lo tanto puede determinar un nivel de atención más alto), la experiencia en la evaluación de videos por parte del individuo (por ejemplo si diferencia entre video en calidad *SD* o *HD*), el tipo de pantalla (*CRT*, *LED*, *LCD*, Plasma y sus propiedades asociadas como tamaño, resolución, contraste, color, brillo, tiempo de respuesta), las condiciones y el ambiente donde se encuentra el usuario (la distancia del televidente a la pantalla o condiciones de iluminación), la calidad y sincronización del audio con el video y la interacción con el servicio o el aparato de visualización (es decir si tiene acceso a servicios como la Guía electrónica de programas, el *cannel zapping time*, o el tipo de control remoto).

A partir de estas definiciones de *QoE*, se puede concluir que ésta hace referencia a preferencias personales de los usuarios, y por lo tanto busca evaluar la percepción subjetiva que tienen ellos del servicio recibido. Sin embargo, esta percepción está influenciada en alto grado por el desempeño de la red en cuanto a los parámetros de *QoS* definidos. Así para la estimación de la calidad en la transmisión de un servicio, se ven involucrados métodos de medición tanto subjetivos como objetivos.

Para sistemas de *IPTV* (*TV por protocolo IP*), los métodos subjetivos son usados para establecer el desempeño de los sistemas de *TV* usando mediciones que están más directamente relacionadas con la percepción de los usuarios. Según Klaue et al[7], para realizar las pruebas subjetivas se emplean en promedio entre 15 y 30 personas que observan una serie de video clips y los evalúan en un rango numérico definido. Los resultados de la prueba permiten calcular el *MOS* (*Mean Opinion Score*), el cual según la ITU Ibid[19] es una medida numérica usada en tráfico multimedia que determina la calidad percibida de los datos que se reciben después de la compresión y/o la transmisión. Esta medida va en un rango de 1 (más bajo) a 5 (el mejor). Según Kuipers et al [20], el mínimo umbral que considera una calidad aceptable corresponde a un *MOS* de 3.5.

A pesar de que las pruebas subjetivas (como el *MOS*) son bastante útiles en la medición de la satisfacción de los usuarios, la implementación de las mismas es compleja debido al consumo de tiempo y los altos costos que implican. Además, si se quieren implementar técnicas de gestión de tráfico en tiempo real, es necesario encontrar una relación de éstas con las métricas objetivas, medibles por los equipo de red.

Las métricas objetivas de calidad, según Winkler Ibid [18], son algoritmos diseñados para caracterizar la calidad del video y predecir la opinión del televidente sin tener en cuenta la percepción del individuo. Las métricas objetivas más utilizadas son las métricas según la cantidad de información de referencia requerida para el cálculo, las cuales se pueden clasificar en *NR* (No Reference), *FR* (Full Reference o referencia completa) y *RR* (Reduced Reference o Referencia Reducida) [21] [22]. Las métricas *FR* miden la degradación en el video de prueba recibido con respecto al video de referencia ubicado en la fuente. Para ello, se requiere acceso al video de referencia completo, generalmente sin deterioro y sin compresión,

imponiendo además un alineamiento espacial y temporal, ya que cada pixel en cada cuadro se compara con el video recibido. La métrica *PSNR* (*Peak Signal to Noise Ratio*), es la más conocida, y evalúa la calidad de la secuencia de video recibida y es luego mapeada en una escala subjetiva. Si se consideran cuadros con un tamaño de $M \times N$ *pixels* y 8 bits/muestra el *PSNR* se puede calcular usando la Ec. (1) [13] [23].

$$PSNR = 20 \log_{10} \left(\frac{255}{\sqrt{\frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \|Y_s(i,j) - Y_d(i,j)\|^2}} \right) \quad (1)$$

donde $Y_s(i,j)$ indica un pixel en la posición (i, j) del cuadro original y $Y_d(i, j)$ se refiere al pixel localizado en la posición (i, j) del cuadro reconstruido en el lado del receptor. El valor 255 es el valor máximo que el pixel puede tomar (255 para imágenes de 8 bits). En forma general, el denominador se conoce como *MSE* o error cuadrático medio, el cual es la media cuadrática de las diferencias entre los valores de niveles de grises de los *pixels* en los imágenes o secuencias Y_s y Y_d . Es decir, la *PSNR* busca comparar pixel a pixel y cuadro a cuadro la calidad de la imagen recibida con la imagen original.

