

ALGORITMO DE LINK ADAPTATION PARA LA TRANSMISIÓN DE VÍDEO EN SISTEMAS DE COMUNICACIONES MÓVILES

Miguel López-Benítez
Área de Teoría de la Señal
y Comunicaciones
Universidad
Miguel Hernández

Javier Gozávez
Área de Teoría de la Señal
y Comunicaciones
Universidad
Miguel Hernández

Óscar Lázaro
IMCO2, Grupo de Comuni-
caciones Móviles
Universidad
Politécnica de Valencia

Javier Evangelio
Área de Teoría de la Señal
y Comunicaciones
Universidad
Miguel Hernández

j.gozalvez@umh.es

Abstract - This paper introduces and evaluates a new Link Adaptation algorithm designed to improve the real-time transmission of video over mobile radio networks. The algorithm selects the most suitable transport mode taking into account not only the throughput but also the transmission delay. The results for a real-time H.263 transmission demonstrate that the proposed algorithm not only improves the transmission delay compared to the traditional Link Adaptation algorithm but also the system throughput.

I. INTRODUCCIÓN

La creciente demanda de los servicios de voz en los sistemas de comunicaciones móviles y la reciente incorporación de servicios multimedia crean nuevos desafíos para los operadores móviles, los cuales se ven obligados a implementar técnicas para conseguir un aprovechamiento, cada vez más eficiente, de los escasos recursos radioeléctricos. Éste es el objetivo que persiguen las denominadas técnicas de gestión de recursos radio, entre las cuales se encuentra *Link Adaptation* (LA).

La base de LA consiste en evaluar de forma periódica la calidad del enlace radio y seleccionar, de entre un conjunto preestablecido, el modo de transporte que, bajo dichas condiciones, resulta óptimo de acuerdo a un criterio predefinido. Uno de los criterios más utilizados consiste en elegir el modo de transporte que proporciona la tasa de transmisión (*throughput*) más elevada [1]. Si bien este criterio puede resultar satisfactorio para servicios con escasos requisitos en términos de calidad de servicio (*best-effort*), su aplicación a servicios de tiempo real, con importantes exigencias de retardo y tasa de errores, resulta cuestionable; por lo que otras alternativas han sido propuestas. En [2] se presenta un algoritmo LA orientado a la consecución de una tasa de errores específica. En [3] se sugiere la elección del modo de transporte que maximiza la calidad de la imagen de vídeo, expresada en términos de la *relación señal a ruido de pico* (PSNR). En [4] se propone un algoritmo que pretende disminuir el retardo de transmisión. Dicho algoritmo escoge el modo de transporte en función del tamaño del paquete a transmitir, factor que influencia el retardo de transmisión.

En este contexto, el presente trabajo presenta y evalúa un nuevo algoritmo LA en el cual se considera para la elección

del modo de transporte óptimo no sólo la tasa de transmisión sino también el retardo, incluyendo este último factor de forma directa en el proceso de selección.

II. ENTORNO DE SIMULACIÓN

A. General Packet Radio Service (GPRS)

El presente trabajo está basado en la interfaz radio del estándar GPRS, el cual contempla cuatro esquemas de codificación (CS) distintos, ver Tabla 1. Cada CS proporciona un grado de compromiso diferente entre robustez ante errores (bits de redundancia) y tasa binaria de información (bits de datos), lo cual posibilita la aplicación de LA sobre GPRS. Al utilizar GPRS una sola modulación, cada CS representa un modo de transporte diferente.

