



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN

Análisis del Comportamiento Dinámico de Redes
Inalámbricas de Banda Ancha Fijas Basadas en el
Estándar IEEE 802.16

Tesis para obtener el grado de

**Maestro en Ingeniería
(Computación)**

Presenta

Ing. Javier Chapa López

Director de Tesis

Dr. Víctor Rangel Licea

Ciudad Universitaria, México D.F.

Mayo 2009



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A mis padres...

....porque gracias a ellos existo, por darme los valores que ahora me permiten ser la persona que soy, por su invaluable cariño, por su amor incondicional, por todo.....esto es, por y gracias a ustedes.

A la Universidad Nacional Autónoma de México...

....por la gran formación humana y profesional que ha dejado en mí desde que era un adolescente.

A mi familia...

....por su entrañable solidaridad y apoyo.

A mis amigos...

....por ser mi soporte emocional

A mi tutor...

....por sus enseñanzas, por su paciencia y sobre todo por la confianza deposita en mí.

A la DGAPA-UNAM...

*....por el apoyo otorgado a los proyectos de investigación **PAPIIT:IN104907** “Diseño de técnicas de mejoramiento de capacidad en redes inalámbricas de banda ancha tipo Mesh” y **PAPIME:PE103807** “Integración de nuevas tecnologías y elaboración de material didáctico para un laboratorio multidisciplinario de redes de la Facultad de Ingeniería”.*

Al Consejo Nacional de Ciencia y Tecnología....

....por el apoyo durante mis estudios de posgrado.



Tabla de Contenido

Capítulo 1-Descripción de la Tesis

- 1.1 Antecedentes – La necesidad de la comunicación inalámbrica
- 1.2 Objetivos
- 1.2 Definición del Problema
- 1.3 Metodología
- 1.4 Contribución y Relevancia
- 1.5 Estructura de la Tesis

Capítulo 2-Redes Inalámbricas de Banda Ancha

- 2.1 Redes Inalámbricas y Redes Inalámbrica de Banda Ancha
- 2.2 WiMAX o 3G...?
 - 2.2.1 Ventajas de un Sistema Celular 3G
 - 2.2.2 Ventajas de un Sistema WiMAX
- 2.3 El Estándar IEEE 802.16
 - 2.3.1 Transición de 802.16-2004 a 802.16e 2
- 2.4 El Foro WiMAX
- 2.5 Certificación de Productos WiMAX
 - 2.5.1 Productos WiMAX certificados

Capítulo 3-Conceptos Básicos y Definiciones

- 3.1 Bandas de Frecuencia
 - 3.1.1 Canales
 - 3.1.2 Tipos de Espectro Radioeléctrico
- 3.2 Tipos de Redes Inalámbricas
 - 3.2.1 Redes Fijas y Móviles
- 3.3 La Capa Física (PHY)
 - 3.3.1 Forward Error Correction (FEC)
 - 3.3.2 Unidades de Datos
- 3.4 Calidad de Servicio (QoS)
 - 3.4.1 Métricas de QoS
 - 3.4.2 Funciones de QoS
 - 3.4.3 Tipos de QoS en el Estándar IEEE 802.16
- 3.5 Arquitectura
- 3.6 Modelo de Referencia
 - 3.6.1 Subcapa de Convergencia (CS)
 - 3.6.2 Subcapa Común de la MAC (MAC CPS)
- 3.7 Tramas y Multiplexado
 - 3.7.1 Ranuras Físicas (Slots)
 - 3.7.2 Tramas TDD
- 3.8 Topologías

Capítulo 4-Modelo Matemático

- 4.1 Antecedentes
 - 4.1.1 ANCHO DE BANDA – Solicitud / Asignación
 - 4.1.1.1 Requerimientos de Ancho de Banda en el UL-Incremental y Agregado
 - 4.1.1.2 Requerimientos de Ancho de Banda en el UL–Autónomo y Piggyback
 - 4.1.2 Mecanismos de Acceso al UL (Solicitud /Asignación)
 - 4.1.2.1 Polling Unicast
 - 4.1.2.2 Polling Basado en Contención (Polling Multicast o Broadcast)
 - 4.1.3.3 Algoritmo para resolver contenciones
 - 4.1.4 Scheduling en el Modelo
- 4.2 Modelo Analítico
 - 4.2.1 Describiendo la Estrategia
 - 4.2.2 Modelo Analítico



- 4.2.2.1 Calculando el Retardo de Contención
- 4.2.2.2 Integrando y generalizando el modelo para el retardo de contención
- 4.2.2.3. Calculando el retardo de espera

Capítulo 5-Resultados Obtenidos

- 5.1 Escenarios Teóricos
 - 5.1.1 Análisis de Resultados Teóricos para un UL-Frame = 1 ms
 - 5.1.1.1 Resultados con paquetes de tamaño fijo (300 bytes)
 - 5.1.1.2 Resultados con paquetes de distinto tamaño, con C=30
 - 5.1.1.3 Resultados variando el tamaño de los paquetes y también C
 - 5.1.2 Análisis de Resultados Teóricos para un UL-Frame=2 ms
 - 5.1.3 Análisis de Resultados de Simulación para un UL-Frame=2 ms
 - 5.1.3.1 Modelo de simulación
 - 5.1.3.2 Comparación y análisis de eficiencia
- 5.2 Comparación de Modelos (GPRS vs IEEE 802.16-2004)
 - 5.2.1 Describiendo El Modelo Para GPRS
 - 5.2.2 Diferencia Analítica en los Modelos
 - 5.2.2 Diferencia Experimental en los Modelos

Conclusiones

Referencias Bibliográficas

Apéndice A-Glosario de Términos

Apéndice B-Programa para Determinar la Eficiencia (Throughput), Modelo con Ecuaciones en Diferencias

Apéndice C-Programa para Determinar el Tamaño de un Bloque en la Red IEEE 802.16 96

Apéndice D-WHITE PAPER: Performance Analysis of Best Effort Support in Broadband IEEE 802.16 Networks

Apéndice E-Programa para Determinar la Eficiencia (Throughput), Modelo con Probabilidad Binomial (usado en GPRS)

Índice de Tablas

Tabla 2.1 – Algunos elementos de comparación entre los más famosos sistemas de comunicación inalámbrica.

Tabla 2.2 – Principales documentos IEEE 802.16 publicados.

Tabla 2.3 – Principales miembros del foro WiMAX.

Tabla 3.1 – Las 5 interfaces físicas definidas en el estándar IEEE 802.16. Tomado de la referencia [13].

Tabla 4.1– Componentes que forma el retardo en la transmisión de un solo paquete.

Tabla 4.2– Notación completa de elementos componentes de nuestro modelo matemático.

Tabla 5.1– Parámetros utilizados para obtener los resultados teóricos con un UL-Frame de 1ms.

Tabla 5.2 – Resultados numéricos completos que generan la Figura 5.1.

Tabla 5.3 – Resultados numéricos que generan las Figuras 5.2 (a) y 5.2 (b).

Tabla 5.4 –Equivalencias para el eje “y” de la Figura 5.6.

Tabla 5.5– Parámetros que cambiarán en relación a la Tabla 5.1 para obtener los resultados teóricos con un UL-Frame de 2ms.

Tabla 5.6– Parámetros de simulación utilizados en OPNET.

Tabla 5.7– Parámetros utilizados para obtener los resultados teóricos con un UL-Frame de 1ms.

Tabla 5.8 – Eficiencia (expresada en porcentaje) del canal de subida para ambos modelos.

Tabla 5.9 – Cuadro comparativo entre los modelos.

Índice de Figuras

Figura 1.1 – Estructura del formato del canal de subida (UL para redes basadas en IEEE 802.16-2004.

Figura 1.2 –Escenario propuesto con las ranuras de contención variando dinámicamente en función a los requerimientos de los usuarios de la red.

Figura 2.1 – Tipos de redes inalámbricas, en cada caso se menciona la tecnología más conocida.

Figura 2.2 –Comparación de los principales sistemas inalámbricos.



- Figura 3.1 – Ejemplo de cómo funcionan las PDUs, SDUs y el SAP en la pila de protocolos.
- Figura 3.2 – Mecanismo de programación de QoS de una EB en donde se muestran la asignación de las aplicaciones a su respectiva clase de servicio. Imagen tomada de IEEE 802.16-2004.
- Figura 3.3 – Modelo de referencia para el estándar IEEE 802.16.
- Figura 3.4 – Formato y tamaño de tramas PHY en el estándar IEEE 802.16.
- Figura 3.5 – Estructura de una trama TDD en una conexión punto-multipunto.
- Figura 3.6 – Topología PMP, en donde la EB cubre sus ESs.
- Figura 3.7 – Topología de Malla (mesh) donde la EB no es el único centro de distribución de tráfico.
- Figura 4.1– Estructura del Grant Management Subheader en un MAC-PDU. Para QoS=UGS.
- Figura 4.2– Estructura del Grant Management Subheader en un MAC-PDU . Para QoS \neq UGS.
- Figura 4.3– Ejemplo del mecanismo de polling unicast.
- Figura 4.4– Ejemplo de un grupo de ESs con polling basado en contención.
- Figura 4.5–Estructura de los canales de subida y bajada (UL y DL).
- Figura 4.6–Componentes del retardo de acceso al medio para ESx.
- Figura 5.1 – Eficiencia del sistema en función del número de ESs activas en la red con UL-Frame= 225 minislots
- Figura 5.2 – Retardo (a -izquierda) en Minislots y (b-derecha)en milisegundos entre cada ciclo de transmisión en función del número del número de ESs activas.
- Figura 5.3 –Eficiencia del sistema IEEE 802.16 utilizando el modelo propuesto variando el tamaño de los paquetes transmitidos.
- Figura 5.4 –Retardo del sistema IEEE 802.16 utilizando el modelo propuesto variando el tamaño de los paquetes transmitidos.
- Figura 5.5 –Eficiencia del sistema IEEE 802.16 utilizando el modelo propuesto variando el tamaño de los paquetes transmitidos.
- Figura 5.6 –Eficiencia del sistema IEEE 802.16 utilizando el modelo propuesto variando el tamaño de los paquetes transmitidos y los slots de contención en función a cada tamaño de paquete.
- Figura 5.7 –Eficiencia del sistema IEEE 802.16 utilizando el modelo propuesto variando el tamaño de los paquetes transmitidos y los slots de contención en función a cada tamaño.
- Figura 5.8 –Comparación de escenarios con UL-MAPs de 225 y 450 minislots.
- Figura 5.9 –Retardo (a -izquierda) en Minislots y (b-derecha)en milisegundos entre cada ciclo de transmisión en función del número de ESs activas. Comparación con el UL-MAP de 1 y 2 ms, usando los mismos slots de contención para ambos casos, se fija el tamaño de los paquetes en 300 bytes.
- Figura 5.10 –Eficiencia del sistema IEEE 802.16 utilizando el modelo propuesto variando el tamaño de los paquetes transmitidos. usando un UL-MAP de 450 minislots.
- Figura 5.11 –Modelo jerárquico de simulación.
- Figura 5.12 –Resultados de la eficiencia del ancho de banda de sistema usando el modelo teórico y la simulación.
- Figura 5.13 –Comparación de eficiencia entre modelos matemáticos iterativos (Ecs. en Diferencias vs Probabilidad Binomial)



CAPÍTULO 1

DESCRIPCIÓN DE LA TESIS

1.1 ANTECEDENTES – LA NECESIDAD DE LA COMUNICACIÓN INALÁMBRICA

Desde la década final del siglo XX las redes de datos han tenido un gran crecimiento. Después de la instalación de redes fijas con acceso a Internet en muchos lugares alrededor del mundo, esta misma necesidad de crecimiento empezó a ser más importante para las redes inalámbricas. Por lo tanto, nos damos cuenta de que a finales de la primera década del siglo XXI los accesos inalámbricos de alta velocidad (Mb/s) están siendo desarrollados ampliamente en todo el **planeta***.

La historia de las comunicaciones inalámbricas se remonta a fines del siglo XIX cuando la ecuación de Maxwell mostraba que la transmisión de información podría llevarse a cabo sin la necesidad de un cable.

Pocos años después, experimentos como el de Marconi probaron que las comunicaciones inalámbricas podían ser una realidad y además alcanzar grandes distancias. Durante el transcurso del siglo XX grandes descubrimientos electrónicos y de propagación generaron ingeniosos inventos dando origen a varios sistemas de transmisión inalámbrica.

En la década de 1970, en los Laboratorios Bell (*Bell Labs*) se propuso el concepto de células aplicado a comunicaciones inalámbricas, una mágica idea que permite la cobertura de una zona tan grande como sea necesario usando un número limitado y fijo de frecuencias. Desde entonces, muchas tecnologías inalámbricas tuvieron gran utilización, siendo la más **exitosa**** hasta ahora **GSM (*Global System for Mobile Communications*)¹**, originalmente fue la segunda generación de los sistemas celulares europeos. GSM es una tecnología utilizada principalmente para transmitir voz y adicionalmente para datos (pero a tasas bajas), por ejemplo los mensajes SMS (*Short Message Service*).

La evolución de GSM ha sido tal que actualmente se utiliza mayoritariamente en muchos países, incluido México. Esta evolución está destinada a facilitar la transmisión de datos a relativamente alta velocidad en este tipo de redes, siguiendo esta directriz, las principales evoluciones de esta tecnología son:

* **Fuente:** Centro de Investigación e Innovación en Telecomunicaciones, Artículo 03-Febrero-2005 por Irazú Muñiz.

** **Fuente:** OECD (Organisation for Economic Co-Operation and Development), ver apartado en las Referencias Bibliográficas.

¹ Todas las abreviaciones técnicas serán incluidas en el glosario de términos, en el Apéndice A al final de este documento.



- GPRS – *General Packet Radio Service*, la evolución de “paquetes conmutados” de GSM.
- EDGE – *Enhanced Data rates for GSM Evolution*, la cual incluye una adaptación para dar eficiencia a través de modulación digital.

En adición a GSM, surgió una tercera generación (3G) de sistemas celulares, originalmente europea y japonesa, la tecnología UMTS (*Universal Mobile Telecommunications Systems*) y la norteamericana CDMA2000 (*Code Division Multiple Access Version 2000*), están listas y desarrolladas, y son dos prometedores sistemas de comunicación inalámbrica.

En paralelo surgió una nueva tecnología inalámbrica de banda ancha, basada en el estándar IEEE (*Institute of Electrical and Electronics Engineers*) 802.16, la cual nació con el fin de transmitir datos (no precisamente voz) a altas velocidades. Con el fin de impulsar el desarrollo de esta nueva tecnología y de igual forma que para el estándar IEEE 802.11 fue creado un foro técnico para certificar productos basados en IEEE 802.16, denominado WiMAX (*Worldwide Interoperability for Microwave Access Forum*) y representa uno de los más prometedores sistemas de **comunicación*****.

Grandes esperanzas e importantes inversiones han sido hechas con el fin de desarrollarla, el cual como un sistema de acceso inalámbrico de banda ancha tiene muchas aplicaciones: acceso de última milla fijo o inalámbrico, redes celulares, telemetría, etc.

El análisis del comportamiento dinámico de una red inalámbrica de banda ancha basada precisamente en el estándar IEEE 802.16 es el motivo de la elaboración de este documento técnico. Una red inalámbrica de banda ancha de este tipo, puede ofrecer mayores tasas de transmisión, sin sacrificar la calidad de servicio. Lo cual significa una enorme ventaja sobre lo que actualmente se ha desarrollado para las redes celulares comerciales.

De aquí, que al proponer soluciones eficientes para la correcta asignación de los recursos (ancho de banda) del sistema IEEE 802.16 y distribuirlos equitativamente en función de la calidad de servicio contratada por los usuarios, este tipo de redes se convertirán en un futuro éxito comercial. Para materializar una verdadera solución eficiente, una de las ideas principales es hacer recomendaciones para tener una correcta utilización de la zona de contención sobre un sistema inalámbrico de banda ancha.

Con esto prestar un servicio superior al que se provee actualmente en forma comercial (3G) a los usuarios con una clase de servicio baja (sin prioridad) y consecuentemente tener un aprovechamiento óptimo de los recursos inalámbricos disponibles.

*** Fuente: OECD (Organisation for Economic Co-Operation and Development), ver apartado en las Referencias Bibliográficas.

1.2 OBJETIVOS

El principal objetivo de la elaboración de este trabajo es la de utilizar un modelo matemático, más sencillo que lo hasta ahora desarrollado con probabilidades binomiales (sobre GPRS, detalladamente descrito en las referencias [1] y [2]). La idea es, generar distintos escenarios teóricos en una red basada en el estándar IEEE 802.16-2004, mismos que podrán ser evaluados dinámicamente (vía simulación).

Para la simulación se podría utilizar *OPNET Modeler®*, el único software autorizado por el WiMAX Forum® para la certificación de equipos. OPNET es un lenguaje de simulación orientado a comunicaciones, utiliza un modelo de simulación jerárquico. El software es desarrollado por la compañía *OPNET Technologies, Inc* (NASDAQ:OPNT).

La Facultad de Ingeniería es, desde Noviembre del 2006 miembro distinguido del *WiMAX Model Development Consortium de OPNET Technologies, Inc.*, gracias a los trabajos de investigación producidos en la UNAM como mostrados en las referencias [3], [4] y [5] que han sido presentados en diversas conferencias internacionales.

Mediremos parámetros como: ancho de banda (*throughput*), retardo (*queue delay*), rápida solución de colisiones en las ranuras (*slots*) de contención en una red con el número de usuarios variable y además con tráfico de tipo CBR (*Constant Bit Rate*), todo esto para usuarios que usan el tipo de servicio **BE²** (*Best Effort*).

Basados en trabajos anteriores, y que fueron realizados para GPRS, pero cambiando el procedimiento para obtener valores de los componentes del retardo para que las ESs (Estación Subscriptora) accedan al medio inalámbrico, podemos determinar el segundo objetivo.

Nuestra intención es la de generar variantes en donde se toman en cuenta aspectos más específicos como analizar el comportamiento dinámico de una red IEEE 802.16-2004 en donde los parámetros mencionados presenten cambios en el tiempo. Una vez generados todos los escenarios posibles se procederá con un análisis detallado de las consecuencias de modificar otros parámetros del nuevo modelo (referencia [3]), tales como, número de *slots* de contención, tamaño de los paquetes que los usuarios transmiten, tamaño del **UL-FRAME³**, usuarios activos en la red, etc.

² **BE - Best Effort** – Es uno de los 5 tipos de calidad de servicio que utiliza el estándar IEEE 802.16 y se usa para aplicaciones que no son de tiempo real, por ejemplo, la navegación en Internet (HTTP) y el correo electrónico. Los usuarios contienden por el ancho de banda.

³ **UL-FRAME** – Es el mapa de múltiple acceso que utilizan los usuarios móviles para transmitir información a la estación base.

Con esto podremos hacer recomendaciones sobre el cómo procesar de la forma más eficiente a los usuarios que compiten por ancho de banda, y así dar una opinión más exacta sobre los parámetros de configuración ideales en una red basada en el estándar IEEE 802.16-2004.

1.2 DEFINICIÓN DEL PROBLEMA

La principal área de oportunidad que nos brindan las redes IEEE 802.16 es, que el alcance del estándar no define como forzoso muchos detalles técnicos, por ejemplo las siguientes cuatro funciones para que una red inalámbrica de banda ancha soporte el concepto de calidad de servicio (QoS- *Quality of Service*):

- El control de admisión de usuarios, así como la reservación y asignación de ancho de banda (tema de nuestro análisis).
- Orden de transmisión de los paquetes en función al tipo de clase de servicio (QoS). Se utilizan algoritmos conocidos como *"Traffic Scheduling"*.
- Clasificación de tráfico.
- Técnicas para monitorear y organizar el tráfico entrante en la red. Los algoritmos utilizados se denominan de *"shaping"* y *"policing"*.

Estas faltas de definición de una regla general hacen la diferencia al momento de desarrollar un algoritmo e implementarlo en un equipo que cumpla con la certificación *WiMAX*.

Esto nos da la pauta de poder aportar ideas y procedimientos con el fin de hacer eficaz aquellas partes que aún no están completamente definidas.

En el caso particular de controlar el acceso al medio (manejo de colisiones de los usuarios que tienen que competir por ancho de banda), como hemos mencionado, el estándar IEEE 802.16 no especifica cómo, ni cuál algoritmo de control de admisión sobre nuevas peticiones de ancho de banda debería implementarse.

Los sistemas basados en este estándar se caracterizan por tener una configuración punto-multipunto (PMP), en donde una sola EB (Estación Base) provee de una conexión (con sus respectivos servicios) a un conjunto de ESs. Por lo tanto varias ESs podrían estar compitiendo por obtener los servicios de dicha conexión al mismo tiempo, por eso la relevancia de tener un buen algoritmo para resolver rápidamente las colisiones que pudieran generarse.