La relación entre la *PSNR* y el *MOS* ha sido desarrollada por Ke et al [24], Khan et al [25] y Zinner et al[26] como se ilustra en la tabla 2, donde los niveles superiores a 33 dB se consideran que representan una buena calidad en la entrega de video comprimido. También se muestra la relación con la métrica *SSIM* (*Structural Similarity Index Metric*)

TABLA 2
EQUIVALENCIA ENTRE PSNR Y MOS [21]

PSNR (dB)	MOS	SSIM
> 45	5(Excelente)	>0.99
33.1 - 45	4 (Bueno)	>= 0.95 & < 0.99
27.5 - 33	3(Aceptable)	>=0.88 & < 0.95
18.8 - 27.4	2 (Pobre)	>=0.5 & < 0.88
< 18.7	1 (Malo)	< 0.5

Existe un conjunto de métricas que buscan además capturar elementos estructurales de la imagen (contornos o distorsiones específicas introducidas bien sea por los codificadores o por la transmisión en la red, así como artefactos espaciales y temporales), a través de la extracción y el análisis de características o de artefactos en el video [18]. En esta categoría, se destaca el *SSIM* que calcula la media, la varianza y la covarianza entre los cuadros enviados y recibidos. Para calcular el *SSIM*, se miden 3 componentes (la similaridad de luminancia, la similaridad del contraste y la similaridad estructural), los cuales se combinan en un solo valor llamado índice *SSIM* que va entre 0 y 1, donde 0 indica cero correlación con la imagen original y 1 significa que es la misma imagen [22]. Otra métrica importante es *VQM* (*Video Quality Metric*)[3], que toma como entrada la imagen del video original y el video procesado, y verifica los niveles de calidad con base en la percepción del ojo humano y aspectos subjetivos, para lo cual divide la imagen en secuencias de bloques temporales – espaciales midiendo elementos como el *blurring*, el ruido global, la distorsión de bloques y la distorsión del color. Los resultados varían de 0 a 5 donde 0 es el mejor valor posible.

La métrica *RR* únicamente selecciona algunos parámetros del video original, como la información del movimiento. El objetivo de esta categoría es usar menos recursos de procesamiento y de red, pero el principal problema es que requieren de un canal alterno para la transmisión de estos

parámetros y de acceso a la imagen de referencia original en algún punto. La métrica NR analiza el video de prueba de salida sin necesidad de acceder al video de referencia. Este método se usa cuando el mecanismo de codificación es conocido y además puede usarse en monitoreo de red u operaciones de diagnóstico especiales, pero el principal problema es que tiene una baja correlación con MOS y un alto consumo de recursos computacionales.

Codificación del Video: Antes de transmitir el video se debe hacer un proceso de codificación y compresión con el objetivo de consumir lo menos posible, los recursos limitados de la red. El método de codificación más utilizado es el formato MPEG (MPEG2, MPEG4-AVC y H.264) [4], que se ha probado exitosamente en la transmisión de audio y video, en sistemas de TV digital. Estos esquemas utilizan los llamados cuadros I, B y P, en los cuales por medio de una combinación de la transformada del coseno discreto (DCT) se establece una predicción de movimiento (a nivel Temporal y Espacial). Generalmente los cuadros I usan una codificación espacial y su compresión es pequeña, los cuadros P o predictivos se codifican con respecto a las imágenes de tipo I o P anteriores y usan compensación de movimiento temporal y los cuadros B que se interpolan bidireccionalmente entre dos imágenes de tipo I o P precedentes y/o siguientes que las enmarcan y usan la mayor compresión. Los cuadros I son más importantes que los cuadros P o B por lo tanto una pérdida por PLR (*Packet Loss Rate*) de un cuadro I en la transmisión podría reflejar muchos más artefactos o distorsiones.

Multiple Description Coding (MDC)

Según Begen et.al [27] y Lee et.al [28] MDC es una técnica que codifica una fuente de video en varios descriptores (*bitstreams*). La principal ventaja es que diferentes niveles de calidad se obtienen desde diferentes subconjuntos de estos descriptores. Cada descriptor se decodifica independientemente. Esta técnica es adecuada cuando se tiene una red sin garantías de QoS . MDC tiene una alta tolerancia a pérdidas de paquetes y retardos debido a congestión en la red, por lo que es una técnica muy eficiente. Con el fin de decodificar el flujo, cualquier descriptor puede ser usado, sin embargo la calidad mejora con el aumento del número de descripciones recibidas en paralelo. La principal ventaja de MDC es proporcionar capacidad de recuperación a errores de los flujos.