Esquema de codificación	Factor del código	Bits de datos (Carga útil)	Tasa binaria (kbps)
CS-1	1/2	181	9.05
CS-2	$\approx 2/3$	268	13.4
CS-3	$\approx 3/4$	312	15.6
CS-4	1	428	21.4

Tabla 1.- Parámetros de los esquemas de codificación de GPRS

Cabe señalar que las velocidades binarias alcanzables sobre GPRS empleando un único intervalo de tiempo por usuario pueden hacer cuestionable la transmisión de vídeo sobre este sistema. No obstante, el objetivo de este trabajo no es demostrar la viabilidad de la transmisión de vídeo sobre GPRS sino evaluar de forma comparativa los algoritmos estudiados, fin para el cual GPRS resulta adecuado. Por otra parte, el funcionamiento de LA se basa en la calidad del canal. La disponibilidad de interfaces avanzadas entre las simulaciones a nivel de enlace y a nivel de sistema que capturen con precisión la variabilidad del canal resulta esencial para la correcta evaluación del rendimiento de LA [5]. La elección de GPRS para este trabajo se ve apoyada por la disponibilidad de dichas interfaces [5]. Esta investigación se ha centrado en el caso en que un sólo intervalo de tiempo es asignado a cada usuario, por la indisponibilidad de interfaces que modelen la calidad del enlace radio para transmisiones *multislot*. Como quedó demostrado en [6], dicha disponibilidad sería necesaria para el tipo de estudios que aquí se presenta. Nuevamente, un escenario con transmisiones *singleslot* resulta suficiente para determinar el grado de mejora aportado por el algoritmo que se propone.

B. Herramienta de simulación

Para este estudio se ha empleado un simulador de eventos discretos que funciona a nivel de ráfaga (*burst*), lo cual permite el modelado de las variaciones repentinas de la calidad del enlace radio y la obtención de resultados precisos. El sistema modelado consiste en una red macrocelular en la que se estudia el enlace descendente. Las solicitudes de asignación de canal se satisfacen en el orden en que surgen y se resuelven eligiendo de forma aleatoria un canal disponible. Cada usuario mantiene dicho canal hasta que finaliza de forma correcta la transmisión. Aunque se ha implementado la movilidad de usuarios dentro de cada sector, no se han implementado mecanismos de *handover*. Tampoco se han incluido técnicas de control de potencia ni salto en frecuencia por su influencia sobre el funcionamiento de LA. La Tabla 2 resume los principales parámetros de la simulación; puede encontrarse una descripción detallada del simulador en [5].

Parámetro	Valor
Tamaño de clúster	4 celdas
Radio de cada celda	1 km
Sectorización	120°
Interferencia modelada	1ª y 2ª corona cocanal
Número de celdas cocanales	25
Time-slots por sector	16
Usuarios por sector	12
Tipo de tráfico	H.263: 6 usuarios por sector WWW: 3 usuarios por sector Email: 3 usuarios por sector
Modelo de propagación	Okumura-Hata
Modelo de shadowing	Distribución log-normal Desviación estándar 6dB Distancia de decorrelación 20m
Velocidad terminal móvil	50 km/h
Protocolo ARQ	Sólo para WWW e email; sin pérdida ni error en las confirmaciones
Tamaño ventana ARQ	64 bloques RLC
Periodo de solicitud de respuesta ARQ	16 bloques RLC

Tabla 2.- Parámetros del simulador

C. Interfaces entre simulaciones a nivel de enlace y a nivel de sistema

Las simulaciones de sistemas de comunicaciones móviles se suelen desglosar en dos niveles con diferente resolución temporal: nivel de enlace y nivel de sistema. Mientras que a nivel de sistema se modela una red de comunicaciones móviles, los trabajos a nivel de enlace modelan un enlace radio a nivel de bit. La interacción entre ambos niveles tiene lugar mediante el empleo de interfaces. Los resultados de una transmisión a nivel de enlace se representan de forma simplificada mediante un conjunto de tablas que reciben el nombre de *Look-Up Tables* (LUTs). Según las indicaciones de [5], la interfaz empleada en este trabajo se compone de dos tipos de LUTs: LUT-1 y LUT-2. Cada LUT-1 representa una CDF (*Cumulative Distribution Function*) del BER (tasa de bits con error) para un valor concreto de CIR (relación señal a interferencia). A partir de la LUT-1 correspondiente al valor