La resolución de colisiones se considera para el caso de aquellos usuarios con una clase de servicio (QoS) no periódica, es decir, en aquellas clases de servicio que no tienen previamente reservada una región para transmitir, por ejemplo BE.

Este tipo de usuarios de la red, tienen que entrar en contención y competir con otras ESs que utilicen esta misma clase de servicio por aquellas ranuras en las cuales le indicarán a la EB cuánto es el ancho de banda que requieren para poder transmitir sus datos.

En la actualidad el uso de la región de contención en el estándar IEEE 802.16 no se explota de forma óptima, sólo se utiliza el esquema general “estático”, en el que siempre hay las mismas ranuras asignadas para la contención, no importando el número de usuarios en la red. En una red IEEE 802.16 una colisión puede ocurrir en una de dos situaciones, durante el proceso de *ranging*⁴ inicial o en los intervalos en los que se solicita ancho de banda en el UL (Up Link-Canal de Subida).

El algoritmo utilizado generalmente para la administración de las probables colisiones es el mismo para ambas situaciones. El método está basado en el algoritmo TBEB (*Truncated Binary Exponential Backoff*), en el cual los valores para la ventana inicial y el máximo valor de la misma los determina la estación base.

En la Figura 1.1 se muestra la estructura general del UL que es utilizada en el estándar IEEE 802.16-2004, en donde podemos observar que las ranuras asignadas para la contención de usuarios con una clase de servicio no periódica (como *BE* o *nrtPS*) es fija.

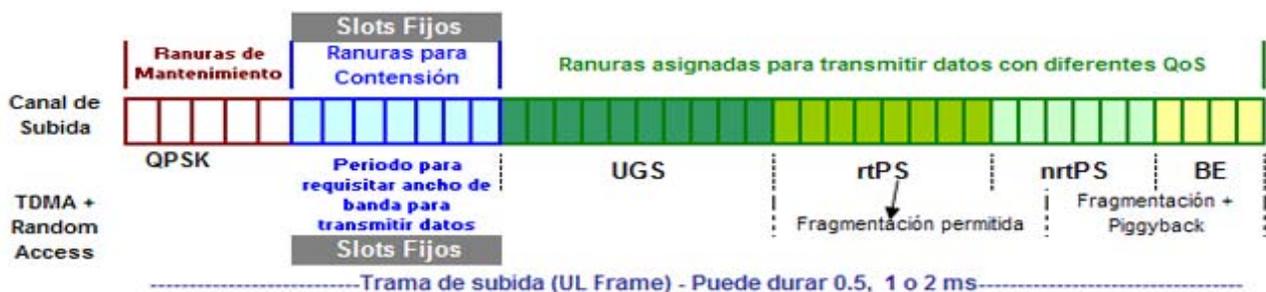


Figura 1.1 – Estructura del formato del canal de subida (Up Link) para redes basadas en IEEE 802.16-2004.

No importa el número de usuarios que estén conteniendo en dicha región para obtener ancho de banda para transmitir datos, lo cual evidentemente representa una seria deficiencia en el manejo de las colisiones durante periodos de alta utilización de la red. Lo anterior, redundará en lentitud dado el número de colisiones y el tiempo que tendrán que esperar los usuarios de estas clases de servicio para poder transmitir datos.

⁴ *Ranging* – Definido como el proceso que permite a una estación subscriptora adquirir los valores correctos de propagación y frecuencia, en adición se ajustan los valores de potencia y además se define el perfil a utilizar para el canal de bajada.



El UL utiliza la técnica de multiplexaje por división de tiempo TDMA (*Time Division Multiple Access*), combinado con un esquema dinámico de acceso aleatorio DAMA (*Demand Assigned Multiple Access*). Como hemos mencionado, el UL incluye los tiempos de reservación de ancho de banda (región de contención).

Nuestro análisis consistirá de un modelo teórico en el que se tratarán varios tamaños o duración de la región de contención, intercambiables dinámicamente en función del número de usuarios utilizando la red con clases de servicio no periódicas en donde no se implemente el algoritmo TBEB.

Dejando la puerta abierta para la utilización del lenguaje de simulación orientado a comunicaciones OPNET para demostrar (vía simulación) que nuestro análisis ayudará a implementar una mejora significativa en la eficacia de la red en periodos de alto tráfico.

El comportamiento dinámico de las ranuras de contención se basará en la cantidad de usuarios en la red así como en la clase de servicio requerida por cada uno de esos usuarios.

Para la parte teórica nos apoyaremos en varios archivos generados a través de varias hojas de cálculo y del lenguaje de programación Visual C++.

En el Capítulo 6 se muestra en forma detallada como se ve afectado la eficiencia del canal inalámbrico de nuestra red IEEE 802.16 cuando vamos variando la duración del UL-Frame y la parte de contención. Se obtuvieron resultados muy interesantes con este respecto.

En la Figura 1.2 se muestra un ejemplo ilustrativo de cómo se modificarán la longitud o duración de la parte asignada a la contención en el UL-Frame para los usuarios bajo el esquema BE en una red.

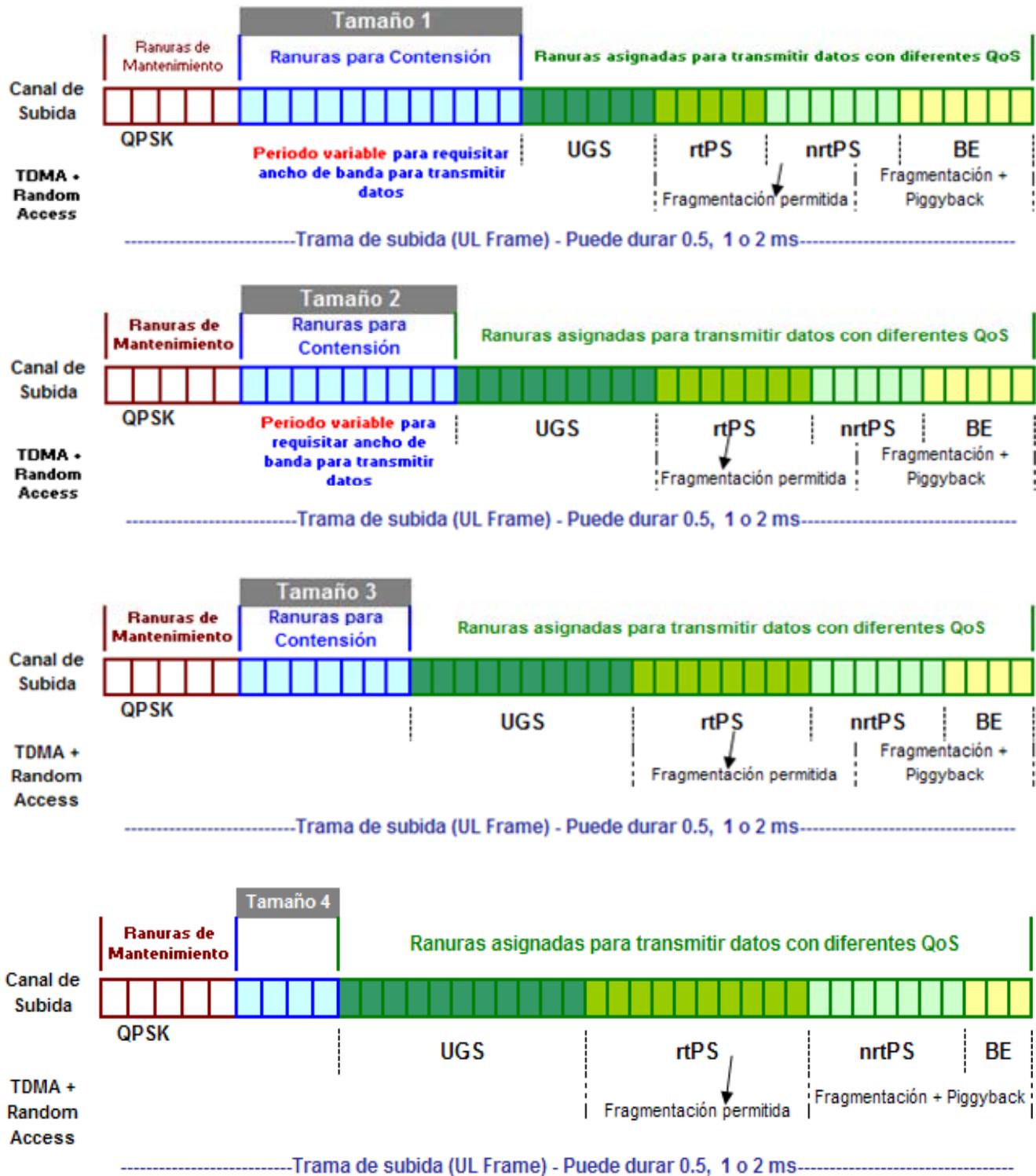


Figura 1.2 –Escenario propuesto con las ranuras de contención variando dinámicamente en función a los requerimientos de los usuarios de la red.



1.3 METODOLOGÍA

Procederemos a ejecutar el método científico con el fin de obtener conocimiento y plasmarlo en este trabajo de Tesis.

Elegir y definir el problema, en nuestro caso es proponer estrategias de mejoramiento en la resolución de colisiones durante el periodo de contención para solicitar ancho de banda para aquellas clases de servicio que no sean periódicas.

Estructurar un marco teórico. Inicialmente nuestro desarrollo se basará en el paradigma de resolución de colisiones en las ranuras de contención utilizadas para solicitar ancho de banda (para clases de servicio no periódicas), analizado a detalle en el rango de referencias marcadas entre [6 y 11]. Por lo tanto tomaremos en cuenta publicaciones que analicen el comportamiento de dichas ranuras para redes inalámbricas, como la descrita en la referencia [7].

Nuestro punto de partida serán algunos estudios previos realizados sobre manejo de tráfico IP sobre una red GSM utilizando las capacidades de GPRS, analizados a detalle en las referencias [1] y [2]. Está demostrado, con la implementación de esta técnica en redes celulares, que se mejora el uso de los recursos de la red, proporcionando a los usuarios una forma de transmitir datos en una red inalámbrica diseñada para transmitir originalmente voz.

Se procederá a realizar un análisis exhaustivo sobre esta tecnología con el fin de obtener los conocimientos y hacer las adaptaciones necesarias para aplicar ciertos conceptos en una red IEEE 802.16. Y así, reusar algunos de los mejores criterios para implementar un sistema eficiente para administrar las colisiones en las ranuras de contención.

Es muy importante considerar los avances relacionados con algoritmos aplicados en la capa MAC para redes IEEE 802.16 con mejoras a los algoritmos ya existentes, como el mostrado en la referencia [4], con la idea de no duplicar o imitar desarrollos anteriores.

Establecer hipótesis acorde a la solución esperada. Nuestra hipótesis inicial se basa en que una vez que se desarrolló y probó un modelo teórico que proponga un manejo eficiente de las ranuras de contención y después de analizar el comportamiento dinámico de una red basada en el estándar IEEE 802.16 podamos obtener mejores tiempos de respuesta a los que actualmente existen, un ejemplo de ello se muestra en la referencia [8].



Proponer una forma eficiente de administrar las colisiones en el área de contención, lo cual se traducirá en una utilización eficiente del espectro radioeléctrico (dado que se disminuirá y en el mejor de los casos se eliminarán las colisiones), no importando del tipo de clase de servicio que tengan los usuarios ni el número de usuarios simultáneos en la red.

Probando nuestras hipótesis. Para la parte del análisis de los algoritmos implementados en el modelo de GPRS se utilizará el modelo de GSM (GPRS) desarrollado para MATLAB Versión 2007b, el cual fue ya previamente demostrado como un caso de éxito en la referencia [12].

Una vez teniendo los resultados usando dicho modelo (iterativo utilizando probabilidades binomiales) aplicado a una red basada en IEEE 802.16, se probará un nuevo modelo basado en ecuaciones en diferencias con el fin de simplificar el desarrollo matemático. Así, reducimos la necesidad de procesamiento para obtener los resultados teóricos similares al modelo binomial original pero con menos iteraciones.

En el caso de probar las adaptaciones y mejoras de los modelos desarrollados para GPRS en una red IEEE 802.16 se utilizarán los ejemplos de simulación en OPNET Modeler® versión 14, OPNET Wireless Suite y el Modelo WiMAX más actualizado para este software (versión Julio 2007) implementadas en el trabajo descrito en la referencia [5].

1.4 CONTRIBUCIÓN Y RELEVANCIA

Dada la complejidad del tema, al día de hoy hay muy pocas publicaciones y desarrollos relacionados con el manejo de las colisiones en las ranuras de contención en redes inalámbricas basadas en el estándar IEEE 802.16. Por esta razón, la importancia que podría tomar este análisis y su posterior desarrollo y continuará proyectando a la Universidad Nacional Autónoma de México internacionalmente.

Otro punto muy importante es que la Facultad de Ingeniería siga siendo un miembro activo del *WiMAX Development Consortium de OPNET Technologies, Inc.*

Los resultados de los trabajos conducidos en la Facultad de Ingeniería en torno al tema del manejo de colisiones que presentan los usuarios al acceder al medio inalámbrico de una red IEEE 802.16 deberán converger en la publicación de un *IEEE Journal*⁵.

⁵ *IEEE Journal*. Publicación técnico-académica publicada por la IEEE que se enfocan a un cierto tema en particular. Son escritos en lenguaje académico y pueden ser arbitrados, lo cual significa que un comité de expertos en la disciplina ha leído el artículo y han determinado que tiene valor y es digno de ser publicado. Los *Journals* arbitrados son considerados de muy buena reputación.



Esta tesis es la primera extensión del artículo *“Performance Analysis of the Best Effort Service Class in Broadband IEEE 802.16 Networks”*, el cual fue enviado y aceptado para su publicación y presentación en la 22nd IEEE Canadian Conference on Electrical and Computer Engineering (IEEE CCECE 2009). La conferencia se llevará a cabo del 3 al 6 de mayo del 2009 en la ciudad de St. John’s, Newfoundland, Canadá. Para obtener detalles de este importante evento de la IEEE se puede acceder a www.ieee.ca/ccece09/comm.php.

1.5 ESTRUCTURA DE LA TESIS

Dividiremos nuestro desarrollo básicamente en 5 Capítulos, en donde describiremos detalladamente los aspectos fundamentales sobre el estándar IEEE 802.16, así como antecedentes técnicos de las redes inalámbricas, su desarrollo y explotación.

En el capítulo 2, haremos una descripción general de los objetivos y alcances de la tesis, de igual forma describiremos el problema de la solución de colisiones que abordaremos a detalle en los capítulos posteriores.

En el capítulo 3, introduciremos los conceptos básicos de los principios de las comunicaciones usando redes inalámbricas. Adicionalmente, daremos detalles sobre las capas física (PHY) y de acceso al medio (MAC), de este tipo de redes. Mencionaremos algunos aspectos importantes del estándar IEEE 802.16 como, el manejo de la calidad de servicio.

Asímismo, definiremos técnicamente la arquitectura de una red basada en el estándar IEEE 802.16. Se describirá el modelo de referencia usado, hablaremos de tramas, ranuras de contención y de las interfaces aéreas del estándar.

Usaremos el Capítulo 4 para explicar por qué es importante el manejo de las colisiones en una red con acceso al medio compartido y aprovecharemos para describir detalladamente el modelo matemático que utilizaremos para hacer el análisis del comportamiento dinámico de la red.

Finalmente, en el capítulo 5 se mostrarán gran parte de los resultados obtenidos (apoyados con gráficas) durante la ejecución de los escenarios propuestos, tomando en cuenta la variación de los parámetros principales.



En las conclusiones se emitirán juicios basados en las hipótesis iniciales, las mismas reflejarán los alcances y las limitaciones del estudio, así como las recomendaciones que puedan ser de utilidad al problema de investigación.

La bibliografía contendrá la relación de documentos y textos utilizados como apoyo. Y en los apéndices incluiremos los desarrollos informáticos utilizados.

Nuestro glosario de términos tendrá como objeto el homogeneizar y racionar la terminología específica empleada al elaborar este trabajo y que no corresponde al lenguaje común.

CAPÍTULO 2

REDES INALÁMBRICAS DE BANDA ANCHA

En este capítulo daremos una introducción general sobre la evolución que han tenido las redes inalámbricas, y como han cambiado para convertir en realidad lo que actualmente podemos explotar en el vasto ámbito de las telecomunicaciones. Daremos una descripción detallada del más nuevo y eficiente sistema inalámbrico de banda ancha disponible actualmente, *WiMAX*. Y se presentará un análisis comparativo entre *WiMAX* y un sistema de comunicación inalámbrica basada en células, 3G.

2.1 REDES INALÁMBRICAS Y REDES INALÁMBRICA DE BANDA ANCHA

Existen un gran número de tecnologías de transmisión inalámbrica. Dichas tecnologías pueden ser clasificadas en distintas familias de redes. La Figura 2.1, muestra las distintas categorías de redes inalámbricas, incluyendo las tecnologías más conocidas para cada tipo de red.

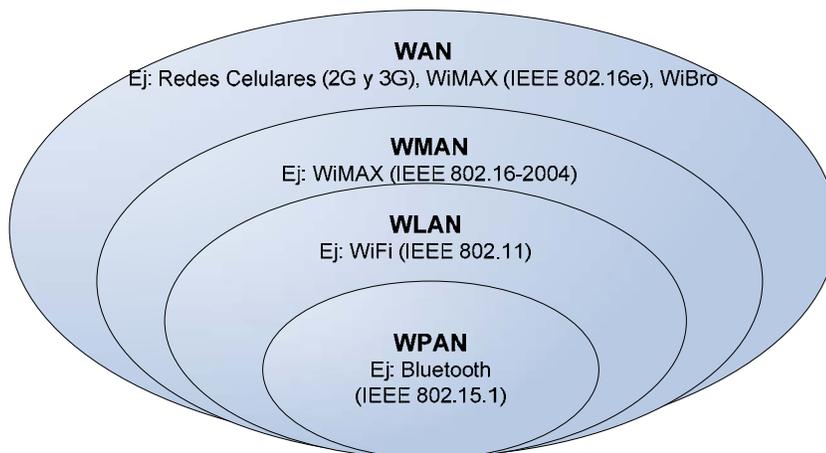


Figura 2.1 – Tipos de redes inalámbricas, en cada caso se menciona la tecnología más conocida.

Una PAN (*Personal Area Network*), es una red de datos (generalmente inalámbrica) usada para la comunicación entre dispositivos que puede manipular una persona por sí misma, por ejemplo un PDA (*Personal Digital Assistant*). Entonces, su alcance es de sólo unos cuantos metros, generalmente se asume que es menor a 10 m.



Una LAN (*Local Area Network*), es una red de datos que se utiliza para la comunicación entre dispositivos como: computadoras, teléfonos e impresoras. Este tipo de redes cubren un área relativamente pequeña, como una casa, una oficina o un segmento pequeño de un campus. Los sistemas LAN más frecuentemente usados son *Ethernet* (para redes alámbricas fijas) y WiFi (*Wireless Fidelity Alliance Consortium*), para redes inalámbricas.

Una MAN (*Metropolitan Area Network*), es una red de datos que puede abarcar varios kilómetros, comúnmente un campus entero o una ciudad inclusive. Algunos ejemplos de tecnologías MAN son: FDDI (*Fiber Distributed Data Interface*), DQDB (*Distributed Queue Dual Bus*), la versión fija de *WiMAX* (IEEE 802.16-2004) puede ser considerada una tecnología MAN.

Una WAN (*Wide Area Network*) es una red de datos que cubre un área geográfica muy grande, tan grande como el planeta Tierra. Una WAN está basada en la conexión de redes LAN, permitiendo a los usuarios de una ubicación específica comunicarse con otros usuarios en otra ubicación remota geográficamente.

Regularmente, una WAN está constituida por una serie interconexiones de nodos conmutados. Éstas conexiones se llevan a cabo usando líneas telefónicas dedicadas (*leased lines*), a través de circuitos conmutados o por métodos de conmutación de paquetes. El ejemplo más claro de una red WAN es, el Internet. Otros ejemplos de este tipo de redes son: una red de telefonía celular 3G o una red *WiMAX* móvil (IEEE 802.16e).

2.2 WIMAX o 3G...?

Es un tanto riesgoso dar comentarios concluyentes es un momento en donde los sistemas de acceso inalámbrico de banda ancha están teniendo grandes avances e innovaciones. Sin embargo, basados en antecedentes técnicos descritos en las referencias [14], [15] y [16], reportes incluidos en [17] y distintos análisis detallados en [1], [2], [3] y [4] podemos dar algunas comparaciones:

Con el fin de hacer una comparación equitativa con las redes 3G, solamente se considera *WiMAX* en su versión móvil, dado que la versión fija representa un mercado completamente distinto al representado por 3G.