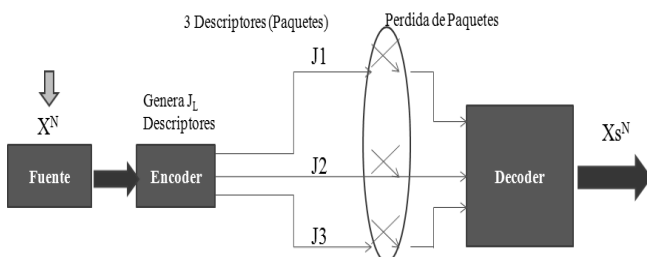


Fig 1. Funcionamiento del MDC

En la figura 1 se muestra que un subconjunto arbitrario de descriptores se pueden usar para decodificar el flujo original y aunque se presente congestión de la red o pérdidas de paquetes, no se interrumpirá el flujo produciendo una pérdida de calidad temporal. La descripción matemática de este modelo es:

- Fuente: Es una variable aleatoria definida en un vector X^N Donde X^N es i.i.d

- Encoder: $X^N \rightarrow \{J_1, J_2, \dots, J_L\}$ Cada Descriptor J_L se envía a tasas R_1, R_2, \dots, R_L
- Descriptor: $J_i = f_i(X^N) \quad i=1, \dots, L$
- Decoder: 2^{L-1} Decoders (Uno por cada subconjunto no vacío de Descriptores Disponibles)
- Salida del Decoder: $X_s^N = f_s(\{J_i: i \in S\})$ donde $S \subseteq \{1, \dots, L\} \quad S \neq \emptyset$

MDC permite además usar una tasa adaptativa para el streaming, por lo tanto los proveedores de contenido envían todos los descriptores del flujo sin prestar atención a las limitaciones de descarga de los clientes y así los receptores que no pueden sostener la tasa de datos únicamente se suscriben a un subconjunto de estos flujos, liberando así al proveedor de contenidos del envío de flujos adicionales a tasas más bajas de datos.

3. TESTBED Y RESULTADOS

Para el *Testbed* se implementaron tres escenarios que facilitaron la simulación. El primero es una red de tipo *Besteffort* (Sin QoS) usando colas de tipo *FIFO* (*First In First Out*). El segundo escenario fue una red con QoS mediante un esquema *Diffserv*, utilizando el algoritmo *WRED* (*Weighted Random Early Detection*) [5], el cual utiliza 3 colas virtuales a las que se les aplico la política *TSW3CM* (*Time Sliding Window With three Colour Marker*) [6], en la cual se debe asignar un valor para el *CIR* (*Committed Information Rate*) y el *PIR* (*Peak Information Rate*) según la capacidad de los enlaces. El tercer escenario fue el desarrollo de una red aplicando la técnica MDC (*Multiple Description Coding*) a la que se le aplicaron 4 descriptores. Todos los escenarios se simularon mediante *NS2* y el *framework Evalvid* [7]. En la figura 2 se presentan los escenarios que se utilizaron, donde se enviaron trazas de tráfico de video usando el protocolo *RTP* (*Real Time Transport Protocol*) que se generaron mediante la herramienta *MP4trace*, en donde se obtiene información como el número del *frame*, el tipo de *frame* (I, B o P), el tamaño del *frame*, el número de la secuencia y el *timestamp*.

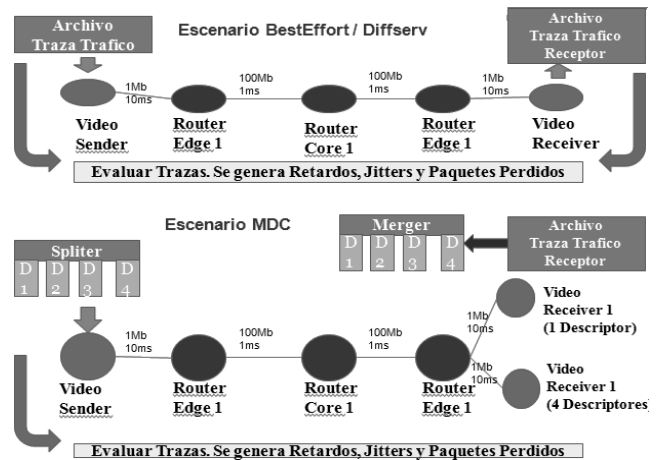


Fig 2. Escenarios simulados con NS-2 (BestEffort / Diffserv y MDC)