de CIR que experimenta una ráfaga durante su transmisión en el nivel de sistema, y haciendo uso de un proceso aleatorio, se extrae de la CDF un valor de BER para dicha ráfaga. El interés de este procedimiento radica en la posibilidad de modelar el efecto del desvanecimiento rápido sobre el BER e incluir así sus efectos dentro de la simulación a nivel de sistema. Repitiendo el proceso para las cuatro ráfagas utilizadas para transmitir un bloque RLC se obtienen cuatro valores de BER a partir de los cuales se calcula el valor medio y la desviación estándar del BER. Haciendo uso de la LUT-2 asociada al CS correspondiente, se obtiene a partir de estos dos valores un determinado valor de BLER (tasa de bloques RLC recibidos con error), que se emplea en la simulación a nivel de sistema para determinar si un bloque RLC es recibido o no con error. En [5] se pueden encontrar ejemplos gráficos de ambas LUTs.

D. Modelos de tráfico

Para el estudio realizado se optó por un modelo de tráfico de vídeo H.263 con una tasa binaria media de 16 kbps. Las simulaciones realizadas, no obstante, incluyen dos tipos más de tráfico: navegación web y correo electrónico.

Las fuentes de tráfico web y correo electrónico se implementan mediante un modelo intermitente (ON/OFF) [5] en el que se alternan periodos de actividad (transmisión de información) y periodos de inactividad. Para ambos modelos de tráfico, la transmisión de un nuevo paquete comienza sólo cuando toda la información del paquete previo se ha transmitido sin errores. El tiempo de actividad de estas fuentes dependerá, por tanto, de la calidad del enlace radio.

El modelo de tráfico de vídeo H.263 [7] emplea tres tipos diferentes de tramas de vídeo: I, P y PB. Cada tipo de trama posee unas propiedades estadísticas diferentes que el modelo empleado captura de forma precisa. Las variables y propiedades manejadas por dicho modelo son las siguientes: tamaño y duración de las tramas de vídeo, correlación entre el tamaño de una trama y su duración, y frecuencia de transición entre los diversos tipos de tramas de vídeo. El modelo se implementa en dos niveles. El primero de ellos determina el tipo de trama a generar en cada momento. Las tramas I se generan a intervalos de tiempo regulares, mientras que la generación de tramas P y PB se determina a partir de una cadena de Markov. Una vez establecido el tipo de trama, el segundo nivel de modelado determina su tamaño y su duración.

III. ALGORITMOS LINK ADAPTATION

El criterio aplicado para la elección del CS óptimo en función de la calidad experimentada en el enlace radio es un factor decisivo en el comportamiento de LA. En esta sección se presentan dos algoritmos LA que difieren en dicho criterio.

A. Algoritmo LA basado en throughput

Este algoritmo, propuesto en [1] y comúnmente utilizado en estudios que involucran el uso de LA, tiene como principal objetivo maximizar la tasa de transmisión global del sistema. Por lo tanto, en cada nueva decisión tomada por el

algoritmo se considerará óptimo aquel CS que maximice el valor del *throughput*, definido según la expresión

$$\text{Throughput} = R_{CS-i} \cdot (1 - \text{BLER}_{CS-i}) \quad (1)$$

donde R_{CS-i} y BLER_{CS-i} representan respectivamente, la tasa binaria y el BLER para el CS considerado. Teniendo en cuenta las interfaces empleadas entre la simulación a nivel de enlace y a nivel de sistema, la expresión (1) proporciona para cada CS, una curva tridimensional del *throughput* en función de la calidad del enlace expresada en términos del valor medio del BER y su desviación estándar. Estas curvas se cortan entre sí en los denominados umbrales de decisión, poniendo de relieve la existencia de regiones en las que cada CS es capaz de proporcionar un *throughput* superior al ofrecido por los demás esquemas de codificación.