En la Tabla 2.1 se muestra la comparación técnica entre los principales y más grandes sistemas de comunicación inalámbrica: 2G de GSM, así como su evolución EDGE, 3G de GSM (UMTS), WiFi (en sus dos variantes, 802.11a y 802.11b) y *WiMAX*. Destacaremos las ventajas de cada uno de los siguientes dos sistemas, 3G y *WiMAX*.

Tecnología	Banda de Frecuencia	Licencia	Ancho de Banda por canal (Frecuencia Portadora)	Número de usuarios por canal	Bit Rate	Alcance
GSM/EDGE	0.9 GHz 1.8 GHz	SI	200 kHz	De 2 a 8	473.6 Kbps	Hasta 30 Km
UMTS	1.9 GHz	SI	5 MHz	Muchos (orden de magnitud=25), se pierde ancho de banda	21 Mbps	Hasta 5 Km
WiFi (802.11g-2003)	2.4 GHz	No	20 MHz	Sólo uno (en un instante dado)	54 Mbps	Hasta 100 m
WiFi (802.11b)	2.4 GHz	No	5 MHz	Sólo uno (en un instante dado)	11 Mbps	Hasta 100 m
WiFi (802.11a)	5 GHz	No	20 MHz	Sólo uno (en un instante dado)	54 Mbps	Hasta 100 m
WiMAX (IEEE 802.16-2004)	2.3 GHz 2.5 GHz 3.5 GHz 5.8 GHz	Se definen bandas con y sin licencia	3.5 MHz 7 MHz 10 MHz 20 MHz	Más de 100	75 Mbps	20 Km (con CPE exterior)

Tabla 2.1 – Algunos elementos de comparación entre los más famosos sistemas de comunicación inalámbrica.

En la Figura 2.2 se presenta una comparación entre la tasa de transmisión y la movilidad de los principales sistemas inalámbricos.

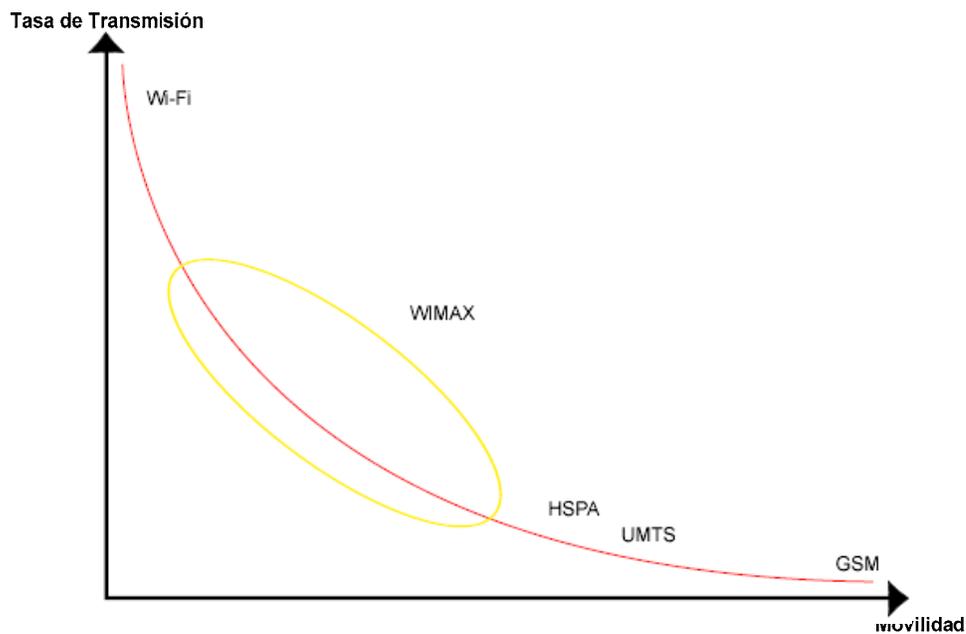


Figura 2.2 –Comparación de los principales sistemas inalámbricos (Tasa de Transmisión contra Movilidad..)

2.2.1 VENTAJAS DE UN SISTEMA CELULAR 3G

- *WiMAX* utiliza frecuencias por arriba de los 2.3 GHz a diferencia de una red 3G la cual opera normalmente en el rango de los 1.8 GHz. La potencia de recepción disminuye cuando la frecuencia se incrementa y la potencia de transmisión es normalmente limitada por cuestiones ambientales y requerimientos regulatorios. El rango de alcance de una red *WiMAX* es en general más pequeño que el rango de un sistema 3G. Sin embargo, el tamaño de la célula no es una limitante común en zonas de alta densidad de usuarios, en donde se localiza la mayor parte del mercado de los operadores móviles.
- 3G es una realidad comercial, desde 2005 su equipamiento incluye redes de tipo HSDPA (*High-Speed Downlink Packet Access*) en varios países. Globalmente, 3G tiene una ventaja de 3 años en relación a *WiMAX*.
- El espectro utilizado por *WiMAX* cambia de un país a otro.
- Algunos países tienen restricciones sobre el uso de *WiMAX*.
- Una red 3G tiene un gran soporte de empresas manufactureras de equipo de comunicaciones líderes a nivel mundial, como Nokia. Estas compañías ahora ponen interés en *WiMAX* aunque permanecen muy interesadas en 3G.

2.2.2 VENTAJAS DE UN SISTEMA WIMAX

- El espectro en frecuencia de *WiMAX* suele ser más barato que en el caso de los sistemas 3G. La licencia de UMTS en Europa, y muy específicamente en Alemania y el Reino Unido, alcanzaron un sorpresivo alto costo.
- *WiMAX* es un sistema sumamente abierto, muchos de los algoritmos son dejados al criterio del fabricante, lo cual abre distintas puertas a la optimización. Aunque podría significar problemas de inter-operación en los primeros años de desarrollo.
- La capa física de *WiMAX* está basada en OFDM (*Orthogonal Frequency Division Multiplexing*), una técnica de transmisión conocida por tener una alta eficiencia espectral, en relación a SC CDMA (*Single Carrier-Code Division Multiple Access*). Aunque existen planes para implementar OFDM y MIMO (*Multiple-Input Multiple-Output*) en las actuales redes 3G. Esta evolución es llamada, LTE (*Long-Term Evolution*). Esto da una ventaja en tiempo a *WiMAX* dado al tiempo que demorará la implementación de OFDM en las actuales redes 3G.



- WiMAX es una tecnología completamente IP. Este no es el caso de las redes 3G en donde existen muchos protocolos intermedios. Una comunicación IP de punta a punta en las redes 3G se tendrá en su evolución (LTE).
- WiMAX tiene el soporte de empresas gigantes en la industria de las telecomunicaciones como, Intel, Nokia, KT, Samsung, Nortel y Motorola entre muchas otras.

2.3 EL ESTÁNDAR IEEE 802.16

Los esfuerzos para estandarizar las redes de datos comenzaron en 1979 en el IEEE. En febrero de 1980 (80/2), se fundó el grupo (comité) de trabajo IEEE 802, el cual se dedicó a la definición de los estándares para las redes LAN y MAN.

Los protocolos especificados en los objetivos del IEEE 802 incluían las dos capas más bajas (*Data Link* y *Física*) de la pila de protocolos del modelo de referencia OSI (*Open Systems Interconnection Reference Model*), de 7 capas. IEEE 802 divide la capa de *Data Link* en dos subcapas llamadas LLC (*Logical Link Control*) y MAC (*Media Access Control*).

Desde aquellos años a la fecha, muchos subcomités IEEE 802 han sido creados, incluido en ellos el IEEE 802.16. Dicho grupo de trabajo se ha dedicado al desarrollo de las redes inalámbricas de banda ancha (BWA- *Broadband Wireless Access*), y su propósito es el de proveer un estándar técnico para los sistemas con altas tasas de transmisión inalámbricas (WMAN).

Como ya se ha mencionado, IEEE 802.16 es la tecnología aplicada a redes inalámbricas de banda ancha usada por *WiMAX*. Las principales características de este tipo de redes se describen como sigue:

Las frecuencias portadoras son menores a 11 GHz. Por ahora, las bandas de frecuencia consideradas son 2.5 GHz, 3.5 GHz y 5.7 GHz. La versión móvil de *WiMAX* está construida basada en OFDM, lo cual permite una alta eficiencia espectral.



Las tasas de transmisión son en promedio de 10 Mbps, aunque algunos reportes como el mostrado en la referencia [18], presentan ejemplos con tasas de hasta 70 u 100 Mbps. Estos valores se obtienen en forma práctica con un canal de transmisión ideal con células pequeñas.

El grupo de trabajo IEEE 802.16 fue creado en 1999, y fue dividido en dos grupos:

- 802.16a, teniendo la frecuencia central dentro del intervalo de los 2 y los 11 GHz.
- 802.16, con un valor de intervalo de frecuencia de los 10 a los 66 GHz.

En el estándar IEEE 802.16-2004 se consolidan los estándares IEEE 802.16-2001, 802.16c-2002 y 802.16a-2003. Antes de publicarse oficialmente el 802.16-2004, existió una versión llamada 802.16d iniciada en Septiembre de 2003 con el objetivo de tomar en cuenta el estándar **ETSI HiperMAN BWA**¹.

El proyecto 802.16d fue después concluido con la aprobación del documento 802.16-2004 y con la posterior retirada de todos los documentos 802.16 previos, incluyendo las enmiendas a, b y c.

Muchos documentos han sido aprobados y publicados por los grupos de trabajo IEEE 802.16, los cuales se presentan en la Tabla 2.2.

Nombre del Documento	Fecha de Publicación	Descripción
802.16	Diciembre 2001	Frecuencias: 10-66 GHz Linea de vista (LOS, Line-Of-Sight) Alcance:2-5 Km Valores de ancho de banda: 20, 25 y 28 MHz
802.16a	Enero 2003	Frecuencias: 2-11 GHz Sin línea de vista (NLOS)
802.16-2004	Octubre 2004	Revisión y consolidación oficial de los documentos anteriores Alcance: 5-50 Km
802.16e	Diciembre 2005	Movilidad OFDMA (SOFDMA)

Tabla 2.2 – Principales documentos IEEE 802.16 publicados.

¹ ETSI HiperMAN BWA – Estándar europeo equivalente al IEEE 802.16. ETSI: European Telecommunications Standards Institute.

2.3.1 TRANSICIÓN DE 802.16-2004 A 802.16e²

El estándar 802.16-2004 fue definitivamente muy útil, dado que reemplazó e integró una serie de publicaciones anteriores que trataban la misma tecnología. Aunque, después de su publicación fueron requeridos cambios y adiciones técnicas, principalmente para incluir en el estándar el concepto de movilidad. Cabe aclarar que en ésta tesis se presenta un análisis sobre el comportamiento dinámico de una red inalámbrica de banda ancha basada en el estándar 802.16-2004. Aquí presentamos algunas de las principales diferencias del estándar 802.16e en función del 802.16-2004:

- Aparece el concepto de Estación Móvil (*MS-Mobile Station*).
- Cambios en los procedimientos de *handover* (cambio de célula) en la capa MAC. Se introduce el concepto de cambio de EB o célula sin interrumpir la comunicación actual (*soft-handover*).
- Se incluyen características de ahorro de energía, para soportar movilidad.
- Se introduce el SOFDMA (*Scalable Orthogonal Frequency Division Multiple Access*).
- Se adiciona una subcapa de privacidad con el fin de incrementar el nivel de seguridad en la comunicación inalámbrica.
- Se utilizan técnicas para mejorar la eficiencia de la transmisión y recepción de la señal inalámbrica, MIMO y AAS (*Adaptive Antenna System*) son usados.
- Características de *multicast*³ y *broadcast*⁴ son implementadas.
- Se adiciona una nueva Clase de Servicio: *ertPS* (*extended real-time Polling Service*).

2.4 EL FORO WIMAX

El estándar IEEE 802.16 sólo provee la tecnología. Entonces, es necesario contar con organismos para la certificación de conformidad y verificación de interoperatividad. En el caso de IEEE 802.11 WLAN. El foro WiFi tuvo el mayor crédito del éxito de la tecnología WiFi (como es conocida actualmente). El problema de la certificación ha sido el más importante para *WiMAX*, dado que muchos fabricantes reclamaban que habían cumplido con el estándar 802.16 en sus productos (equipos llamados *pre-WiMAX*).

² El detalle técnico de algunas de las características mencionadas en esta sección se describen en el Capítulo 3.

³ *Multicast* – Tipo de difusión (en este caso radiodifusión) en una red, en donde la información transmitida llega a múltiples nodos, no a todos.

⁴ *Broadcast* – Tipo de difusión (en este caso radiodifusión) en una red, en donde la información transmitida llega a todos los nodos.

El foro *WiMAX* (www.wimaxforum.org) fue creado en Junio del 2001 con el objetivo tener el mismo rol que tuvo el foro WiFi con el estándar IEEE 802.11. El foro *WiMAX* provee entonces la certificación de conformidad, compatibilidad e interoperatividad de los productos fabricados basados en el estándar IEEE 802.16. En la Tabla 2.3 se muestran algunos de los miembros principales del foro *WiMAX*:

Fabricantes de equipo		Proveedores de servicios WiMAX	
Airspan	Motorola	British Telecom	WorldMax
Alcatel	Navini	France Telecom	NZWireless
Alvarion	Nokia	Korean Telecom	Wipet
Broadcom	Nortel	Sprint	Volare PR
Cisco	NEC	Nextel	Summa Telecom
Ericsson	Proxim	Telmex	QMAX
Fujitsu	Sagem	Axtel	Telkom
Intel	Samsung	Ultranet2Go	Iberbanda
LG	Siemens	Orbitel	Telia Sonera
Lucent	ZTE	Maxtel	Chunghwa Telecom
		Yozan Inc	AT&T

Tabla 2.3 – Principales miembros del foro WiMAX.

2.5 CERTIFICACIÓN DE PRODUCTOS WiMAX

El primer reconocimiento del foro *WiMAX* fue otorgado al Centro de Tecnología de Comunicaciones (Cetecom Lab - www.cetecom.es) en Málaga, España; como el primer laboratorio de certificación de productos *WiMAX*. En Febrero de 2006, el foro designó al laboratorio TTA (*Telecommunications Technology Association's IT Testing and Certification*) en Seúl, Korea; como el segundo para certificar compatibilidad e interoperatividad de productos *WiMAX*.

El cumplimiento del estándar IEEE 802.16 no debe ser confundido con la interoperatividad. La combinación de estas dos pruebas generan un examen de certificación real.

El cumplimiento del estándar 802.16, consiste de un proceso en donde los fabricantes de EBs y ESs prueban sus productos para asegurar que cumplen con la eficiencia de acuerdo a las especificaciones de los llamados PICS (*Protocol Implementation Conformance Specification*) de *WiMAX*.

Ahora bien, la certificación *WiMAX* de interoperatividad es un procedimiento mediante el cual se ponen a prueba los equipos en el laboratorio en donde interactúan con dispositivos de otros fabricantes (usando la misma tecnología). Esto es, se aplica una prueba de eficiencia de una EB de un fabricante que transmite y recibe información desde/hacia una o varias ESs de un fabricante distinto.



2.5.1 PRODUCTOS WIMAX CERTIFICADOS

El proceso de certificación de equipos comenzó en el verano de 2005 en Cetecom. La primera certificación *WiMAX* para un equipo tuvo lugar el 24 de Enero de 2006. La lista completa de equipos puede ser encontrada en www.wimaxforum.org/kshowcase/view. Todos los equipos aquí ubicados fueron certificados con un perfil de uso basado en el estándar IEEE 802.16-2004.

La certificación actual de equipos sólo se ofrece a tres diferentes tipos de fabricantes:

1. Compañías con experiencia en la fabricación de equipos *pre-WiMAX*
2. Compañía previamente especializada en productos para redes celulares.
3. Nuevas empresas que inicializan sus actividades comerciales basadas en la fabricación de equipos *WiMAX* exclusivamente.



CAPÍTULO 3

CONCEPTOS BÁSICOS Y DEFINICIONES

En este capítulo definiremos algunos de los principales conceptos básicos que conforman una red inalámbrica, dado que en la comprensión de estos conceptos se basará la futura interpretación de los resultados que se obtengan.

La comunicación inalámbrica es un tipo de comunicación que técnicamente no utiliza un método de propagación físico, es decir sólo se utiliza la modulación de ondas electromagnéticas, las cuales se propagan sin un medio físico (cable) que comunique cada uno de los extremos. De aquí que, los únicos dispositivos físicos presentes son los emisores y receptores de la señal, por ejemplo: antenas, teléfonos móviles, módems inalámbricos, etc.

Adicionalmente escribiremos algunas características fundamentales de la arquitectura de una red inalámbrica de banda ancha basada en el estándar IEEE 802.16. Se dará una descripción del modelo de referencia, así como detalles sobre sus componentes.

Con la definición de estos rasgos podremos ver claramente la importancia de mejorar la eficiencia en la administración del acceso al medio inalámbrico compartido.

3.1 BANDAS DE FRECUENCIA

Uno de los parámetros básicos que caracterizan a un sistema inalámbrico es la banda de frecuencia usada, generalmente expresada en términos de megahertz (MHz) o gigahertz (GHz).

3.1.1 CANALES

Una banda de frecuencia es regularmente dividida en múltiples canales de frecuencia con el fin de soportar actividades de comunicación independientes. Un canal es usualmente definido con una frecuencia central específica y el canal de ancho de banda para ocuparlo. Por ejemplo, el canal 2437 MHz con 22 MHz de ancho de banda ocupa el espectro radioeléctrico desde la frecuencia 2426 MHz hasta la 2448 MHz.

3.1.2 TIPOS DE ESPECTRO RADIOELÉCTRICO

Típicamente, el espectro de frecuencia en el cual opera un sistema de comunicación inalámbrica puede ser clasificado en: espectro con licencia y espectro exento de licencia.

En el caso de las bandas con licencia, son asignados permisos individuales emitidos por una entidad oficial regulatoria (en México, estas licencias son asignadas por la Secretaría de Comunicaciones y Transportes) con el fin de garantizar la utilización legal y exclusiva.

Por ejemplo, la mayoría de los operadores de telefonía móvil mantienen derechos exclusivos sobre las bandas de frecuencia sobre las que ellos operan, normalmente estos derechos exclusivos fueron obtenidos por medio de una subasta pública.

Contrariamente, la operación del espectro exento de licencia no requiere de un permiso explícito, aunque generalmente las transmisiones están permitidas solo utilizando dispositivos de comunicación que hayan sido certificados para cumplir las especificaciones regulatorias. Pero en general, la operación de estas bandas de frecuencia exentas no ofrece protección legal contra interferencia perjudicial.

Comúnmente, los dispositivos de comunicación inalámbrica que puedan operar tanto en ambas bandas (con y sin licencia), requieren una certificación expedida por la entidad regulatoria gubernamental. Dicha certificación usualmente requiere que los dispositivos cumplan ciertos requerimientos específicos sobre algunas características físicas como: potencia de transmisión, selección de canal y control de emisiones *in-band* y *out-band*. Aunque podrían también especificar el tipo de modulación y el detalle de los protocolos a utilizar.

3.2 TIPOS DE REDES INALÁMBRICAS

Los dispositivos inalámbricos pueden ser clasificados basándonos en si los equipos están físicamente estáticos en una sola localidad o tienen permitido moverse mientras se lleva a cabo la sesión de comunicación.

3.2.1 REDES FIJAS Y MÓVILES

Por ejemplo, la ITU (*International Telecommunication Union*) define como servicio inalámbrico fijo a: “un servicio de radio comunicación entre dos puntos fijos específicos”. En las redes inalámbricas fijas, los dispositivos inalámbricos deberán permanecer estáticos en ambos lados mientras dure la comunicación.



En las redes inalámbricas móviles, al menos uno de los dispositivos puede moverse durante la operación. Un buen ejemplo de una red inalámbrica móvil, es una red de telefonía celular, en donde la EB es un dispositivo fijo y los teléfonos (ESs) son usuarios subscriptores.

Los requerimientos regulatorios de las redes inalámbricas fijas y móviles son diferentes por distintas razones. Por ejemplo, dado que las redes fijas son estacionarias y no son transportadas por seres humanos durante su operación, entonces podría permitirse al dispositivo usar una alta potencia de transmisión. De igual forma, los dispositivos móviles, especialmente los teléfonos celulares son transportados por humanos durante su operación, por esto el dispositivo podría tener un conjunto diferente de regulaciones con respecto a la potencia de transmisión.

A parte de los requerimientos regulatorios, los dispositivos para redes fijas y móviles difieren en su diseño técnico. Las redes fijas pueden usar transmisores de alta potencia e incluye antenas direccionales de alta ganancia con línea de vista. Esta situación generada por un canal de propagación simple.