En la figura 2 se observa además la velocidad establecida por el *sender* (*Servidor de Video*) y *receiver* (simulando una última milla), y la red que se compone de 2 enrutadores de borde y 1 enrutador *core*. Según Flierl et al[8] para el escenario usando MDC , es necesario utilizar la herramienta *splitter* que permitirá generar los descriptores deseados antes de la transmisión y luego en cada *receiver* se reconstruirán mediante la herramienta *merger* para posteriormente evaluar cada métrica. Se evaluaron

4 videos con diferentes niveles de actividad los cuales se codificaron en formato *MPEG-4*. Los niveles de actividad (Bajo, Medio y Alto) se refieren a la cantidad de cambios de fondo y movimiento en el video, por ejemplo un noticiero tendría un nivel de actividad bajo mientras que una carrera de autos sería un nivel alto, esto es importante porque puede influir en la calidad final medida. Para las simulaciones se utilizaron 4 muestras de videos (News, Highway, Coastguard y Trailer). Estos videos se probaron con diferentes tipos de resolución *QCIF* (176x144p), *CIF* (352x288p) y otra de 480x280 pixels. La tabla 3 y 4 muestran los parámetros del *encoder* y de la simulación.

TABLA 3
PARÁMETROS DEL ENCODER UTILIZADOS

Longitud GOP	30 Frames
Tasa del Frame	25 fps
BitRate	700 kbps
Tipos de Frames soportados	I, B, P
Resolución	QCIF 176 x 144 p y CIF 352 x 288 p
Modo Color de Video	YUV (4-2-0)

Después de ejecutar la simulación, se obtuvieron varios archivos que contienen datos como el número de *frames* perdidos, el retardo y *jitter*, que permitieron hacer un análisis de cada video con respecto a cada escenario implementado. Se utilizó además la herramienta *MSU VQMT* [10] para obtener las métricas *PSNR*, *SSIM*, y *VQM*, por medio del video de referencia original y el video con distorsión recibido.

TABLA 4
PARÁMETROS DEL SIMULADOR

Velocidad de Enlace	1/100 Mbps
Retardo de enlace	1 ms
Tamaño de Buffer	50
Comportamiento de Cola	DropTail / WRED
Tamaño de Paquete	1052 bytes (8 bytes UDP y 20 bytes IP)
Max Tamaño de Fragmentación	1024 bytes

La figura 3 muestra la relación entre la métrica *MOS* y *PSNR* para los videos utilizados, donde se puede observar un aumento considerable del *PSNR* para los videos que se transmitieron con el esquema *Diffserv*, por lo tanto se tiene un *MOS* entre 4 y 5 lo cual indica un video con bajas distorsiones. En cambio los videos con el esquema *Besteffort* presentan un *PSNR* y *MOS* muy bajos. El esquema con *MDC* fue aplicado al video *coastguard* en resolución *CIF*, en el cuál usando 1 descriptor se obtuvo un *PSNR* de 27.3dB, mientras que usando 4 descriptores se presentó una mejora obteniendo un *PSNR* 42.46 dB. A pesar de estos resultados es de destacar que con 1 sólo descriptor se obtuvo un mejor resultado en la calidad que cuando se envió el video sin *QoS* (16.35dB), debido a que la utilización de *MDC* es recomendable en redes con alta posibilidad de pérdida de paquetes según se explicó antes.

En la figura 4 se presenta la relación encontrada entre las métricas *MOS* y *SSIM*. Es importante recordar que un valor bajo de *SSIM* indica una muy baja similaridad con la imagen original, por lo tanto el valor ideal debe acercarse a 1. En este caso para los videos sin *QoS* se obtuvieron valores en promedio por debajo de 0.8 en cambio los videos a los que se les aplico *QoS* tuvieron un *SSIM* muy cercano a 1. Por otra parte se pudo establecer una correlación lineal con un buen ajuste, con lo que se demuestra la estrecha relación de estas métricas. Además siguiendo la tendencia presentada en la relación *MOS* - *PSNR* al aplicar los descriptores con *MDC* se mejora sustancialmente el incremento de la calidad con respecto a la red *besteffort*.