B. Algoritmo LA propuesto

En contraste con el algoritmo anterior, el algoritmo propuesto en este trabajo se diseña considerando no sólo el *throughput* sino también el retardo, parámetro de vital importancia para servicios en tiempo real. En el diseño del algoritmo se define el *retardo de trama de vídeo* como sigue. Sea una trama de vídeo generada en el instante t_0 con un tamaño T y una duración D , de tal forma que la siguiente trama de vídeo se genera en $t_0 + D$. La transmisión de la trama generada en t_0 finalizará en el instante $t_0 + D + \delta$, donde δ representa el retardo anteriormente citado. Observar que el valor de δ repercute directamente sobre la calidad de un servicio en tiempo real: si $\delta < 0$, existe un margen de tiempo antes de $t_0 + D$ durante el cual el canal puede ser utilizado por otros usuarios, mientras que si $\delta > 0$, la transmisión de una trama de vídeo se solapa con la siguiente, con la consecuente degradación de la calidad percibida por el usuario (retardo o pérdida de datos). Para los servicios de tiempo real (H.263), no se considera el uso de ARQ. Bajo esta condición, el retardo de trama δ depende, no del *throughput* experimentado durante su transmisión, sino de la tasa binaria R_{CS-i} de los CSs empleados. Para un determinado CS, el retardo vendrá dado por la expresión

$$\text{Retardo de trama} = \frac{T}{R_{CS-i}} - D \quad (2)$$

donde se aprecia que una reducción del retardo de trama requiere el empleo de CSs con menor protección ante errores.

Cuando $\delta > 0$, interesa que el algoritmo minimice el valor de δ , mientras que si $\delta < 0$ entonces interesa que el algoritmo maximice $|\delta|$. Teniendo en cuenta que el retardo de trama δ no depende del *throughput* y que el objetivo perseguido consiste en maximizar el *throughput* al mismo tiempo que se minimiza el retardo de trama, interesa que el algoritmo maximice (3a) cuando $\delta > 0$ o que maximice (3b) cuando $\delta < 0$.

$$\frac{\text{Throughput}}{\text{Retardo}} = \frac{\text{Throughput} \cdot R_{CS-i}}{T - D \cdot R_{CS-i}} \quad (3a)$$

$$\text{Throughput} \cdot |\text{Retardo}| = \text{Throughput} \cdot \left(D - \frac{T}{R_{CS-i}} \right) \quad (3b)$$

El algoritmo, por lo tanto, se define como sigue. Cada vez que se genera una nueva trama de vídeo el algoritmo calcula, para cada CS, el retardo de trama a partir de la expresión (2). Dependiendo de si dicho retardo es mayor o menor que cero se evalúa, respectivamente, la expresión (3a) o (3b) de forma dinámica. El algoritmo considerará óptimo aquel CS que proporcione el valor más elevado resultante de dicha evaluación, ya que ese será el CS capaz de ofrecer el mayor *throughput* y el menor retardo. Una característica interesante y deseable del algoritmo es que las tramas de vídeo con mayor tamaño y/o menor tiempo entre tramas fomentan un mayor grado de utilización de aquellos CSs con mayor tasa binaria, hecho que favorece la reducción del retardo a expensas de un riesgo más elevado de sufrir errores. Como posible contrapartida, la necesidad de realizar cálculos para cada trama de vídeo transmitida acarrea un coste adicional. No obstante, dicho coste se reduce significativamente en transmisiones descendentes (*downlink*), ya que el cómputo se puede realizar íntegramente en la estación base. En cualquier caso, y como se verá en el siguiente apartado, esta posible desventaja se ve compensada por las mejoras que pueden llegar a obtenerse al aplicar el presente algoritmo.