Por otro lado, las redes inalámbricas móviles, pueden involucrar dispositivos como un teléfono celular con una pequeña antena omnidireccional, de ahí que el canal pudiese ser de calidad deficiente. Además, como el dispositivo puede moverse rápidamente, el sistema debe acoplarse pronto a las variaciones de canal. En el nivel físico, el sistema debe lidiar con cambios en la señales de radio frecuencia (RF) que son resultado de dicho movimiento del dispositivo.

Una efectiva comunicación móvil requiere una optimización de las capas física y de acceso al medio. Otro importante requerimiento adicional para estas dos capas es, un rápido *handover*. Un ejemplo de un algoritmo de *handover* es presentado en la referencia [19].

3.3 LA CAPA FÍSICA (PHY)

El estándar IEEE 802.16 describe un sistema inalámbrico de banda ancha, de ahí que los datos son transmitidos a alta velocidad en la interfaz aérea a través de ondas electromagnéticas con una frecuencia dada (frecuencia de operación).

La capa física establece la conexión física entre ambos lados, comúnmente en dos direcciones UL y DL.

Como el estándar IEEE 802.16 es evidentemente una tecnología digital, la capa física es la responsable de la transmisión de las secuencias de bits. Por esto, se define el tipo de señal utilizada, el tipo de modulación y demodulación, la potencia de transmisión y otras características físicas.

El estándar IEEE 802.16 considera la banda de frecuencias 2-66 GHz. Esta banda está dividida en dos partes:

- Primer rango está entre los 2 y los 11 GHz y está destinada para transmisiones sin línea de vista (NLOS).
- Segundo rango está entre los 11 y los 66 GHz y se utiliza para transmisiones con línea de vista (LOS).

Seis interfaces físicas son descritas en el estándar IEEE 802.16, las cuales son resumidas en la Tabla 3.1.

Designación	Banda de Frecuencia	Multiplexado	Opciones para capa MAC ¹
WirelessMAN-SC Single Carrier LOS	10-66 GHz (LOS)	TDD y FDD	
WirelessMAN-SCa Single Carrier NLOS	Frecuencias menores a 11 GHz (NLOS) Banda con Licencia	TDD y FDD	AAS (6.3.7.6) ARQ (6.3.4) STC (8.2.1.4.3) Movilidad
WirelessMAN-OFDM	Frecuencias menores a 11 GHz Banda con Licencia	TDD y FDD	AAS (6.3.7.6) ARQ (6.3.4) STC (8.3.8) Mesh (6.3.6.6) Movilidad
WirelessMAN-OFDMA	Frecuencias menores a 11 GHz Banda con Licencia	TDD y FDD	AAS (6.3.7.6) ARQ (6.3.4) HARQ (6.3.17) STC (8.4.8) Movilidad
WirelessHUMAN	Frecuencias menores a 11 GHz Banda Exenta de Licencia	TDD	AAS (6.3.7.6) ARQ (6.3.4) STC (8.4.8) Mesh (6.3.6.6) Movilidad

Tabla 3.1 – Las 5 interfaces físicas definidas en el estándar IEEE 802.16. Tomado de la referencia [13].

3.3.1 FORWARD ERROR CORRECTION (FEC)

Una forma de manejar los errores en las transmisiones inalámbricas es la de detectar y corregir automáticamente los símbolos o bits erróneos en lugar de depender de la detección y retransmisión de los paquetes con errores, por ejemplo como cuando se utilizan algoritmos como *checksum* o CRC (*Cyclic Redundancy Check*).

La técnica de corrección usada más frecuentemente es llamada “*Forward Error Correction*”, la cual es capaz de detectar y corregir sobre la señal recibida mediante la adición de redundancia a las señales transmitidas. FEC es el tipo de codificación utilizada por el estándar IEEE 802.16 y es la codificación utilizada en nuestro análisis como se muestra en el Apéndice C.

¹ Las Opciones para la capa MAC son descritas a detalle en el Capítulo 8 de la referencia [16].

Los protocolos de comunicación inalámbrica deben considerar los gastos y costos asociados con los temas de control de errores en la capa física (como FEC) y con los asociados a los de la capa de acceso al medio, como ARQ (*Automatic Repeat Request*). Por ejemplo, mientras que FEC es capaz de corregir ciertos tipos de error rápidamente con una probabilidad de éxito muy alta, la adición de FEC a cada bloque de datos transmitido reduce la eficiencia del canal y podría incrementar la demora en la entrega de las PDUs (*Protocol Data Unit*).

Por otro lado, en la capa de acceso al medio el protocolo ARQ incrementará las demoras en un canal con una alta tasa de errores, en donde la detección, comunicación y la retransmisión consumirán una importante porción del ancho de banda disponible.

3.3.2 UNIDADES DE DATOS

Las unidades de datos son las unidades básicas intercambiadas entre las distintas capas de la pila de protocolos. Las SDUs (*Service Data Unit*) y las PDUs son las dos unidades fundamentales usadas con este fin. Las SDUs son unidades de datos intercambiadas entre dos capas adyacentes de la pila de protocolos de un mismo dispositivo. El SAP (*Service Access Point*) define la interfase entre dos capas de protocolos adyacentes, en donde los servicios de la capa inferior están disponibles para la capa superior. Las SDUs son intercambiadas entre dos capas de protocolos adyacentes a través del SAP. PDU es la unidad de datos intercambiada entre entidades pareadas en la misma capa de protocolos. Cuando un PDU es transferido a una capa inferior para transmitirse, este se convierte en un SDU de la capa inferior.

La Figura 3.1 muestra el concepto de PDU, SDU y SAP.

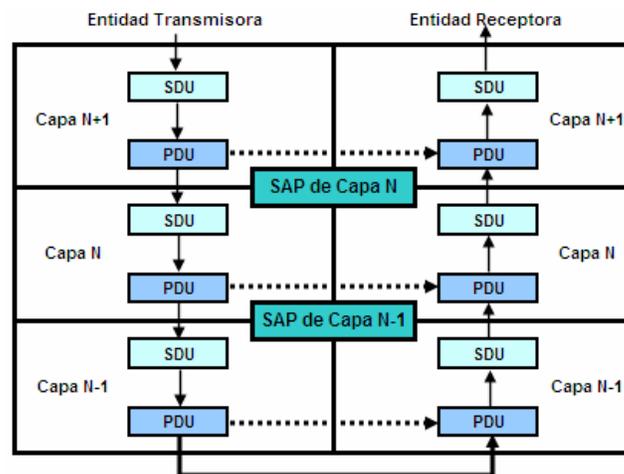


Figura 3.1 – Ejemplo de cómo funcionan las PDUs, SDUs y el SAP en la pila de protocolos.

Una SDU recibida de una capa superior puede sufrir una gran cantidad de transformaciones para llegar a ser una PDU, y así transmitirse a la capa de pareo (en donde la conexión lógica entre dos entidades está ejecutándose). Cuando las capas física o de acceso al medio reciben una SDU, se le pueden adicionar encabezados e información de control adicionales para hacer posible la transmisión de la PDU.

En adición, una SDU puede ser fragmentada, entonces más de una PDU puede ser formada de una sola SDU. De igual manera, fragmentos de varias SDUs pueden ser empaquetadas en una sola PDU. El único requerimiento indispensable para la fragmentación de SDUs es que el proceso inverso (defragmentación) se efectúe en la misma capa de pareo en el receptor antes de entregar el PDU a una capa superior.

3.4 CALIDAD DE SERVICIO (QoS)

En un sistema inalámbrico de banda ancha soportando múltiples servicios, debe ser tomado en cuenta apropiadamente el hecho de que diferentes servicios requieren distintos niveles de eficiencia con el fin de mantener una apropiada calidad de servicio.

Una red que soporta tipos de servicios diferenciados puede proveer ciertas clases de servicio que requieren niveles de eficiencia muy específicos, utilizando un conjunto de medidas cualitativas y cuantitativas.

Aplicaciones como voz, video y transferencia de datos demandan distintos requerimientos de QoS. Dichos requerimientos pueden ser especificados o forzados con un conjunto de parámetros.

En el modelo de QoS por priorización, el tráfico es tratado en la red basado en su prioridad relativa en algún instante. Sin embargo en el modelo de QoS por parametrización, la red intenta garantizar el conjunto de parámetros de QoS para cierto tipo de tráfico, de acuerdo a lo previamente configurado.

3.4.1 MÉTRICAS DE QoS

Las métricas de QoS comúnmente utilizadas en el modelo de parametrización son:

Ancho de Banda (*Throughput*) – El ancho de banda es típicamente expresado en bits por segundo, aunque puede ser expresado también en *bytes* por segundo o en paquetes por segundo, dependiendo de la aplicación. La aplicación actual del concepto ancho de banda como un parámetro en la QoS depende de varios factores como el tamaño del paquete, retransmisiones, etc. El ancho de banda alcanzable, está ligado a la máxima capacidad instantánea de la red, también expresada en bits por segundo.

Retardo (*Delay*) – El retardo, es el tiempo que tarda la información en llegar de la fuente al destino y viceversa. El retardo es expresado en unidades de tiempo como milisegundos.

Jitter – El *jitter* es la variación en el retardo. Este parámetro es muy importante en aplicaciones de tiempo real como voz o video.

Pérdida de paquetes – La pérdida de paquetes en cualquier sistema es típicamente definida como un porcentaje. Las aplicaciones de datos pueden tolerar algo de demora y de *jitter*, pero no pueden tolerar pérdida de paquetes. Si un paquete se pierde, este debe ser retransmitido por alguna de las capas para que la aplicación funcione adecuadamente.

3.4.2 FUNCIONES DE QoS

Un acercamiento muy simplista hacia la QoS es la asignación estática de servicios y la reservación de recursos. Esto es, de formal similar a como opera la QoS en una red de circuitos conmutados, cuando una conexión se establece por medio de la reservación de recursos exclusivos durante la duración de la comunicación. Sin embargo, este no es un método eficaz en las redes de paquetes conmutados. De ahí que, para que una red de soporte completo de QoS a los usuarios, en la capa de red son necesarias las siguientes cuatro funciones:

Control de Admisión y Asignación– Respecto de cómo una red soluciona el problema de QoS, debemos ocuparnos de la máxima cantidad de tráfico y servicios que la red puede soportar en un instante en particular. De aquí, que es muy importante regular qué usuarios pueden acceder a los servicios de la red y de cómo lo harán. Este control debe ser dinámico, basado en la carga de la red y otros parámetros.

Clasificación de Tráfico – Con el fin de que la red pueda proveer un tratamiento diferenciado al tráfico basado en los requerimientos de QoS, se debe tener la posibilidad de clasificar los paquetes entrantes. Dicha clasificación puede llevarse a cabo mediante la colocación de información en el encabezado de cada paquete para determinar como deberá ser tratado.

Vigilancia (*policing*) y Formato (*shaping*) – A pesar de los controles de admisión, el tráfico requiere ser regulado dinámicamente para estar seguros de que cumple con las especificaciones del servicio acordado. *Policing* meramente descarta o reclasifica los paquetes que no concuerden con la clase de servicio indicada. *Shaping*, por otro lado, puede encolar aquellos paquetes que no concuerden con la clase de servicio indicada y demorar su entrega. Este tipo de algoritmos no son parte del estándar IEEE 802.16.

Programación de Tráfico (*scheduling*) – El *scheduler* de tráfico es responsable de determinar el orden de la transmisión de los paquetes basado en los requerimientos de QoS, cuando existen fructuaciones en el tráfico entrante que resultan en encolamiento del tráfico. Por esta razón, es un componente muy importante de cualquier arquitectura de QoS. Una red sin QoS trata a todas las aplicaciones con un manejo de cola de tipo FIFO (*First-In First-Out*). En cambio, un *scheduler* de QoS asegura que el tráfico sea tratado en función a su QoS y no solamente en el orden de arribo del paquete. Un ejemplo de uno de estos algoritmos externos aplicado al estándar IEEE 802.16 se describe en la referencia [10].

En la Figura 3.2 se muestra un ejemplo de cómo funciona un *scheduler*.

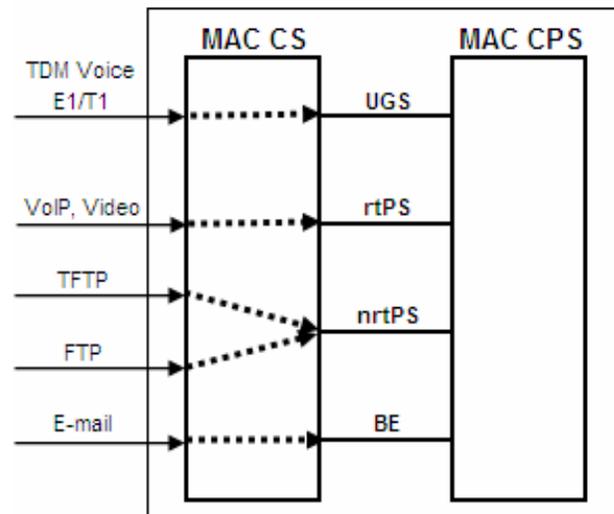


Figura 3.2 – Mecanismo de programación de QoS de una EB en donde se muestran la asignación de las aplicaciones a su respectiva clase de servicio. Imagen tomada de IEEE 802.16-2004.

3.4.3 TIPOS DE QoS EN EL ESTÁNDAR IEEE 802.16

En el estándar IEEE 802.16-2004 se definen cuatro clases de servicio, y una quinta fue adicionada en la enmienda IEEE 802.16e:

BE *Best Effort* – Usado para las aplicaciones de más baja prioridad como correos electrónicos, navegación WEB, etc.

ertPS *Extended real-time Polling Service* – Esta es una nueva clase definida en la enmienda IEEE 802.16e. Esta clase es intermedia entre rtPS y UGS.

nrtPS *Non-real-time Polling Services* – Esta clase es utilizada para servicios que no operan en tiempo real pero que tienen alguna limitación de tiempo.

rtPS *Real-time Polling Services* – Usada para datos que se transmiten con una tasa variable y que además son de tiempo real. Por ejemplo, video con compresión MPEG (Moving Pictures Experts Group).

UGS *Unsolicited Grant Services*. Esta clase es dedicada al tráfico de tasa constante (CBR). UGS garantiza tamaños fijos de ancho de banda en ciertos intervalos de tiempo. Ejemplo de esto sería un enlace de tipo T1/E1 privado.

3.5 ARQUITECTURA

El estándar IEEE 802.16 define las características técnicas de las capas física (PHY) y de acceso al medio (MAC). Consecuentemente, con el fin de asegurar la interoperabilidad entre redes construidas con equipos fabricados por distintos proveedores, para operaciones como el *roaming*², es muy importante definir estándares sobre la amplitud de los rangos de las interfaces y los equipos. Por ejemplo 3GPP (*3rd Generation Partnership Project*) utilizado para GSM.

La arquitectura de una red basada en el estándar IEEE 802.16 está basada en la utilización de protocolos IP estandarizados y su compatibilidad con ciertos esquemas de servicio. Los aspectos clave de la arquitectura de una red IEEE 802.16 son resumidos como sigue:

Flexibilidad con un control de acceso al medio (MAC) común – El estándar IEEE 802.16 define una capa MAC común que puede funcionar con distintas tecnologías de la capa física (PHY). La capa MAC es lo suficientemente flexible para soportar muchas de las actuales tecnologías de capa física. Y está preparada para adecuarse a otras extensiones de MAC conforme sea necesario.

Arquitectura Modular – Ambas capas PHY y MAC del estándar IEEE 802.16 son enormemente modulares, es decir, tienen una serie de características obligatorias y opcionales que pueden ser utilizadas para configurar una amplia variedad de configuraciones fijas o móviles. Muchas de las características opcionales son negociables entre la ES y la EB.

Soporte de múltiples topologías de red – Se refiere al soporte de topologías de red de tipo salto simple o multi-salto, es decir conexiones Punto a Punto, Punto a MultiPunto o de malla pueden ser construidas en las capas PHY y MAC de una red IEEE 802.16. Esto habilita una implementación eficiente de cualquiera de las topologías inalámbricas soportadas.

Calidad de servicio integrada – La calidad de servicio (QoS) integrada a la capa MAC del estándar IEEE 802.16 define múltiples tipos de QoS los cuales pueden ser controlados y configurados desde capas superiores. El modelo de QoS del estándar IEEE 802.16 define comportamientos para varios tipos de QoS por lo tanto se requiere de un algoritmo de QoS *Scheduling* en la capa MAC, un excelente ejemplo de este tipo de algoritmos (que están fuera del estándar) se muestra en la referencia [3].

² *Roaming* – Concepto utilizado en comunicaciones inalámbricas que está relacionado con la capacidad de un dispositivo para moverse de una zona de cobertura a otra. El término más adecuado para expresar este concepto en nuestro idioma es: itinerancia.

3.6 MODELO DE REFERENCIA

Existen dos grandes componentes en los sistemas inalámbricos de banda ancha: el plano de datos/control y el plano de administración. El plano de datos y control define los procedimientos de cómo los datos son encapsulados y desencapsulados en la capa de acceso al medio y como son modulados y demodulados en la capa física. Se requiere de un conjunto de funciones de control para soportar otras varias de configuración y coordinación.

Mientras otros estándares se refieren a esto como “administración” (por ejemplo Administración de las capas PHY y MAC en el estándar IEEE 802.11), en el caso del estándar IEEE 802.16 se refiere a este asunto como plano de control con el fin de diferenciarlo de los sistemas de administración externos. Cada sistema inalámbrico de banda ancha requiere funciones de administración, para controlar eficientemente la seguridad, la calidad de servicio, las configuraciones iniciales y muchas otras funciones vitales. En la Figura 3.3 se describe el modelo de referencia para el estándar IEEE 802.16.

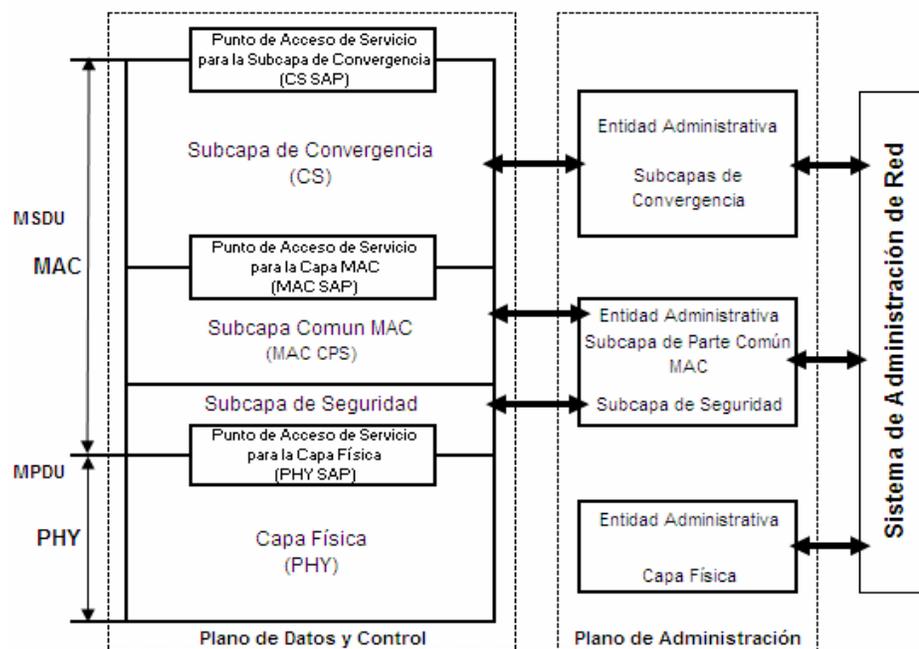


Figura 3.3 – Modelo de referencia para el estándar IEEE 802.16.

Como ya hemos mencionado, actualmente el estándar IEEE 802.16 soporta un conjunto de subcapas de convergencia como interfaz con protocolos como *IP*, *Ethernet* y *ATM*. La subcapa de convergencia para un servicio en particular provee todas las funciones que son específicas para ese protocolo de capa superior y así soportarlo. Por ejemplo, cuando la Capa MAC de un equipo envía una MPDU hacia un equipo correspondiente, este MPDU es recibido como una PSDU (PHY SDU) por la capa física.

3.6.1 SUBCAPA DE CONVERGENCIA (CS)

La subcapa de convergencia (CS) es la capa superior de la subcapa MAC CPS. La CS utiliza los servicios provistos por la citada capa MAC CPS a través del punto de acceso de servicio MAC (MAC SAP), entonces la subcapa CS ejecuta las siguientes funciones:

- Aceptar PDUs de las capas superiores. En la versión del estándar IEEE 802.16-2004, se proveen especificaciones para manejar dos tipos de capas superiores: CS para ATM y CS para paquetes. Para el caso de CS para paquetes, los protocolos de capa superior pueden ser IP v4 o IP v6.
- Clasificar y asignar las MSDUs hacia los identificadores de conexión (CID) apropiados. Esta es una función básica para los mecanismos de QoS de las redes IEEE 802.16.
- Procesar, en caso de ser necesario, las PDUs de capas superiores basándose en la clasificación de los mismos.
- Entregar las CS PDUs hacia los MAC SAP apropiados y recibir CS PDUs desde la entidad destino.