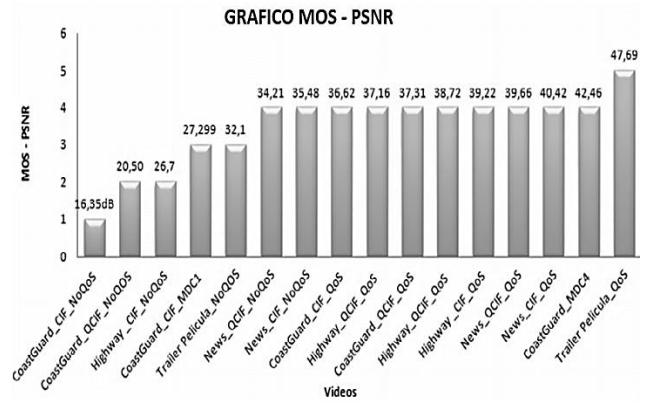


Fig 3. Relación entre MOS y PSNR

La figura 5 muestra la relación encontrada entre las métricas *VQM* y *SSIM*. Recordemos que el valor ideal de *VQM* debe tender a 0, y como se observa hay una estrecha relación entre el valor máximo de *SSIM* (Con buena similaridad estructural) y el *VQM* (tendiendo a cero) para el caso los videos transmitidos con *QoS* y con *MDC* de 4 descriptores.

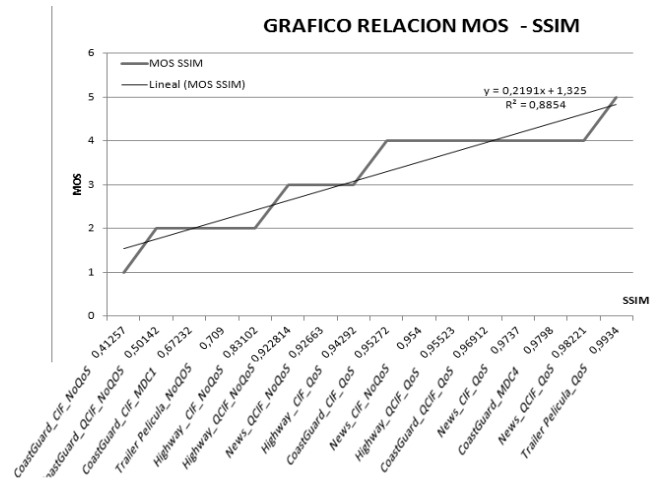


Fig 4. Relación entre MOS y SSIM

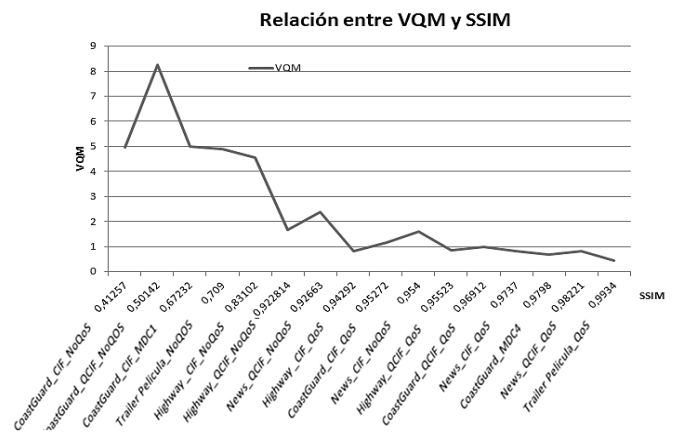


Fig 5. Relación entre VQM y SSIM

La figura 6 presenta el comportamiento entre la métrica *PSNR* y *SSIM*, en donde se puede observar que los videos transmitidos sin *QoS* (con *SSIM* tendiendo a 0) tienen bajo *PSNR*, a diferencia de los videos enviados en un esquema *Diffserv* que obtuvieron valores de *PSNR* más altos. También se puede ver un ajuste lineal de esta relación.