IV. RENDIMIENTO DEL ALGORITMO

A. Parámetros de calidad

Las prestaciones de los dos algoritmos LA se compararán principalmente mediante los parámetros de *throughput* y retardo. Su valor medio es un parámetro de importancia; no obstante, también se considerará el *throughput* mínimo y el retardo máximo experimentados por el 95% e incluso el 99% de las muestras analizadas, ya que constituyen indicaciones más representativas de la calidad de servicio experimentada por los usuarios. El *throughput* se define como la cantidad de bits correctamente transmitidos por unidad de tiempo. También se empleará el retardo de transmisión normalizado, definido como el cociente entre el tiempo necesario para transmitir un bloque de datos y el tamaño de dicho bloque. Otros parámetros útiles son el porcentaje de ocasiones en que se emplea el CS óptimo o el número de cambios de CS por unidad de tiempo, valor representativo de la carga de señalización asociada al uso de LA.

La calidad de vídeo se evalúa a partir del BLER y BER de las tramas de vídeo recibidas. Según [8], un BLER por debajo del 5% no produce un deterioro perceptible por el ojo humano para transmisión de vídeo H.263. En cambio, en [9] se especifican requisitos más exigentes. Concretamente, el BLER máximo aceptable en videotelefonía es del 1%, siendo del 2% para aplicaciones de vídeo en tiempo real o de tipo *streaming*. Como referencia adicional, también se considerarán las indicaciones dadas en [10], según las cuales un BER igual a 10^{-6} no produce degradación visible, un BER igual a 10^{-5} provoca una distorsión ligeramente perceptible, un BER igual a 10^{-4} da lugar a artefactos visibles y un BER superior a 10^{-3} limita la aplicación práctica.

B. Resultados de la simulación

Según se puede apreciar en la Tabla 3, el algoritmo propuesto ofrece un *throughput* superior al ofrecido por el pro-

pio algoritmo basado en *throughput*. La mejora obtenida resulta especialmente significativa para aquellos parámetros más restrictivos. La mejora relativa es del 2.3% en valor medio y del 8.3% y 7.4% en el *throughput* mínimo experimentado por el 95% y 99% de las muestras, respectivamente. Este hecho es debido a que no sólo se utilizan con mayor frecuencia los CSs con mayor velocidad binaria sino que, además, el porcentaje de ocasiones en que la elección de los codificadores resulta ser óptima es también mayor, tal y como puede observarse en la Tabla 4. Otro parámetro, indicador del mejor comportamiento del algoritmo propuesto, es la frecuencia con la que se producen cambios en el CS usado. Mientras que el algoritmo basado en *throughput* requiere 4.69 cambios de CS por segundo, el algoritmo propuesto solicita tan sólo 3.94, lo que implica una reducción del 16% sobre la carga de señalización asociada al uso de LA.

Las prestaciones de retardo que ofrece el algoritmo basado en *throughput* también se ven mejoradas por el algoritmo propuesto, según se muestra en la Tabla 3. El retardo normalizado medio se reduce en un 5%, mientras que la mejora apreciada para el retardo normalizado máximo experimentado por el 95% de las muestras es del 14%. Esta reducción del retardo normalizado es consecuencia directa del mayor porcentaje de utilización de los CSs con mayor velocidad binaria. Como resultado y según se muestra en la Tabla 5, el algoritmo propuesto es capaz de transmitir hasta casi un 5% más de tramas de vídeo sin retardo, es decir, de tramas de vídeo cuya transmisión finaliza por completo antes de que se genere la siguiente trama de vídeo a transmitir.

Por último, el mayor grado de utilización de CSs con menor protección ante errores origina un incremento del BLER en las tramas de vídeo recibidas, según se observa en la Tabla 3. A pesar de ello, la Tabla 6 pone de manifiesto que el algoritmo propuesto consigue transmitir un mayor porcentaje de tramas de vídeo sin retardo que satisfacen al mismo tiempo los requisitos de calidad señalados en [8], [9] y [10].