3.6.2 SUBCAPA COMÚN DE LA MAC (MAC CPS)

La subcapa común (CPS) se encuentra en medio de la capa MAC. La CPS representa el núcleo del protocolo de control de acceso al medio (MAC) y es responsable de:

- Asignación de ancho de banda.
- Establecer la conexión.
- Mantener la conexión establecida.

El estándar IEEE 802.16-2004, define un conjunto de mensajes de transferencia y de administración. Los mensajes de administración con intercambiados entre la ES y la EB antes y durante el establecimiento de la conexión. Cuando la conexión ha sido establecida, los mensajes de transferencia pueden entonces intercambiarse para permitir la transmisión de datos.

3.7 TRAMAS Y MULTIPLEXADO

Como se ha descrito en la sección 3.6, la CS convierte las PDUs de capas superiores en MPDUs, y la MAC CPS es responsable de transportar estas MPDUs. Los datos encapsulados de las MPDUs de las redes IEEE 802.16 son eventualmente transportados en tramas físicas (PHY).

El estándar IEEE 802.16 soporta múltiples tamaños de tramas PHY. La cantidad de información de capas superiores que puede ser transportada dentro de una trama PHY depende de varios factores, incluyendo la modulación y la codificación utilizadas. Distintos MPDUs dirigidos a múltiples conexiones y direcciones pueden ser transportados en la misma trama PHY. La duración de las tramas PHY es fija para una configuración de sistema específica y no cambia durante la operación de la red, como se muestra en la figura en la Figura 3.4.

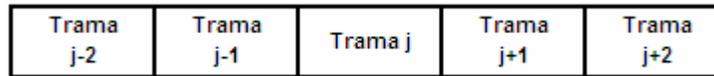


Figura 3.4 – Formato y tamaño de tramas PHY en el estándar IEEE 802.16.

El estándar IEEE 802.16 soporta ambas formas de multiplexar, TDD y FDD, el caso FDD es sólo permitido en las bandas de operación con licencia, esto se describe en la sección 4.4. Por defecto, los sistemas TDD o ATDD (*Adaptive Time Division Duplexing*) son fijos. Asimismo, el estándar soporta la operación en FDD en modo *Full Duplex* y H-FDD (*Half Duplex*) y ambos tipos de conexiones pueden coexistir en la misma red al mismo tiempo. En una red mixta, la EB es responsable de que las ESs operando en modo H-FDD no reciban asignaciones traslapadas en el UL o el DL. Es importante describir los siguientes conceptos antes de pasar a la explicación de la operación de TDD y FDD.

Perfiles de Ráfaga – Las ráfagas de datos del DL o UL se caracterizan por un conjunto de parámetros denominados perfiles de ráfaga, que incluye información como el tipo de modulación, FEC, tipo de preámbulo y tiempos de guarda.

Descriptor de canal DL (DCD-Dowlink Channel Descriptor) – Describe las características del canal DL PHY, el cual es enviado a todas las ESs desde la EB (*broadcast*) en forma periódica. Contiene información como, la duración de la trama, así como los perfiles de codificación y de ráfaga DL.

Descriptor de canal UL (UCD-Uplink Channel Descriptor) – Describe las características del canal UL PHY, el UCD también es transmitido desde la EB en *broadcast* en intervalos periódicos. El UCD contiene perfiles de ráfaga UL que definen los códigos de intervalo usados en el UL y el UIUC (*UP Interval Usage Codes*) y sus características físicas asociadas. Adicionalmente, el UCD contiene los parámetros de *backoff*³ que se utilizarán para los *slots* de contención en el UL. Todas las transmisiones en el UL y el DL del estándar IEEE 802.16 son programadas, es decir, ninguna ES puede transmitir información a través de estos canales a menos que así se le indique en ciertos periodos de tiempo. Aunque, hay algunos casos en donde una ES puede enviar información a la EB sin una asignación de recursos previa, por ejemplo, cuando una ES quiere registrarse y formar parte de la red.

³ *Backoff* – Término que se refiere al periodo de tiempo que una ES deberá esperar antes de intentar transmitir de nuevo una vez que colisionó uno de los paquetes que envió.

Las ESs deben seguir las políticas impuestas por los algoritmos de *backoff*, las cuales están indicadas en los parámetros que transmite la EB con el fin de minimizar la probabilidad de que una colisión se presente en los *slots* de contención. El esquema de tener ranuras (*slots*) de contención se define en el UL para soportar este tipo de comunicación.

Mapa del Canal de Bajada (DL MAP) – El DL MAP describe las asignación en el canal de bajada. Dichas asignaciones son especificadas en elementos físicos específicos dentro del DL MAP. El DL MAP incluye información sobre el tiempo de inicio de las ráfagas de DL y DIUC (*Down Interval Usage Codes*) para indicar el perfil de ráfaga a utilizar.

Mapa del Canal de Subida (UL MAP) – El UL MAP describe la asignación de espacios TDMA en el canal de subida, en donde se especifican los tiempos exactos junto con los perfiles de ráfaga asignados para transmitir con el fin de que la ESs puedan enviar sus datos en el UL.

3.7.1 RANURAS FÍSICAS (*SLOTS*)

Las ranuras o *slots* son utilizadas en el estándar IEEE 802.16 para expresar cantidades de tiempo, lo cual se traduce en ancho de banda. El *slot* es la unidad básica de asignación de recursos (ancho de banda) en la red. La definición de *slot* varía conforme a la interfaz física usada. Para el caso (nuestro caso) de *Single-Carrier*, un *slot* se define como la duración de cuatro símbolos QAM (*Quadrature Amplitude Modulation*).

3.7.2 TRAMAS TDD

La estructura de una trama TDD del estándar IEEE 802.16 se ilustra en la Figura 3.5. Como se puede apreciar la longitud fija de las tramas TDD se divide en las porciones DL y UL, llamadas subtrama DL y subtrama UL.

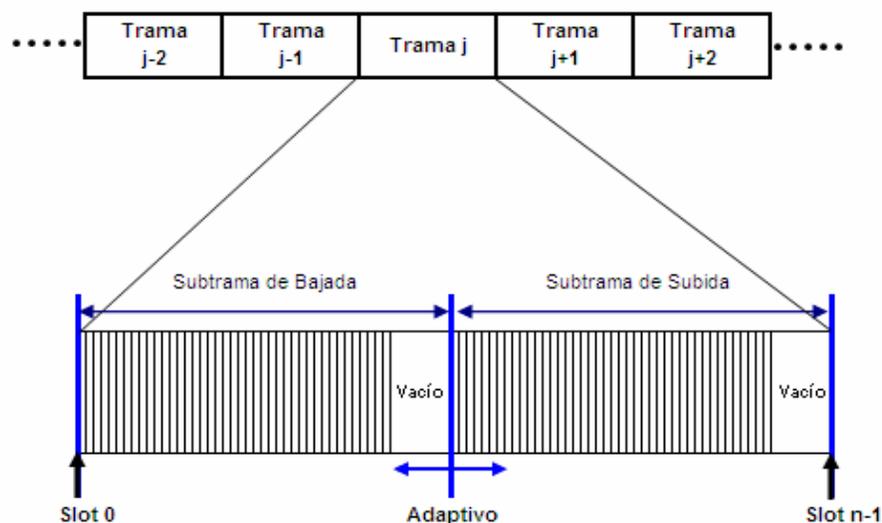


Figura 3.5 – Estructura de una trama TDD en una conexión punto-multipunto.

La trama TDD se repite en intervalos regulares, aunque la duración de una subtrama DL o UL puede ser ajustada dependiendo de los parámetros del sistema. Las subtramas DL o UL pueden transportar una o más MPDUs. Los DL MAPs y UL MAPs son transmitidos en *broadcast* al inicio de cada trama.

Los sistemas TDD también requieren un espacio de transición RTG (*Receive Transmit Transition Gap*) entre la subtrama UL y la subsecuente subtrama DL de la siguiente trama. La duración de dichos espacios varía dependiendo de la interfaz aérea usada.

Este espacio es utilizado por la EB como transición para cambiar del modo receptor al modo transmisor y por la ES como transición para cambiar de modo transmisor al modo receptor. De forma similar, un espacio de transición TTG (*Transmit Receive Transition Gap*) se requiere entre el DL y el UL, en la misma trama.

Típicamente la subtrama DL comienza con la información requerida para la sincronización física, como el preámbulo. En resumen, la subtrama DL contiene cualquier información a ser transmitida desde la EB a una o más ESs. La subtrama UL, es la porción TDMA, la cual es utilizada por una o más ESs para transmitir información hacia la EB.

3.8 TOPOLOGÍAS

El estándar IEEE 802.16 define dos posibles topologías de red:

- Topología punto-multipunto (PMP). Ver Figura 3.6.

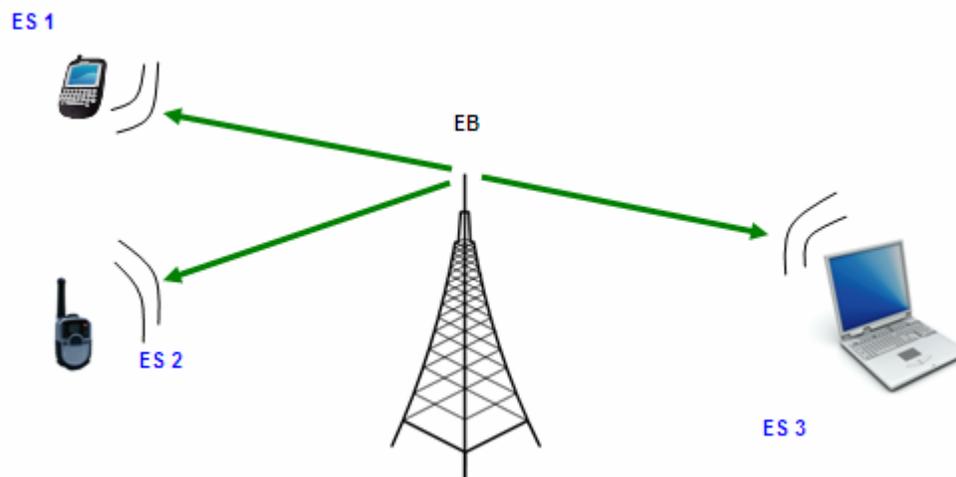


Figura 3.6 – Topología PMP, en donde la EB cubre sus ESs. Todas las comunicaciones inician o terminan en la EB.

- Topología de Malla (*mesh*). Ver Figura 3.7.

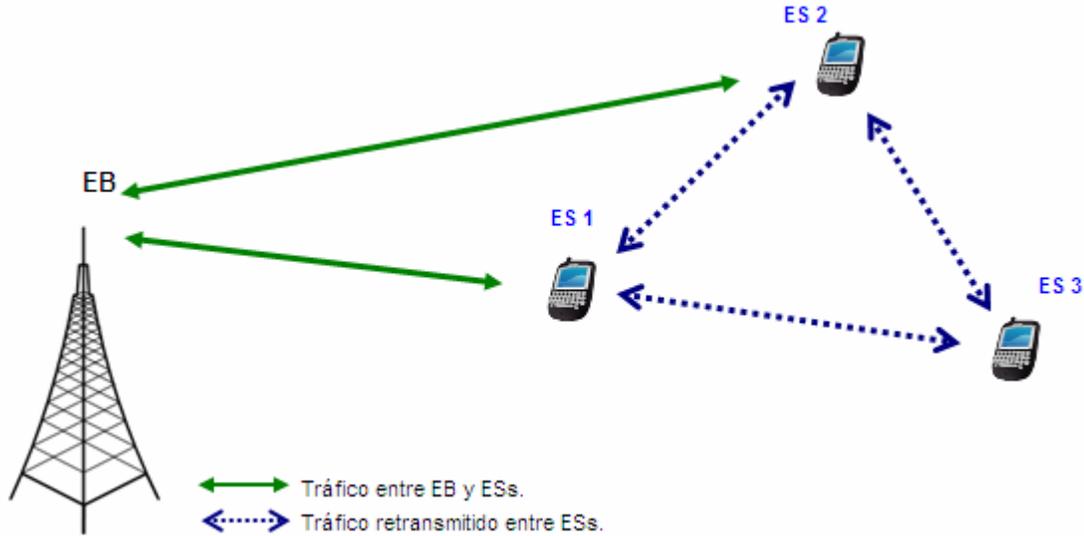


Figura 3.7 – Topología de Malla (*mesh*) donde la EB no es el único centro de distribución de tráfico.

La principal diferencia entre las dos topologías mencionadas es que, en el modo PMP, el intercambio de tráfico sólo puede tener lugar entre la EB y sus ESs. Mientras que en el modo de malla, el tráfico puede ser redireccionado a través de las distintas ESs que forman la red hasta que la EB pueda tener el control de todas las ESs.

El modo PMP representa una topología centralizada en donde la EB es el centro del sistema mientras que en el modo malla no. Los elementos de una red de tipo malla son llamados nodos.

En la topología de malla, cada nodo puede crear su propia comunicación con cualquier otro nodo en la red, dicha comunicación no está restringida a realizarse a través de la EB. De este modo, una gran ventaja del modo de malla es que el alcance de una ES puede ser mucho mayor, dependiendo del número de saltos hasta la ES más distante.

Los primeros desarrollos de las redes basadas en IEEE 802.16 fueron planeados bajo la topología PMP.

CAPÍTULO 4

MODELO MATEMÁTICO

En esta sección se dará una descripción detallada del nuevo modelo matemático, basado en ecuaciones en diferencias (publicado oficialmente en la referencia [3]) el cual fue utilizado en lugar del modelo previamente implementado para el modelado de una red con usuarios en GSM (usando probabilidades binomiales para GPRS, referencias [1] y [2]). Adicionalmente se incluye un análisis con el fin de mostrar las ventajas de nuestro modelo.

Asimismo, se introducirán algunos detalles técnicos sobre el manejo y la asignación de los recursos de la red, como el ancho de banda, la administración de los *slots* de contención, etc. Una de las principales aportaciones de este documento es el proponer un manejo eficiente del canal de subida en una red basada en el estándar 802.16-2004.

4.1 ANTECEDENTES

4.1.1 ANCHO DE BANDA – SOLICITUD / ASIGNACIÓN

El mecanismo de solicitud y asignación para el estándar IEEE 802.16 fue elegido para ser eficiente, con bajo retardo y flexibilidad además de encajar con los niveles de calidad de servicios (QoS).

Las solicitudes son hechas en base a una arquitectura “por conexión” para asegurar que puedan ser correctamente utilizadas en algoritmos imparciales usados por el *scheduler* del canal de subida (UL) de la EB, aunque las asignaciones son hechas a la ES no a la conexión.

Esto incrementa la eficiencia, el retardo también se reduce. La ES puede utilizar el ancho de banda asignado para su necesidad de transmisión más importante en ese momento y no siguiendo una instrucción de la EB.

Por ejemplo, si una ES requiere ancho de banda para una conexión con calidad de servicio de tipo BE, pero repentinamente surge una necesidad urgente por cambiar el tipo de calidad de servicio. Así, la ES puede usar el ancho de banda previamente asignado para enviar un mensaje DBPC-REQ (*Down Link Burst Profile Change-Request*) indicando esta petición de cambio.

Este importante mensaje es enviado con un retardo muy bajo en comparación a si la ES tuviera que esperar una nueva contención exitosa para ese propósito.

El retardo disminuye también porque se eliminan los ACK (positivos o negativos) a las requisiciones de ancho de banda, la ES no necesita esperar por el ACK, sólo espera por la asignación.

Aunque la asignación del ancho de banda sea proporcionada directamente a las ES y no a la conexión se podría presentar inconsistencia en el comportamiento de la red. Aún si las ESs no utilizan en forma diferente intencionalmente su asignación de recursos.

Una ES con varios servicios simultáneos no sabe cuánto ancho de banda de una asignación está siendo distribuido por la EB para una aplicación u otra. Adicionalmente, la falta de ACKs puede causar que la percepción de la EB del estatus de la cola de espera por ancho de banda de las ES sea distinta de la real y se presente una falta de sincronía.

La sincronía entre la EB y las ES puede perderse por varias razones:

- La EB no escucha las solicitudes de ancho de banda de la ES (error de capa física).
- La ES no escucha la asignación que la EB le dio en el mapa.
- La ES utiliza el ancho de banda asignado para un propósito distinto al que originalmente planteó.

Para rectificar cualquier discrepancia entre la realidad del estado de las ES y la percepción de la EB, la combinación de los procesos de solicitud y asignación de ancho de banda está diseñada para auto corregirse en caso de ser necesario. Después de un periodo, potencialmente basado en la calidad de servicio (QoS), si la ES sigue consumiendo y solicitando ancho de banda, simplemente continúa requiriéndolo. Para evitar que la percepción de la EB se aleje de la realidad por requerimientos duplicados, la ES emite un requerimiento agregado.

4.1.1.1 Requerimientos de Ancho de Banda en el UL-Incremental y Agregado

Un requerimiento agregado le indica a la EB el estado actual de la cola de espera de una ES para un servicio dado, permitiendo a la EB restaurar su idea sobre los requerimientos de servicio.

Los requerimientos agregados son necesarios para garantizar la naturaleza de auto corrección del procedimiento para requerir y asignar ancho de banda. Estos requerimientos agregados son necesarios y suficientes para que el mecanismo de requerimiento/asignación opere, aunque existe una potencial posibilidad de falla que puede ser fácilmente solventada con la adición del concepto de requerimientos incrementales.



Con los requerimientos incrementales, la ES simplemente está requiriendo más ancho de banda para un servicio en particular. Entonces la EB adicionará esta petición de ancho de banda a su actual idea de los requerimientos de ancho de banda en la red para ese servicio en lugar de usar esa solicitud de ancho de banda para restaurar su percepción de los requerimientos.

En general, el medio aéreo suele ser confiable. De ahí que, la mayoría de las solicitudes de ancho de banda suelen ser incrementales con requerimientos agregados periódicos para evitar que la EB no se aleje de la realidad en lo que a solicitudes de ancho de banda se refiere.

4.1.1.2 Requerimientos de Ancho de Banda en el UL–Autónomo y Piggyback¹

Existen algunas otras opciones para requerir ancho de banda. Un requerimiento autónomo es transmitido en una trama MAC dedicada (sin carga/payload), este tipo de requerimientos están incluidos en los dos principales métodos de petición y asignación de ancho de banda:

- *Polling² Unicast*
- *Polling* basado en contención

Para cualquier asignación de UL, la ES puede opcionalmente decidir utilizar dicha asignación para:

- Datos
- Requerimientos
- Requerimientos *Piggyback* en los datos

Para conexiones de tipo UGS, no son necesarios los requerimientos por parte de las ESs, dado que ambas partes saben de las necesidades recurrentes de ancho de banda. Sin embargo, la falta de sincronía puede ser un problema para este tipo de usuarios.

Aunque para manejar esta situación si la ES detecta que el tráfico UGS está comenzando a detenerse, puede proceder a encender el *bit* marcado como “SI” (*Slip Indicator*) en el GMSH (*Grant Management Subheader*) de un MAC PDU para esa conexión. Con esto la ES está solicitando un incremento ligero en la tasa de transmisión autorizada por la EB. Estas solicitudes continuarán incrementándose hasta que la ES apague el *bit* SI. Esto se muestra en la Figura 4.1.

¹ *Piggyback* – Mensaje para solicitar ancho de banda implícito en los datos que se están transmitiendo, en conexiones con QoS distintas a UGS, se utiliza el GMSH-*Grant Management Subheader* de un MAC-PDU para incluir el requerimiento sin afectar los datos.

² *Polling* - Término que se utiliza para hacer referencia a una operación de interrogación continua, con el fin de crear una actividad síncrona.

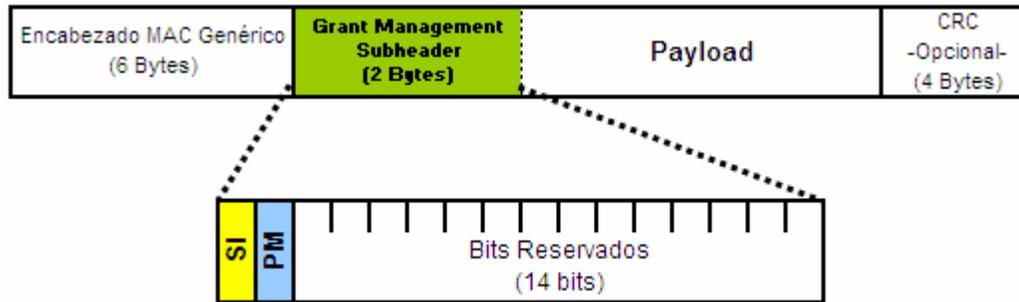


Figura 4.1- Estructura del Grant Management Subheader en un MAC-PDU. Para QoS=UGS.