RELACION PSNR - SSIM

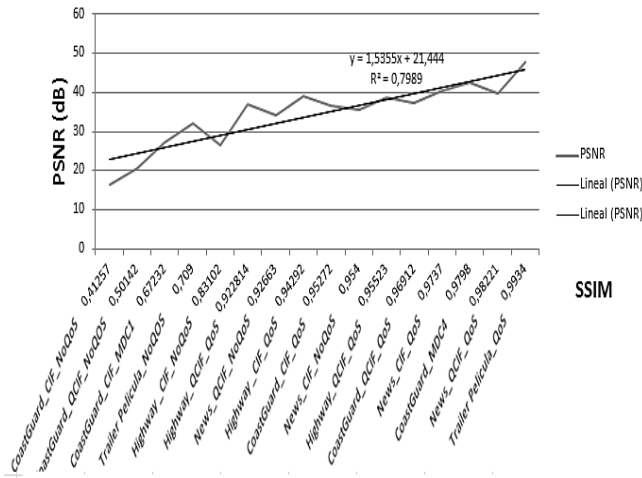


Fig 6. Relación entre PSNR y SSIM

Resultados similares se obtuvieron para la relación entre *PSNR* y *VQM* mostrados en la figura 7.

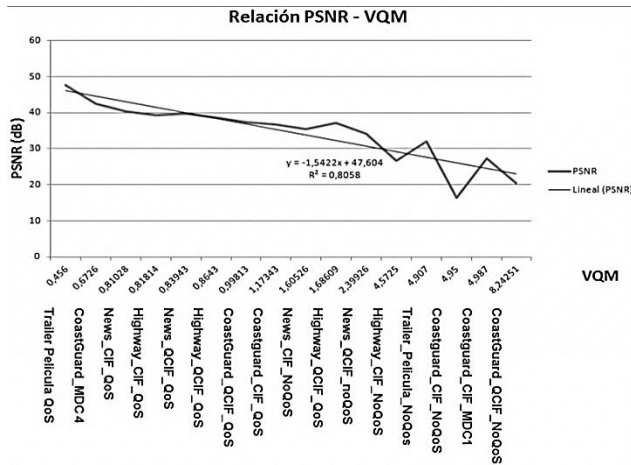


Fig 7. Relación entre PSNR y VQM

Por último se presenta en la figura 8 el comportamiento de los retardos medidos tanto en la transmisión tipo *Besteffort* como en la red con *Diffserv*. Se evaluaron todos los videos pero se presenta el comportamiento para el video de mayor duración y resolución (Tráiler de una Película). Se puede observar como el retardo se mantuvo bajo en toda la transmisión sobre la red con *WRED*, mientras que el video con cola *FIFO* mantuvo altos retardos conllevando a una baja calidad del video.

4. CONCLUSIONES

La información transmitida en las redes de *iDTV*, en donde en el extremo final se encuentra un humano, debe tener en cuenta las preferencias, sensaciones y percepciones del televidente para satisfacer sus necesidades. Este hecho ha traído nuevos retos a las redes de *iDTV*, puesto que sus parámetros de desempeño ya no se limitan únicamente a los parámetros medidos en la red, sino que es necesario involucrar un aspecto subjetivo, como es la percepción de los usuarios. En este contexto, surge el concepto de *QoE*, y la necesidad de correlacionarla con los parámetros de red establecidos en los modelos de *QoS* como *Diffserv*. Las principales métricas subjetivas que existen como el *MOS*, *PSNR*, *SSIM* y *VQM*, tienen grandes ventajas al

momento de estimar la *QoE*, pero todas manejan características distintas.