Parámetro	Algoritmo	Medio	95%	99%
Throughput (kbps)	Throughput	16.56	10.46	7.53
	Propuesto	16.94	11.33	8.09
BLER (%)	Throughput	11.18	25.90	39.73
	Propuesto	12.20	28.91	42.71
Retardo normalizado (ms/kbit)	Throughput	58.32	99.08	129.44
	Propuesto	55.36	85.32	119.59

Tabla 3.- Comparación del rendimiento de los dos algoritmos estudiados

Algoritmo	CS-1	CS-2	CS-3	CS-4	CS óptimo
Throughput	6.63	7.14	26.41	59.82	65.99
Propuesto	3.37	2.92	27.98	65.73	71.17

Tabla 4.- Porcentaje de utilización de cada CS y del CS óptimo

Algoritmo	Sin retardo	Con retardo
Throughput	69.16	30.84
Propuesto	73.92	26.08

Tabla 5.- Porcentaje de tramas de vídeo transmitidas con y sin retardo

	Throughput	Propuesto
BLER ≤ 1%	44.73	46.22
BLER ≤ 2%	44.75	46.24
BLER ≤ 5%	44.98	46.45
BER ≤ 10 ⁻³	45.93	48.10
BER ≤ 10 ⁻⁴	40.40	42.14
BER ≤ 10 ⁻⁵	39.33	41.02
BER ≤ 10 ⁻⁶	39.23	40.92

Tabla 6.- Porcentaje de tramas de vídeo transmitidas sin retardo y que cumplen los diversos requisitos de calidad

V. CONCLUSIONES

En este trabajo se ha presentado el diseño de un nuevo algoritmo LA pensado para mejorar la calidad de servicio en aplicaciones de tiempo real, con estrictos requisitos de retardo. El criterio seguido por el algoritmo para la elección del modo de transporte óptimo se basa no sólo en el *throughput* sino también en el retardo experimentado por la información transmitida. El algoritmo se ha evaluado con tráfico de vídeo H.263 y su rendimiento se ha comparado con el del algoritmo LA tradicional basado en la maximización del *throughput*. Los resultados obtenidos muestran que el algoritmo propuesto mejora no sólo las prestaciones de retardo ofrecidas por dicho algoritmo sino también las prestaciones de *throughput*.

AGRADECIMIENTOS

Los autores quieren agradecer el apoyo económico de Banca-UMH para la realización de este trabajo.

REFERENCIAS

- [1] ETSI-SMG, "EDGE Feasibility Study, Work Item 184; Improved Data Rates through Optimized Modulation", Tdoc 97-331, Diciembre 1997.
- [2] K.K. Leung et al., "Link Adaptation and Power Control for Streaming Services in EGPRS Wireless Networks", *IEEE Journal on Selected Areas in Communications*, vol. 19, n° 10, pp. 2029-2039, Oct. 2001.
- [3] C. Kodikara et al., "Performance Improvement for Real-Time Video Communications by Link Adaptation in EGPRS Networks", *Proceedings of the IEE 3G Mobile Communications Technologies Conference*, pp. 489-494, Mayo 2002.
- [4] W. Luo et al., "Packet Size Dependent Link Adaptation for Wireless Packet Data", *Proceedings of IEEE Globecom*, pp 53-56, Noviembre 2000.
- [5] J. Gozávez y J. Dunlop, "Link Level Modelling Techniques for Analysing the Configuration of Link Adaptation Algorithms in Mobile Radio Networks", *Proceedings of European Wireless 2004*, pp 325-330, Feb. 2004.
- [6] J. Gozávez y J. Dunlop, "On the Effect of Correlation in Multislot Link Layer Analysis for GPRS", *Proceedings of IEEE VTC-Fall 2000*, pp. 444-450, Septiembre 2000.
- [7] O. Lázaro et al., "H.263 Video Traffic Modelling for Low Bit Rate Wireless Communications", *Aceptado para su publicación en PIMRC2004*, Sept. 2004.
- [8] L. Hanzo et al., "Wireless Video Communications: Second to Third Generation Systems and Beyond", *IEEE Press*, 2001.
- [9] 3GPP TS 22.105, v. 6.2.0, "TSG Services and System Aspects; Service Aspects; Services and Service Capabilities".
- [10] 3GPP TS 23.107, v.5.2.0, "Technical Specification Group Services and System Aspects; QoS Concept and Architecture".