Si una ES requiere más ancho de banda para un servicio que está actualmente transmitiéndose sobre el medio aéreo, la ES puede usar *piggyback* nuevamente en el GMSH el cual se convierte en *Piggyback Grant-Request Subheader*, cómo se muestra en la Figura 4.2.

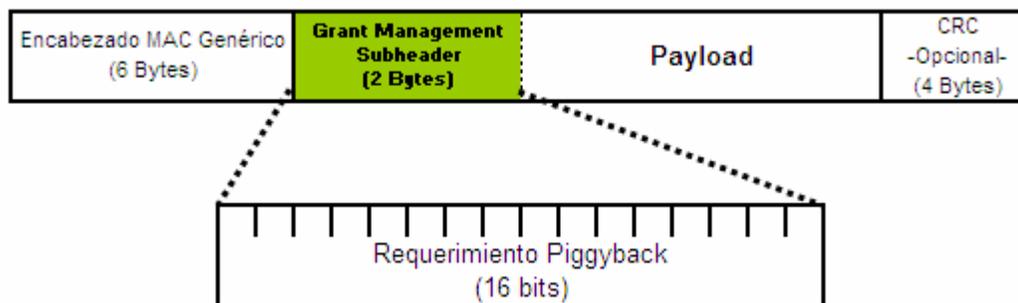


Figura 4.2- Estructura del Grant Management Subheader en un MAC-PDU. Para QoS ≠ UGS.

Asimismo, si una ES utilizando servicios UGS está recibiendo asignaciones de ancho de banda en exceso a sus necesidades, puede solicitar el ser puesta es un estado de espera dinámica (*polled*), es decir, su ancho de banda UGS puede ser utilizado por otra conexión, lo conseguirá encendiendo el *bit* PM del GMSH pero en el momento que exista un requerimiento UGS, su ancho de banda original le será asignado de inmediato.

4.1.2 MECANISMOS DE ACCESO AL UL (SOLICITUD /ASIGNACIÓN)

Como se adelanta informalmente en párrafos previos, el estándar IEEE 802.16 define dos métodos principales de solicitud y asignación de ancho de banda para el canal de subida (UL).

- *Polling Unicast*
- *Polling* Basado en Contención

Existe un tercer método llamado asignación de ancho de banda no solicitado, y se utiliza exclusivamente para conexiones UGS y no es tema de nuestro desarrollo.

4.1.2.1 Polling Unicast

El *polling* es el proceso mediante el cual la EB asigna ancho de banda a las ESs. Estas asignaciones pueden ser dadas a una ES en particular o a un grupo de ESs. El uso del *polling* simplifica la operación de acceso y garantiza que las aplicaciones puedan recibir servicio en forma determinística, si así es necesario.

Esta técnica de asignación es usada cuando la demanda de los recursos de ancho de banda no es lo suficientemente relevante para tener asignaciones de ancho de banda no solicitado para todos los usuarios. La EB puede entonces asignar directamente el monto requerido conforme el requerimiento completo de cada ES.

Cuando una ES es interrogada individualmente por la EB, se presenta un *polling unicast*, en este caso no se transmiten mensajes explícitos al conjunto de ESs. En lugar de eso, a la ES le es asignado, en el UL-MAP, suficiente ancho de banda. El estándar entonces indica que la ES, puede opcionalmente decidir cómo utilizar dicha asignación de recursos para transmitir datos, requerimientos o requerimientos de *piggyback* al momento de transmitir. Tomando en cuenta la posibilidad de que se tendrán distintos tipos de transmisiones pendientes a ser enviadas en el canal de UL, la ES utilizará *scheduling* para decidir que tipo petición adicional transmite (autónomo o *piggyback*) junto con los datos. La Figura 4.3 representa el mecanismo de *polling unicast*.

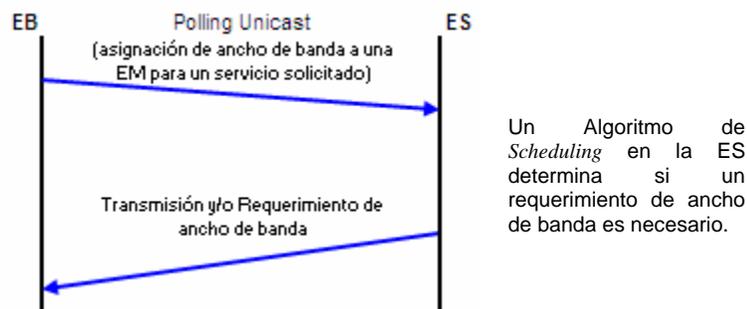


Figura 4.3- Ejemplo del mecanismo de *polling unicast*.

4.1.2.2 Polling Basado en Contención (Polling Multicast o Broadcast)

El ancho de banda disponible en el sistema pudiera ser no suficiente para que la EB interrogue a todas las ESs individualmente. Un mecanismo de petición y asignación de ancho de banda basado en la contención de los usuarios es permitido en una pequeña parte de cada trama (para FDD) o sub-trama (para TDD) de UL conocidos como *slots* de contención.

Con estos *slots* de contención, una ES puede acceder a la red solicitándolo a la EB a través de un UL *slot*. Si la estación base recibe la solicitud (lo cuál quiere decir que no se presentó una colisión), entonces la EB evalúa el requerimiento de la ES tomando el contexto del nivel de servicio contratado, el estado de la red, el algoritmo de *scheduling* y la posibilidad de asignarle slots en donde la ES pueda transmitir datos. Un ejemplo de un *polling multicast* es mostrado en la Figura 4.4.

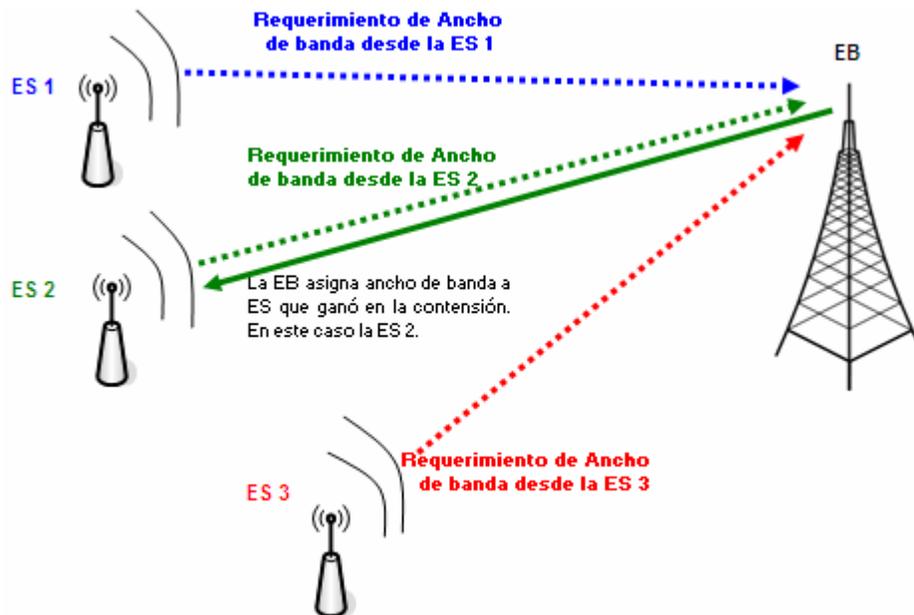


Figura 4.4- Ejemplo de un grupo de ESs con polling basado en contención.

4.1.3.3 Algoritmo para resolver contenciones

En un medio inalámbrico controlado, una colisión puede ocurrir durante el *ranging* y durante los intervalos de requerimiento de ancho de banda en el UL-MAP. El medio de transmisión (UL-MAP) y el algoritmo de resolución de contenciones es el mismo para los dos procesos.

Como una ES puede tener múltiples conexiones activas con la EB, y como consecuencia tener múltiples identificadores de conexión (CID), esto hace que los procesos de *ranging* y requerimiento se decidan por identificación de conexión o por calidad de servicio.

El método de solución de contenciones requerido por el estándar IEEE 802.16 está basado en el algoritmo binario de espera exponencial truncado (*Truncated Binary Exponential Backoff*).

Este algoritmo determina los tiempos que deberán esperar las ESs, después de una colisión, antes de volver a retransmitir sus peticiones, utilizando los valores para la ventana de espera inicial y de espera máxima dados por la EB. Estos dos parámetros son especificados en el mensaje descriptor de canal de subida (UCD).

Y está dado en valores de potencias de 2, menos 1. Por ejemplo, un valor de $k=4$ indica una ventana de espera de entre 0 y 15. Un valor de $k=10$ indica un valor de entre 0 y 1023. El algoritmo es válido hasta para un número de máximo de 16 colisiones por usuario antes de tener una contención exitosa.

El algoritmo utiliza la siguiente expresión completa para el cálculo de los tiempos de espera (*backoff*), los cuales indican el número de *slots* de contención que la ES deberá esperar antes de transmitir nuevamente:

si ($c < 16$) entonces
 $k = \min(c + i, t)$;
 b = número aleatorio entre $[0, 2^k - 1]$;
 ES espera “b” slots de contención antes del siguiente intento de transmisión;
si no
 Error de colisiones excesivas en la capa MAC.

Donde

i = Valor inicial del algoritmo (determinado por la EB).
 t = Valor truncado del algoritmo (determinado por la EB).
 c = Número de colisiones que ha experimentado el usuario.
 b = Tiempo de espera (*slots*) en caso de colisión.

4.1.4 SCHEDULING EN EL MODELO

Los algoritmos de *scheduling* fueron intencionalmente dejados fuera del alcance del estándar IEEE 802.16, ejemplo de uno de estos algoritmos externos se describe en la referencia [10]. El conjunto de servicios ofrecidos combinado con los algoritmos de *scheduling* es el área de mayor diferenciación entre fabricantes de equipos. Sin embargo, algunos conceptos son comunes para todas las implementaciones.

Un sistema IEEE 802.16 tiene tres principales puntos en donde se utiliza el concepto de *scheduling*. El más obvio, y el que se relaciona con una parte medular de nuestro trabajo de tesis es el *scheduling* que es generado por EB sobre el UL.

En la Figura 4.5 se describe gráficamente la estructura de los canales UL y DL. La EB recibe peticiones de las ESs, procesa dichas requisiciones, entonces crea y distribuye mensajes del UL-MAP. El modo en el que esté operando la capa física del UL impacta en la cantidad de ancho de banda disponible.

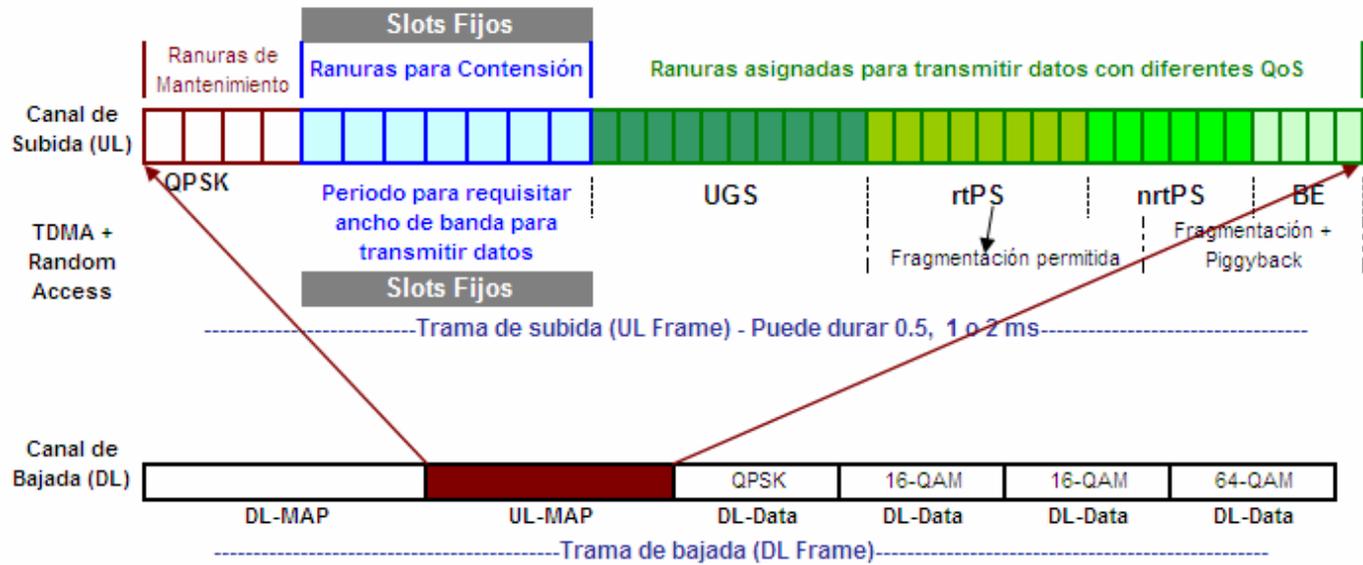


Figura 4.5-Estructura de los canales de subida y bajada (UL y DL).

Como hemos mencionado previamente en la sección 4.1.1, los mensajes de UL-MAP asignan ancho de banda a ESs no a las conexiones. Esto nos muestra el segundo punto en donde se utilizan *scheduling*. Una vez que una ES recibe una asignación, debe decidir que datos transmitir durante esas tramas. Para hacer esto, la ES ejecuta algún algoritmo de *scheduling*, análogo a los que se ejecutan en la EB, pero acotado al alcance de sus propios servicios.

El tercer punto de empleo de *scheduling* es empleado por la EB durante la creación del DL. Como en el caso de la generación de *scheduling* para el UL, esta tarea debe tomar en consideración no únicamente los datos en espera a ser transmitidos en el DL, si no que también debe considerar el modo en el que está operando la capa física de cada ES. Por esta razón, los mensajes de administración de DL MAC deben ser incluidos en el *scheduling*.

4.2 MODELO ANALÍTICO

4.2.1 DESCRIBIENDO LA ESTRATEGIA

Partiendo de que nuestro trabajo se orientará básicamente al tratamiento de las ESs con calidad servicio de tipo BE, en dónde básicamente los usuarios finales compiten por el ancho de banda disponible. Entonces empezaremos por describir la secuencia de actos que se llevan a cabo cuando una ES requiere utilizar esta clase de servicio para transmitir datos.

Cuando una ES está activa (ES_x), se forma un ciclo continuo con la serie de acciones mostradas en la Figura 4.6.

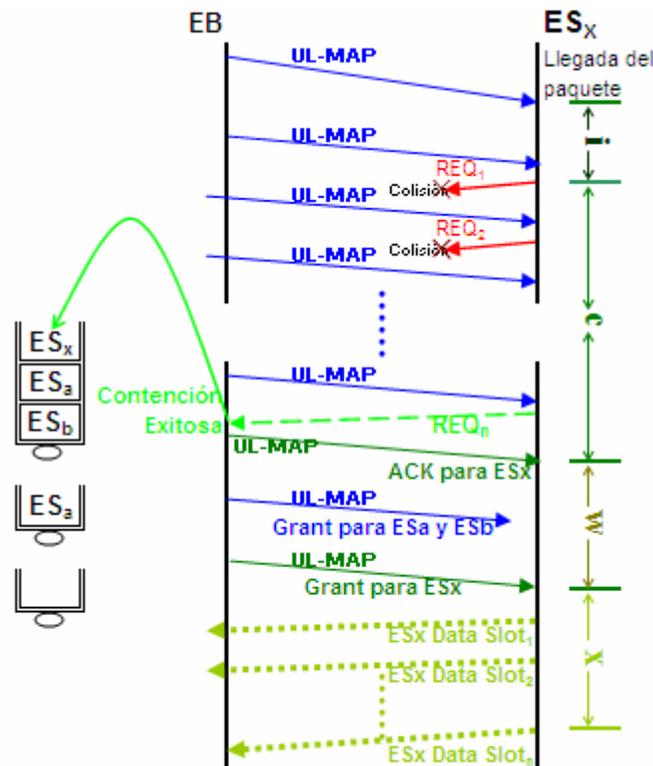


Figura 4.6–Componentes del retardo de acceso al medio para ES_x .

Una vez que un paquete llega desde una capa superior, la ES_x tiene que esperar por el siguiente UL-MAP proveniente de la EB que contenga un periodo de contención.

Entonces, la ES_x elige arbitrariamente uno de los *slots* de contención disponibles y transmite un requerimiento de ancho de banda (REQ) en donde indica el tamaño del paquete a transmitir. En el caso de que alguna otra ES elija el mismo *slot* de contención que la ES_x entonces ocurre una colisión y ninguna de las dos ES recibirá una asignación de ancho de banda para transmisión inmediata ni un mensaje de enterado (ACK) en el siguiente UL-MAP. De esto modo, la ES_x retransmite su requerimiento de ancho de banda (REQ) hasta que es exitosamente recibido por la EB, nótese que en nuestro ejemplo esto ocurre en el intento REQ_n .

Una vez recibida la solicitud de ancho de banda (proveniente de ES_x), la EB convierte el tamaño de paquete a su equivalente número de *slots*, los cuales podrían ser reservados en los subsecuentes UL-Frames.

En caso de que el espacio (en *slots*) en el siguiente UL-Frame no sea suficiente para alojar el total del ancho de banda solicitado por la ES_x (a través del REQ), la EB envía una asignación nula a la ES_x para informarle que está enterada de su requerimiento de ancho de banda, pero que en ese preciso momento no tiene recursos suficientes para satisfacer el REQ enteramente.

El modelado de tal secuencia de eventos se puede llevar a cabo dividiendo una sola transmisión de un paquete en los componentes que forman su retardo. La Figura 4.6, muestra gráficamente la relación de los retardos que utilizaremos, mismos que son descritos detalladamente en la Tabla 4.1.

ID	Definición del retardo (en minislots)
i	Retardo desde el tiempo en el que el paquete llega procedente una capa superior hasta el inicio del bloque de contención en donde ES_x transmitirá una petición de ancho de banda (REQ).
C	Representa el retardo total que el paquete pasa en contención, el cual inicia con el principio del bloque de contención en donde ES_x transmite la primera petición de ancho de banda (REQ) y termina cuando dicha ES_x recibe un mensaje ACK desde la EB (puede ser un "null grant").
W	Tiempo que tarda el algoritmo que asigna la programación de transmisión para conceder la petición de ancho de banda (REQ) a ES_x . Se mide desde que la ES_x está enterada de que tiene una asignación de ancho de banda por parte de la EB. El algoritmo que asigna la programación de las transmisiones (en la EB) atiende las peticiones de ancho de banda usando el método de manejo de colas FIFO. Si una petición de ancho de banda no puede ser asignada en el siguiente UL-Frame, entonces se espera a que las solicitudes de ancho de banda previas de otras ES sean atendidas. En caso de que un REQ pueda ser atendido inmediatamente después de ser recibida, la EB en lugar de responder con un mensaje de "null grant" (ACK), entonces la EB responde con una asignación indicando el número de slots que fueron reservados para que la ES transmita. En este caso, la recepción de tal asignación también señala el fin del tiempo de contención "c" y entonces "w" es igual a cero.
X	Tiempo que pasa mientras un paquete es transmitido desde la ES_x .

Tabla 4.1- Componentes que forma el retardo en la transmisión de un solo paquete.

Asimismo, en la Tabla 4.2 se detalla la notación completa de los componentes del modelo matemático usado.

Símbolo	Definición
C	Tamaño de un bloque de contención en un UL-Frame
R	Tamaño de un periodo de reservación en un UL-Frame
M	Tamaño total de un UL-Frame [$M = C + R$]
i	Retardo inicial
c	Retardo por contención hasta transmitir un REQ exitosamente
w	Retardo del algoritmo para asignar ancho de banda a la ES
x	Retardo por la transmisión de un paquete
t	Tiempo total para completar el ciclo de transmisión; $t = i + c + w + x$
$\bar{t}, \bar{i}, \bar{c}, \bar{w}, \bar{x}$	Esperanza matemática de t, i, c, w, x respectivamente
P_{sc}	Probabilidad de tener una contención exitosa en un UL-Frame
n	Número de ESs conteniendo en el actual UL-Frame.
s	Número de ESs están transmitiendo un REQ exitosamente en el actual UL-Frame
\bar{n}, \bar{s}	Esperanzas matemáticas de n y s respectivamente
\bar{b}	Número esperado de periodos de contención por ciclo t
N	Número máximo de ESs activas
P_c	Probabilidad de que una ES contienda en un periodo de contención
P_s	Probabilidad de que una contención sea exitosa en un minislots
d	Minislots disponibles para que una ES transmita por ciclo
\bar{m}	Tamaño medio de un paquete en la capa LLC
MAC_OV	Encabezado de capa MAC para el estándar IEEE 802.16
PHY_OV	Encabezado de capa física (PHY) para el estándar IEEE 802.16
γ	Rendimiento del sistema

Tabla 4.2- Notación completa de elementos componentes de nuestro modelo matemático.