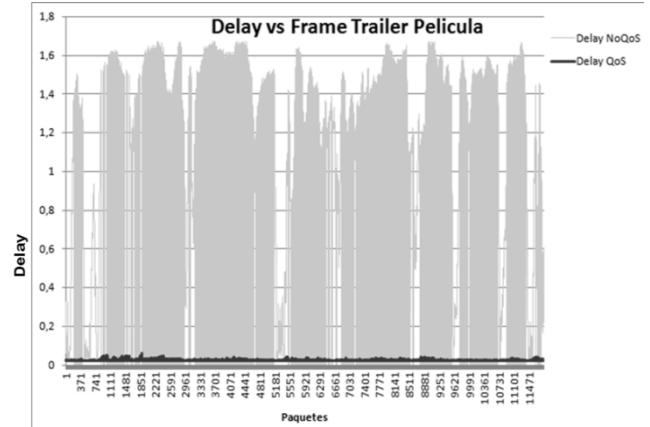


Fig 8. Retardos sin QoS y con QoS para Video Trailer

Por tanto, es recomendable si se quiere aplicar los modelos de correlación de *QoS/QoE*, realizar un completo estudio y selección cuidadosa de las métricas a aplicar y que puedan influir en un mayor o menor grado la percepción del usuario final. Si bien a la fecha no hay una metodología estandarizada para determinar la relación entre la *QoS* y la *QoE*, hay varios trabajos que hacen importantes aportes al respecto, y además de demostrar cómo la primera afecta la segunda, proponen métodos para correlacionarlas, que se convierten así en una herramienta clave para las *Telcos* que pretendan implantar y desarrollar nuevos servicios y aplicaciones sobre la red. Otra conclusión importante es que los niveles de compresión de video representan un factor crítico en el consumo de recursos y se deben tener en cuenta para una adecuada percepción del usuario así mismo el tipo de mecanismo de control de congestión utilizado en la red como p.e *WRED* aumenta la probabilidad de tener alta *QoE*. Por último se obtuvo un incremento notable de la calidad del video al momento de utilizar la técnica *MDC*, en donde él envió de un único descriptor mejoró los resultados de las métricas *QoE* con respecto a si se enviaba en la red *BestEffort*. Se recomienda además elevar el número de descriptores si se supone que la red presenta pérdidas de paquetes o alto nivel de retardos ya que así se tiene la probabilidad de obtener mejores resultados.

AGRADECIMIENTOS

Los autores agradecen el apoyo brindado por *ARTICA* (Alianza Regional para las *TIC* aplicadas), del cual hacen parte la Universidad de Antioquia, *EAFIT*, *UPB*, *UNAL* sede Medellín, *IPS* Universitaria, la Universidad *ICESI* de Cali y *UNE* Telecomunicaciones así como a *Colciencias* y al Ministerio de las *TIC*. También agradecemos el apoyo brindado por el Centro de Investigaciones Ambientales y de Ingeniería (*CIA*) de la Facultad de Ingeniería de la Universidad de Antioquia.

REFERENCIAS

- [1] Recomendación ITU Y.1541 QoS. [online] http://wftp3.itu.int/packet/Background%20Docs/Y_1541-200602-prepub.doc. Consultado el 01/11/2010
- [2] D. Luna, V. García, J. Igor, Medición y análisis de tráfico en redes MPLS. Pontificia Universidad Católica Del Perú. Facultad de ciencias e Ingeniería. 2008
- [3] F. Xiao, VQM – DCT Based Video Quality Evaluation. [online]

- http://compression.ru/video/quality_measure/vqm.pdf. 2000
- [4] H. Schwarz, D. Marple, Overview of the Scalable Video Coding Extension of the H.264/AVC Standard. IEEE Transaction on Circuits and Systems for Video Technology. September 2007.
- [5] M. Barbera, A. Lombardo, G. Schembra, A. Trecarichi, Improving fairness in a WRED-based DiffServ network: A fluid-flow approach. ACM Digital Library. Journal Performance Evaluation. Vol 65 Issue 10. Oct 2008.
- [6] W. Fang, N. Seddigh, A time sliding window three colour marker (TSWTCM) . RFC 2859. Junio. 2000
- [7] J. Klauwe, N. Rathke, A. Wolisz, "EvalVid - A Framework for Video Transmission and Quality Evaluation", 13th International Conference on Modelling Techniques and Tools for Computer Performance Evaluation, pp. 255-272, Urbana, Illinois, USA, September 2003
- [8] S.E. Flierl, T. Sikora, P. Frossard, Unbalanced Quantized Multiple State Video Coding, Hindawi Publishing Corporation EURASIP Journal on Applied Signal Processing. Volume 2006, Artículo ID 14694, Páginas 1–10
- [9] Technische Universität Berlin . CIF – QCIF Reference Videos. [online] <http://www.tkn.tu-berlin.de/research/evalvid/qcif.html>. Consultado el 01/Octubre/ 2010
- [10] D. Vatin, MSU Video Metric Quality Tool. MSU Graphics and media Lab [online]http://compression.ru/video/quality_measure/video_measurement_tool_en.html. Consultado el 01/Octubre/2010
- [11] A. Davy, D. Botvich, B. Jennings, Empirical Effective Bandwidth Estimation for IPTV Admission Control. Telecommunications Software & Systems Group, Waterford Institute of Technology Cork Rd., Waterford, Ireland, 2007
- [12] V. Ojansivu, O. Silvén, R. Huotari, Digital Video Quality Evaluation Method. [online] http://www.ee.oulu.fi/mvg/files/pdf/pdf_425.pdf University of Oulu, Finland. 2001.
- [13] S. Winkler, Digital Video Quality – Vision, Models and Metrics. Ed Jhon Wiley & Sons. Suiza. 2005
- [14] CISCO NETWORKS, Internetworking Technology Handbook. [online] <http://www.cisco.com/en/US/docs/internetworking/technology/handbook/QoS.html>. USA. Consultado 30-09-2010.
- [15] L. Li-yuan, Z. Wen-an, S. Jun-de, The Research of Quality of Experience Evaluation Method in Pervasive Computing Environment. 1st International Symposium on Pervasive Computing and Applications, 3-5 Aug. 2006, p.178 – 182.
- [16] D. López, F. González, L. Bellido, A. Alonso, Adaptive multimedia streaming over IP based on customer oriented metrics, International Symposium on Computer Networks, 16-18 June 2006, p. 185 – 191.
- [17] F. Kuipers, R. Kooij, D. Vleeshauer, K. Brunnstrom, Techniques for Measuring Quality of Experience. Book Series: Lecture Notes in Computer Science Ed. Springer ISSN0302-9743 Pag 216-227. Mayo 2010. [online] <http://www.springerlink.com/content/5028804658914365>
- [18] S. Winkler, P. Mohandas, The evolution of Video Quality Measurement: From PSNR to Hybrid Metrics. IEEE Trans. Broadcasting Vol 54 Nro 3. Septiembre 2008.
- [19] K. Kilkki, Next Generation Internet and Quality of Experience[Online]kilkki.net/files/50ajatelmaa.ajatukset.fi/.../kilkki_santander_v1.0.ppt, Junio 2007
- [20] S. Winkler, Video Quality Measurement Standars – Current Status and Trends. IEEE. 2009.
- [21] F. Boavida, E. Cerqueira, R. Choderek, M. Grega, M. Leszczuk, Z. Papir, P. Romaniak, C. Guerrero, Benchmarking the Quality of Experience of Video Streaming and Multimedia Search Services: The content Network of Excellence. KSTiT 2008 - XXIII Symposium of Telecommunications and Teleinformatics, Institute of Telecommunications and Electrotechnics of University of Technology and Life Sciences, Bydgoszcz, Poland, 10-12 September 2008.
- [22] E. Cerqueira, L. Veloso, M. Curado, E. Monteiro E, Quality Level Control for Multi-user Sessions in Future Generation Networks. University of Coimbra INESC Porto. Portugal
- [23] A. Khan, L. Sun, E. Ifeachor, Content Classification Based on Objective Video Quality Evaluation for MPEG4 Video Streaming over Wireless Networks. Proceedings of the World Congress on Engineering 2009 Vol I WCE 2009, London, U.K. July 2009
- [24] K. Chih H, C. Shieh, W. Hwang, A. Ziviani, An Evaluation Framework for More Realistic Simulations of MPEG Video Transmission. Journal of Information Science and Engineering Nro 24. Pag 425-440. 2008
- [25] T. Zinner, O. Abboud, O. Hohlfeld, T. Hossfeld, P. Train-Gia, Towards QoE Management For Scalable Video Streaming. In 21st ITC Specialist Seminar on Multimedia Applications Traffic, Performance and QoE. Marzo 2010.
- [26] A. Begen, O. Altunbasak, Multipath Selection for Multiple Description Encoded Video Streaming. Georgia Institute of Technology. 2002
- [27] I. Lee, Y. He, L. Guan, Centralized P2P Streaming with MDC. 2006
- [29] M. Nieblas, G. Bressan. A Proposed Approach for Quality of Experience Assurance for IPTV. International Conference on the Digital Society, Korea University of Technology and Education. Multimedia Communication and Network Lab. 2007.
- [30] P. Romaniak. Quality of Experience Assessment of Video-Based Applications – Introduction to Research. Department of Telecommunications. Faculty of Electrical Engineering, Automatics, Computer Science and Electronics AGH University of Science and Technology, Krakow, Poland, Mayo 2007
- [31] E. Nordström. Overview of IPTV systems. Ernst Consulting & Education. [online] <http://www.du.se/~eno/itv/iptv-system.pdf>. Suecia. 2009