4.2.2 MODELO ANALÍTICO

La idea principal del modelo analítico parte de la premisa de tener una buena administración de los *slots* de contención, para manejar de una forma eficiente las colisiones que puedan presentarse dentro de los *slots* asignados para que las distintas ESs soliciten ancho de banda a la EB.

Basados en la Figura 4.6 y tablas 4.1 y 4.2, el tiempo (t) que toma transmitir un solo paquete puede ser directamente obtenido sumando todos los componentes de retardo, así:

$$t = i + c + w + x$$

Tomando la esperanza matemática de ambos lados de esta ecuación, tenemos que:

$$\bar{t} = \bar{i} + \bar{c} + \bar{w} + \bar{x} \quad (1)$$

De estos parámetros, la eficiencia de la red (γ) normalizada para canal de UL puede ser determinado como una fracción del tiempo (*slots*) que actualmente tarda una sola ES en transmitir datos, y escalándolo al total de ESs (\mathcal{N}), obtenemos:

$$\gamma = \frac{N \cdot \bar{x}}{\bar{t}} \quad (2)$$

Uno de los objetivos de nuestro trabajo es el proveer un método eficaz para calcular esta forma de medir la eficiencia de una red basada en el estándar IEEE 802.16.

El cálculo de \bar{t} implica determinar otros componentes del retardo que se describirán posteriormente. El retardo medio generado por la transmisión de un paquete, dado en *slots*, puede ser directamente obtenido de la distribución de probabilidad del tamaño del paquete y tomando en consideración que el tamaño de los *slots* puede ser 2, 4, 8, 16, 32 o 64 *bytes*. Por lo tanto:

$$\bar{x} = \left[\frac{\bar{m} + MAC_OV + PHY_OV}{Tamaño_Slot} \right] \quad (3)$$

En el caso del encabezado de la capa de acceso al medio para el estándar IEEE 802.16 el valor por defecto es $MAC_OV = 6$ *bytes*. Sin embargo, el encabezado de la capa física depende de la técnica de codificación involucrada, ver referencia [15], capítulo 8 (páginas 307-630) para obtener detalles adicionales sobre tasas de codificación y técnicas de modulación aplicables.

Con fines prácticos, en nuestro análisis utilizamos tamaños fijos de paquetes para los usuarios simultáneos de la red (paquetes de 100, 150, 200, 250 y 300 *bytes*), considerando adicionalmente distintos tamaños del UL-Frame, 225 y 450 *slots*, es decir tramas con duración de 1 y 2 *ms* respectivamente.

Asímismo, a través de un programa realizado en Visual C++, anexo y detallado en el Apéndice C de este documento, utilizamos codificación FEC-Reed Solomon para obtener el valor promedio del tamaño de paquete mencionado en el párrafo anterior, incluyendo los encabezados de la capas de acceso al medio y física, todo esto basado en las referencias [15].y [20].

Continuando con el análisis de la ecuación (1), en la estimación del retardo inicial promedio \bar{i} se deben considerar los siguientes dos factores.

Primero, un paquete que llega de una capa superior puede hacerlo en cualquier momento dentro del actual UL-Frame.

Segundo, todas la ESs que están previamente en espera de ser atendidas (en *backlog*) tienen que aguardar el tiempo que toma la transmisión del paquete actual. Tomando los promedios de varios experimentos con los programas realizados, encontramos que este retardo promedio puede ser muy bien aproximado con la siguiente expresión:

$$\bar{i} \cong 2 \cdot M \quad (4)$$

Mientras que el cálculo de \bar{x} e \bar{i} es directo, determinar los valores para \bar{c} y \bar{w} es más complicado. Ambos procedimientos son explicados en las secciones 4.1.2.1 y 4.1.2.3.

4.2.2.1 Calculando el Retardo de Contención (c)

Por simplicidad asumiremos que el número de contenciones fallidas que requiere una ES para tener una contención exitosa sigue una distribución geométrica, tomada de [21]. Así, dada la probabilidad de contención exitosa (P_{sc}) en el periodo de contención de un UL-Frame, el número promedio de minislots (retardo por contención) \bar{c} está dado por la expresión:

$$\bar{c} = C + M \frac{(1 - P_{sc})}{P_{sc}} \quad (5)$$

En donde, C es el tamaño del bloque de contención y M el tamaño total del UL-Frame, ambos medidos en *slots* y haciendo que \bar{n} y \bar{s} tomen el valor de la esperanza matemática del total de usuarios conteniendo y de contendientes con éxito en un bloque respectivamente:

Entonces podemos estimar la probabilidad de éxito en la contención (P_{sc}) en un UL-Frame, mediante:

$$P_{sc} = \frac{\bar{s}}{n} \quad (6)$$

Lo que nos lleva ahora a distraer nuestra atención en la estimación de \bar{N} y \bar{s} .

Denotando a P_{sc} como la probabilidad de que una ES decida contender en un bloque de contención. Por lo tanto, el valor esperado del número de contendientes en un bloque de contención está dado por:

$$\bar{n} = P_c \cdot N \quad (7)$$

y el número esperado de contendientes exitosos por bloque de contención es:

$$\bar{s} = (P_c \cdot N) \cdot P_{sc} \quad (8)$$

Podemos darnos cuenta entonces que hay \bar{t} minislots por cada ciclo de transmisión y un bloque de contención cada M minislots. De ahí que, el promedio del número de bloques de contención \bar{b} por ciclo de transmisión está dado por:

$$\bar{b} = \frac{\bar{t}}{M} \quad (9)$$

Y para evitar incongruencias en la obtención de los cálculos subsecuentes, asumiremos que el sistema está operando en forma constante basado en las aseveraciones mencionadas en los párrafos anteriores. Bajo esta premisa, cada ES activa tiene la misma oportunidad de transmitir cada \bar{t} *slots* y el parámetro \bar{s} puede ser estimado dividiendo el número total de ESs entre el número de bloques de contención por ciclo de transmisión, como sigue:

$$\bar{s} = \frac{N}{\bar{b}} \quad (10)$$

Una vez que tenemos las expresiones anteriores, podemos comenzar a sustituir y desarrollar algunas fórmulas con variables conocidas con el fin de ir obteniendo resultados tangibles, tomando (9), sustituyendo en (10) y desarrollando tenemos:

$$\bar{s} = \frac{N \cdot M}{\bar{t}} \quad (11)$$

Tomando (7), (11) y sustituyendo en (6) resulta:

$$P_{sc} = \frac{N \cdot M}{(P_c \cdot \bar{t})}$$

Pero nos interesa obtener la probabilidad de que una ES contienda en un periodo de contención dado (P_c), por lo tanto desarrollando:

$$P_{sc} = \frac{M}{(P_c \cdot \bar{t})}$$

Y finalmente despejando...

$$P_c = \frac{M}{(P_{sc} \cdot \bar{t})} \quad (12)$$

Ahora bien, definamos la probabilidad de éxito en la contención (P_s) como la probabilidad de que en un número promedio de usuarios solicitando ancho de banda \bar{n} , sólo uno de ellos contienda en un *slot* dado.

Para esto usamos probabilidad binomial, con el número de ESs contendientes y el tamaño del bloque de contención como parámetros. La expresión que nos representa dicha probabilidad se ve simplificada considerablemente dado que, sólo nos interesa por el momento, conocer la probabilidad de contención de un grupo de contendientes en un solo *slot* de contención dado, quedando expresada como:

$$P_s = \binom{\bar{n}}{1} \frac{1}{C} \left(1 - \frac{1}{C}\right)^{\bar{n}-1} \quad (13)$$

Nota: No está permitida la captura de un paquete (probabilidad de captura).

Por consiguiente, el número esperado de ESs que transmiten un REQ exitoso, es decir, contienen exitosamente en el actual UL-Frame puede ser calculado de la siguiente forma:

$$\bar{s} = (C \cdot P_s) \quad (14)$$

Finalmente podemos utilizar variables conocidas para determinar la probabilidad de que una ES tenga una contención exitosa en un UL-Frame, sustituyendo (7) y (14) en (6):

$$P_{sc} = \frac{(C \cdot P_s)}{(P_c \cdot N)} \quad (15)$$

En la propuesta hay un detalle de complejidad interesante, siendo observadores notemos que para calcular el retardo de contención \bar{c} , en un UL-Frame en la expresión (5) es necesario conocer el valor P_{sc} , que a su vez depende de conocer los valores de P_c y P_s como se hace notar en la ecuación (15).

Además de que el cálculo de P_c en (12) depende del retardo total generado por la transmisión de un paquete t cuyo valor se vuelve dependiente de los retardos de contención \bar{c} y de asignación de ancho de banda una vez que la ES recibió un ACK \bar{w} , por parte de la EB, tal y como se muestra en la ecuación (1). Ambos valores son incógnitas.

Hasta este punto, podemos mencionar que los autores de las referencias [1], [2] y [22] enfrentaron el mismo problema en el análisis de eficiencia realizada sobre una red GPRS con servicios IP. De esos trabajos, trasladamos algunos criterios como, el de asumir un valor inicial para \bar{w} y entonces determinar \bar{c} .

Aunque fue necesario realizar ajustes adicionales en \bar{w} , cuando el valor asumido para esta variable no satisfacía todas las condiciones del sistema. Ellos también propusieron un procedimiento iterativo para determinar el valor de P_{sc} . Sin embargo, y aunque el modelo matemático que usaremos es distinto en mucho sentidos, decidimos probar la solución propuesta en dichos trabajos en nuestro contexto con la red basada en IEEE 802.16.

El hecho de que en nuestro modelo no se permitiera la captura de un paquete para calcular la probabilidad de tener una contención exitosa en un *slot* dado (P_s) es una diferencia fundamental, la cual nos permite dar un paso adicional en el desarrollo.

Tal y como se señala en las referencias [5] y [16], asumimos que \bar{w} es conocida. El procedimiento inicia cuando damos un valor inicial estimado para la probabilidad de éxito en la contención (P_{sc}).

Entonces utilizando las ecuaciones (1), (5), (12) y (13), determinamos P_c y P_s . Estos valores se utilizan en la expresión (15) para entonces determinar el siguiente valor para P_{sc} . Con este nuevo valor, se repite el procedimiento (haciéndose iterativo), comenzando en la expresión (5). El cual se detiene cuando dos valores consecutivos para la probabilidad de éxito en la contención P_{sc} están dentro de un cierto umbral determinado.

El valor final de la probabilidad P_{sc} nos permite entonces calcular el valor para el retardo de contención, usando nuevamente la expresión (5).

4.2.2.2 Integrando y generalizando el modelo para el retardo de contención

En esta sección trabajaremos en construir una ecuación que represente el modelo que describimos en la sección 4.2.2.1, con el fin de poder utilizar herramientas informáticas y así obtener resultados para comprobar que el planteamiento realizado es correcto.

Definamos con el subíndice n como la n -ésima iteración de nuestro procedimiento. De este modo las expresiones (5) y (1) se convierten en las siguientes:

$$\bar{c}_n = C + M \frac{1 - P_{SCn}}{P_{SCn}} \quad (16)$$

$$t_n = \bar{i} + \bar{c}_n + \bar{w} + \bar{x} \quad (17)$$

Entonces sustituyendo la expresión (16) en la (17), obtenemos:

$$t_n = \bar{i} + C + M \frac{1 - P_{SCn}}{P_{SCn}} + \bar{w} + \bar{x} \quad (18)$$

Aplicando el mismo criterio usado en (5) y (1) de expresar con el subíndice n a la n -ésima iteración sobre la ecuación (12), tenemos:

$$P_{Cn} = \frac{M}{(P_{SCn} \cdot \bar{t}_n)} \quad (19)$$

La cual combinada con la expresión (18) produce:

$$P_{Cn} = \frac{1}{P_{SCn} \left[\left(\frac{\bar{i} + C + \bar{w} + \bar{x}}{M} \right)^{-1} + 1 \right]}$$

Con fines prácticos

$$P_{Cn} = \frac{1}{(P_{SCn} \cdot \beta + 1)} \quad (20)$$

En donde:

$$\beta = \left(\frac{\bar{i} + C + \bar{w} + \bar{x}}{M} \right) - 1$$

Usamos la expresión anterior (20) y la sustituimos en la ecuación (7) con el fin de obtener el promedio del número de usuarios esperados en la n -ésima iteración:

$$\bar{n}_n = \left[\frac{1}{(P_{SCn} \cdot \beta + 1)} \right] \cdot N \quad (21)$$

Continuamos determinando el valor esperado de la probabilidad de tener éxito en la contención en la n-ésima iteración, para ello sustituimos (21) en (13). Recordando que esta probabilidad se refiera a un solo contendiente en un minislot cualquiera. Es decir, no hay probabilidad de captura, quedando:

$$P_{sn} = \left[\frac{N}{(P_{SCn} \cdot \beta + 1)} \right] \frac{1}{C} \left(1 - \frac{1}{C} \right) \left[\frac{N}{(P_{SCn} \cdot \beta + 1)} \right]^{-1} \quad (22)$$

Concluyendo que la intención de todo el desarrollo descrito en esta sección de nuestro trabajo es la de encontrar la probabilidad de que una ES contienda exitosamente en un UL-Frame cualquiera, procedemos a sustituir las expresiones (20) y (22) en (15), quedando como resultado la siguiente ecuación en diferencias que puede utilizar con ese fin.

$$P_{SC(n+1)} = \left(1 + \frac{1}{C} \right) \left[\frac{N}{(P_{SCn} \cdot \beta + 1)} \right]^{-1} \quad (23)$$

En resumen, dados ciertos valores iniciales arbitrarios para el retardo que se genera mientras la ES espera a que la EB le asigne ancho de banda una vez que una solicitud REQ es exitosamente recibida por la EB y para la probabilidad de que una ES contienda exitosamente en un UL-Frame. Entonces iniciamos a iterar con la expresión (23) hasta que obtenemos un valor consistente (cuando el algoritmo converge) para la probabilidad de contender exitosamente (P_{sc}).

Con este valor de P_{sc} podemos calcular el retardo de contención \bar{c} utilizando la expresión (5), ya sólo tenemos pendiente determinar el retardo de espera \bar{w} .

4.2.2.3. Calculando el retardo de espera (\bar{w})

Recordando, en la sección 4.2.2.1 el valor para el retardo de espera \bar{w} de una ES mientras la EB le asigna ancho de banda una vez que una solicitud REQ es exitosamente recibida fue asumido.

Por esta razón es necesario determinar si este valor de \bar{w} requiere ser ajustado.

Con este fin, procederemos a calcular el número promedio de minislots disponibles para que cada ES transmita en un ciclo de transmisión promedio. Denotemos ese número con \bar{d} .

El cual está representado por:

$$\bar{d} = \frac{\bar{t} + C \cdot \bar{b}}{N} \quad (24)$$

Adicionalmente, sabemos que el número promedio esperado de minislots de transmisión requeridos por cada ES es \bar{x} . De aquí que, el número esperado de minislots de transmisión en un ciclo de transmisión debe ser al menos $(N \cdot \bar{x})$.

Por lo tanto, si $\bar{d} > N \cdot \bar{x}$

Se propone que el valor inicialmente asumido para el retardo de espera w se incremente de acuerdo a:

$$w_{m+1} = w_m + 1 \quad (25)$$

Y entonces se procedería a calcular nuevamente el valor del retardo por contención, descrito en la sección 4.2.2.1.

Pero, si $\bar{d} < N \cdot \bar{x}$

Entonces podremos tomar el correspondiente valor de w y proceder a calcular la eficiencia del sistema.

Los resultados del análisis de eficiencia usando el modelo matemático con ecuaciones en diferencias descrito, así como una comparación teórica con el modelo anterior usado para GPRS son mostrados a detalle en el Capítulo 5.

CAPÍTULO 5

RESULTADOS OBTENIDOS

Usando el nuevo modelo matemático basado en ecuaciones en diferencias, tenemos ya una forma teórica alterna para determinar todos los valores de las variables del ciclo de transmisión de un paquete en nuestra red basada en IEEE 802.16-2004. Por lo cual, en este capítulo procederemos a obtener resultados numéricos y a elaborar un análisis de eficiencia de nuestro sistema inalámbrico en función de varios parámetros, tales como:

- El tamaño (duración) del UL-Frame.
- El número de usuarios en la red.
- El tamaño de los paquetes que transmiten las ES.
- El número de *slots* de contención en el UL-Frame.

5.1 ESCENARIOS TEÓRICOS

A continuación analizamos la eficiencia obtenida en función del número de ESs activas en forma simultánea con la cantidad de *slots* de contención como parámetro, aplicando el modelo matemático descrito anteriormente, sobre una red basada en el estándar IEEE 802.16.

Es importante mencionar que para obtener los resultados teóricos se utilizaron 2 programas. Uno para obtener la eficiencia del sistema (*throughput*) y otro para determinar la cantidad de bloques que se requieren para transmitir paquetes de cierto tamaño (fijo para todos los usuarios de cada simulación), mismos que se anexan en los apéndices B y C respectivamente.

Ambos programas contienen los comentarios respectivos con el fin de explicar la lógica utilizada en la programación y así poderla asociar con el modelo y con la obtención teórica de los bloques. El primer parámetro importante utilizado para obtener los resultados teóricos fue la longitud del UL-Frame. El tamaño de cada minislots es fijado a 16 *bytes*. De aquí, se originaron 2 criterios de análisis:

- UL-Frame de 2 *ms* de duración, es decir, $\mathcal{M}=450$ minislots de longitud.
- UL-Frame de 1 *ms* de duración, es decir, $\mathcal{M}=225$ minislots de longitud
 - En ambos caso el canal es de 28.8 *Mbps*.

En ambos casos se grafica la eficiencia de una red IEEE 802.16-2004 desde 4 hasta 100 usuarios aumentando de 4 en 4, se toma como parámetro el número de *slots* de contención (C). Todas las ESs utilizan la técnica de BE para solicitar ancho de banda a la EB. De igual forma utilizamos un retardo inicial de 1 minislot ($w_0=1$), la probabilidad inicial de contender exitosamente en un UL-Frame se toma como $P_{SC0}=0.01$. Todos los usuarios transmiten ráfagas simétricas en la capa MAC.

Al momento de que los paquetes que transmite cada ES son codificados (usando Reed -Solomon) cumpliendo con la norma de un sistema IEEE 802.16, se utilizan *Codewords* (palabras codificadas) de 245 *bytes*, una paridad Reed-Solomon de 10 *bytes*, un valor de preámbulo de 4 *bytes* y una banda de guarda de 6 *bytes*. El tiempo esperado para transmitir cada paquete (x) dependerá del tamaño fijo utilizado en cada escenario para todos los usuarios.

5.1.1 ANÁLISIS DE RESULTADOS TEÓRICOS PARA UN UL-FRAME = 1 ms

Todos los escenarios que se grafican en esta sección consideran los parámetros mostrados en la Tabla 5.1:

PARÁMETRO	VALOR
<i>Slots</i> de contención (C)	Variable en función al tamaño del paquete a transmitir por las ESs
Tamaño de un minislot en el UL-Frame	16 <i>bytes</i>
Duración del UL-Frame	1 <i>ms</i> / 225 minislots
<i>Reed Solomon</i> entre <i>Codewords</i>	6 <i>bytes</i>
<i>Reed Solomon</i> longitud de <i>Codewords</i>	10 <i>bytes</i>
<i>Reed Solomon</i> preámbulo	6 <i>bytes</i>
<i>Reed Solomon</i> banda de guarda	4 <i>bytes</i>
Encabezado de Capa MAC para 802.16	6 <i>bytes</i>
RS tamaño de las asignaciones (<i>grants</i>)	245 <i>bytes</i>
Número máximo de ESs activas (N)	100
Tipo de Tráfico	CBR
Ancho de banda del Canal	28.8 Mbps
Valor inicial P_{SC} (Probabilidad de contender con éxito)	0.01
Retardo inicial (w_0) en ser asignado un <i>grant</i> a una ES	1 minislot

Tabla 5.1– Parámetros utilizados para obtener los resultados teóricos con un UL-Frame de 1ms.

5.1.1.1 Resultados con paquetes de tamaño fijo (300 bytes)

Para el caso de un mensaje de 300 *bytes*, se requieren 21 bloques codificados con la técnica FEC Reed Solomon descrita. Ver apéndice C. Por esta razón vamos aumentando de 21 en 21 los *slots* de contención, dado que así podemos lograr que se desperdicien la menor cantidad de minislots posibles, y en lugar de ello las ESs las ocupen para solicitar ancho de banda o para transmitir sus datos.

Como se muestra en la Figura 5.1, para el número inicial de 30 *slots* de contención por UL-Frame, la eficiencia del sistema es de aproximadamente el 88 % del total de la capacidad del canal, mismo que se alcanza aproximadamente cuando hay 24 ES activas en la red (coincidiendo con el número de *slots* de contención). Dicha eficiencia prácticamente se mantiene constante hasta las 60 ESs, de ahí se observa una rápida caída dado el excesivo número de colisiones en cada UL-Frame.

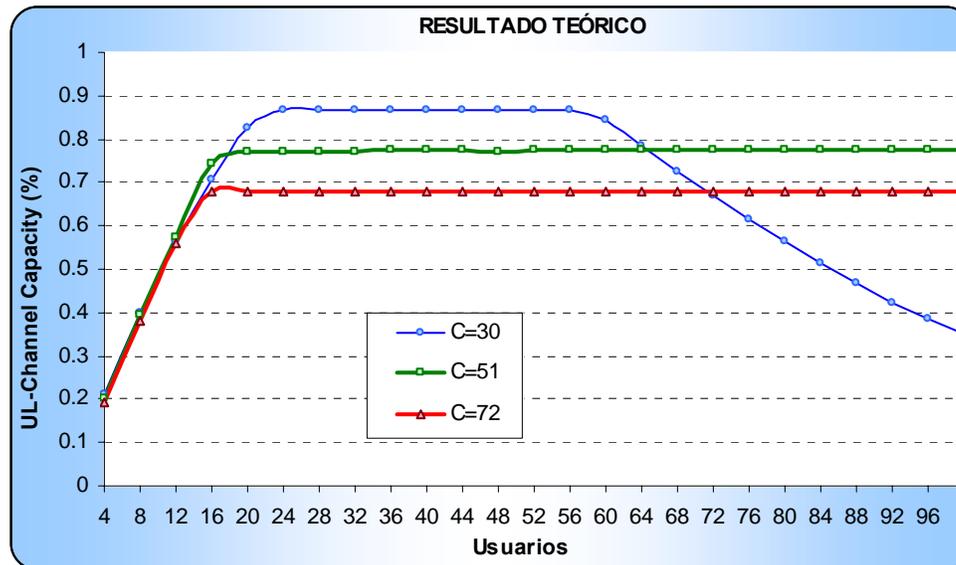


Figura 5.1 – Eficiencia del sistema en función del número de ESs activas en la red con UL-Frame= 225 minislots

Si no se toma la precaución arriba mencionada y aumentáramos arbitrariamente los *slots* de contención para cada caso la eficiencia del sistema disminuye. En todos los casos subsecuentes se utiliza el mismo criterio de aumentar los *slots* de contención en función al tamaño del mensaje que transmiten las ES y a su vez al número de bloques codificados requeridos para dicha transmisión en la red IEEE 802.16 propuesta.

Al incrementar el número de *slots* de contención (C) a 51 se observa un comportamiento interesante dado que se obtiene una eficiencia máxima menor (77 %) pero, a diferencia de cuando se tienen 30 *slots* de contención, está se mantiene constante a partir de que hay 20 ESs activas y hasta los 100 ESs.

En el caso de tener 71 *slots* de contención se ve reflejado un comportamiento muy similar a cuando $C=51$, solamente que la máxima eficiencia del sistema es solamente del 67 % y es rápidamente alcanzada a partir de cuando hay 16 ESs en la red, sosteniéndose inclusive para muchos más usuarios. De aquí, nuestra primera prueba concluyente de que no por el hecho de tener muchos *slots* de contención la eficiencia del sistema mejora.

En los tres casos (con $C=30$, $C=51$ y $C=72$) se puede percibir un evidente “desperdicio” de recursos cuando se tienen pocos usuarios en la red. Esto se explica porque cuando el número de usuarios activos en la red es menor al número de *slots* de contención (por ejemplo con $C=30$ y $N=8$), hay 22 minislots que no son utilizados. Entonces la eficiencia del sistema no es óptima debido a que esos 22 *slots* de contención implican prácticamente un 10% de los recursos no utilizados (usando UL-Frame de 225 minislots).

Otro dato muy importante que debe ser considerado al momento de diseñar un sistema de este tipo es el retardo presente entre cada ciclo de transmisión completo.

En las Figuras 5.2 (a) y 5.2 (b) se muestra el retardo, en minislots y en segundos respectivamente, presente entre cada ciclo de transmisión, para la misma cantidad de minislots de contención usados en la Figura 5.1. Se considera que se están transmitiendo también paquetes de 300 *bytes*.

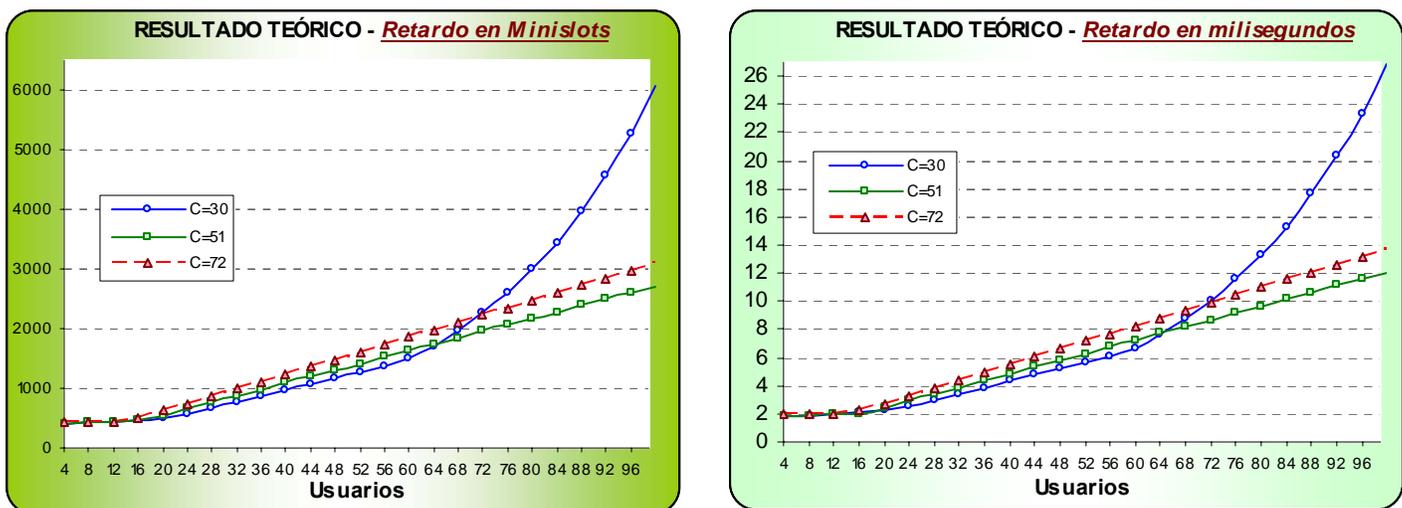


Figura 5.2 – Retardo (a -izquierda) en Minislots y (b-derecha) en milisegundos entre cada ciclo de transmisión en función del número del número de ESSs activas.

Podemos mencionar que aunque en el caso de usar $C=30$ minislots se observa un dramático incremento en el retardo a partir de los 60 usuarios, la eficiencia del sistema no se ve seriamente afectado porque en el peor escenario (100 ESSs activas) el retardo no es mayor a 27 milisegundos. Aunque este valor de retardo en casos de alta utilización podría llegar a impactar la eficiencia general del sistema.

Para convertir minislots a segundos, únicamente se divide la duración del UL-MAP (1 *ms*) entre el número de minislots en dicho mapa (225). Con esto, se obtiene que cada minislot tiene una duración de 4.4×10^{-6} segundos. Posteriormente al multiplicar el número de *slots* de retardo por su duración unitaria en segundos se obtiene el resultado en milisegundos.

Con fines ilustrativos se anexa los resultados numéricos para el experimento mencionado. Dada la extensión de los datos a la cantidad de variantes que se aplicarán posteriormente, este será el único caso en el cual se mostrarán los resultados teóricos completos (Tablas 5.y y 5.3). En los demás, únicamente se mostrarán las gráficas.

\mathcal{N}	C=30	C=51	C=72
4	20.94%	20.19%	19.31%
8	39.89%	39.36%	38.01%
12	56.53%	57.46%	56.06%
16	70.77%	74.42%	67.96%
20	82.52%	77.27%	67.98%
24	86.52%	77.29%	67.94%
28	86.63%	77.25%	67.95%
32	86.57%	77.25%	67.97%
36	86.63%	77.33%	67.95%
40	86.58%	77.30%	67.95%
44	86.60%	77.32%	67.96%
48	86.64%	77.29%	67.96%
52	86.64%	77.31%	67.97%
56	86.67%	77.31%	67.98%
60	84.25%	77.31%	67.98%
64	78.49%	77.30%	68.00%
68	72.69%	77.31%	67.98%
72	66.98%	77.31%	67.97%
76	61.47%	77.33%	67.98%
80	56.22%	77.33%	67.99%
84	51.26%	77.31%	67.98%
88	46.63%	77.31%	67.99%
92	42.32%	77.31%	68.00%
96	38.35%	77.32%	68.00%
100	34.69%	77.32%	67.98%

Tabla 5.2 – Resultados numéricos completos que generan la Figura 5.1.

\mathcal{N}	Retardo en Minislots			Retardo en milisegundos Cada minislots=0.001 s/225= 0.00000444444 s		
	C=30	C=51	C=72	C=30	C=51	C=72
4	401.079415	416.0099101	435.007612	1.782573397	1.848931085	1.933365231
8	421.192953	426.8084559	442.0373076	1.871966807	1.896924574	1.964608291
12	445.817734	438.5879063	449.4966243	1.98141017	1.949277634	1.997760777
16	474.765682	451.4661451	494.4212092	2.110067588	2.006514194	2.197425399
20	508.964128	543.5767975	617.8500803	2.26206053	2.415894462	2.745997611
24	582.527502	652.0708558	741.8260511	2.589008529	2.898089794	3.297001375
28	678.776751	761.1189133	865.3959662	3.016782542	3.382747343	3.846200448
32	776.253084	869.9127852	988.6110506	3.450010259	3.866275179	4.393822498
36	872.726969	977.6678029	1112.527497	3.878782652	4.34518589	4.944561707
40	970.201675	1086.62433	1236.206603	4.312003133	4.829436636	5.494246074
44	1066.92007	1195.049527	1359.715235	4.741862233	5.311325922	6.043172779
48	1163.37837	1304.238435	1483.126483	5.170565344	5.79660947	6.591666668
52	1260.3536	1412.515725	1606.519716	5.60156594	6.277841391	7.140080488
56	1356.9425	1521.235355	1729.981172	6.030849537	6.76103926	7.688797519
60	1495.6235	1629.781423	1853.604133	6.647208887	7.243465747	8.238232354
64	1712.32544	1738.568079	1976.489759	7.610327666	7.726961511	8.784390145
68	1964.50236	1847.03944	2100.746718	8.731112862	8.209055971	9.336642741
72	2257.23777	1955.669601	2224.491807	10.03215783	8.691856202	9.886620367
76	2596.34135	2063.962925	2347.850043	11.53928337	9.173159382	10.43487864
80	2988.45059	2172.454842	2470.954762	13.28198932	9.6553452	10.98201018
84	3441.18278	2281.713302	2594.947366	15.29413038	10.14093787	11.53308787
88	3963.26856	2390.340983	2717.978425	17.61450933	10.62372708	12.07989203
92	4564.6881	2498.978309	2841.206057	20.2874824	11.10655915	12.62756985
96	5256.9098	2607.307233	2964.7968	23.36402019	11.58802056	13.17686149
100	6053.06341	2716.057043	3088.925278	26.90247714	12.07135256	13.72854306

Tabla 5.3 – Resultados numéricos que generan las Figuras 5.2 (a) y 5.2 (b).

5.1.1.2 Resultados con paquetes de distinto tamaño, con $C=30$

Procediendo a realizar un análisis más detallado sobre el comportamiento de la red, decidimos crear un escenario teórico en el que se presenta una comparación variando el tamaño de los paquetes que están transmitiendo las ESs.

De igual forma que en las figuras anteriores se utilizó un UL-Frame de 225 minislots, sólo que en este experimento se fijaron los *slots* de contención en $C=30$ para todos los casos. El tamaño de los paquetes utilizados fue 64, 128, 256, 300 (nuestro parámetro base inicial), 512 y 1500 bytes.

La idea de este escenario fue la tener en una sola figura el comportamiento del sistema cuando se están transmitiendo paquetes con tamaños distintos pero con un número de *slots* de contención fijo ($C=30$), este valor es el que usamos como punto de partida en todos los ejercicios teóricos.

Dado que para los siguientes valores de C , la variación del número de *slots* de contención se genera en función al tamaño del paquete (y de los bloques codificados requeridos para transmitirlo).

En la Figura 5.3, se mantienen fijos los valores del UL-Frame (225 minislots) y de *slots* de contención ($C=30$) para los distintos tamaños de paquetes transmitidos por las ESs.

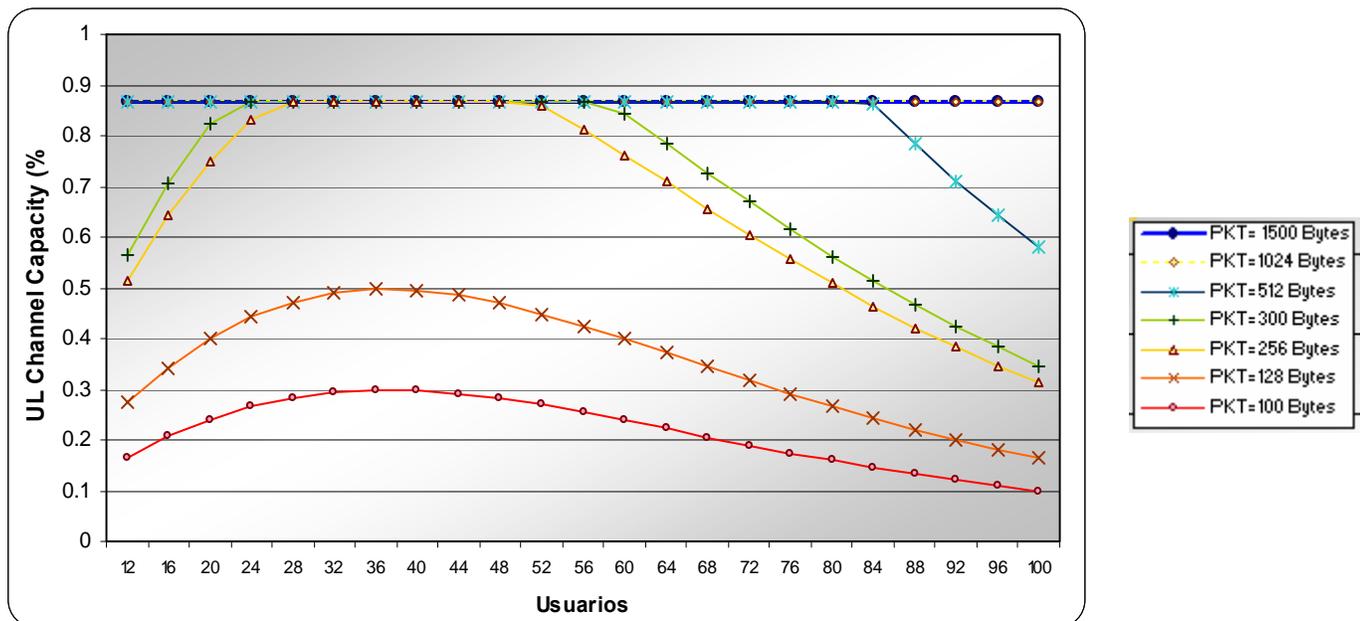


Figura 5.3 – Eficiencia del sistema IEEE 802.16 utilizando el modelo propuesto variando el tamaño de los paquetes transmitidos. Usando un UL-MAP de 225 minislots y 30 slots de contención para todos los caso.

Como el comportamiento de la eficiencia es muy similar para paquetes grandes (512, 1024 y 1500 *bytes*), no es posible apreciar claramente estas tres gráficas (las tres superiores) en una impresión en blanco y negro.

De aquí podemos concluir que, el comportamiento de la eficiencia del sistema es muy similar para todos los casos en los que se transmiten paquetes menores a 512 Kb, se comienza con poca eficiencia la cual va incrementando hasta un punto en donde es máxima (este punto varía dependiendo el tamaño del paquete).

A partir de ese punto máximo la eficiencia comienza a descender dada la cantidad de colisiones. Solamente en los casos de los paquetes de 1024 y 1500 *bytes* se observa el mismo comportamiento sin importar el número de ESs activas en la red.

Ahora analizando la Figura 5.4 en donde se muestra el efecto en el retardo del sistema al transmitir paquetes grandes, podemos decir que aunque la eficiencia del sistema es muy buena y constante el retardo aumenta considerablemente (en forma lineal) en función al número de ESs activas.

Al momento y basándonos en las Figuras 5.3, 5.4 y 5.5 podemos mostrar que la eficiencia del sistema analizado, depende del número de ESs activas en la red como al tamaño de los paquetes que son transmitidos.

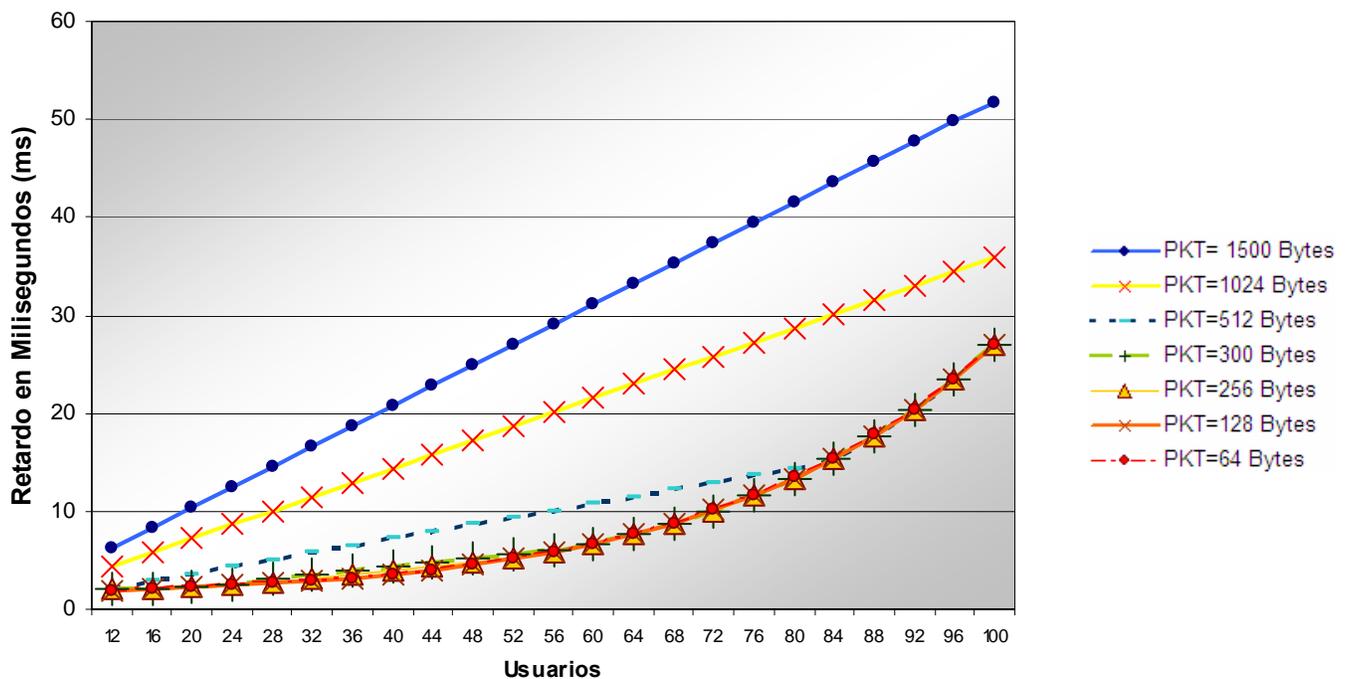


Figura 5.4 – Retardo del sistema IEEE 802.16 utilizando el modelo propuesto variando el tamaño de los paquetes transmitidos. Usando un UL-MAP de 225 minislots y 30 slots de contención para todos los casos.

La Figura 5.5 muestra más claramente las diferencias en la eficiencia entre los paquetes más grandes. Esto se debe a que el ancho de banda se aprovecha más si lo llenamos con pocos paquetes de tamaño grande que con muchos de tamaño pequeño.

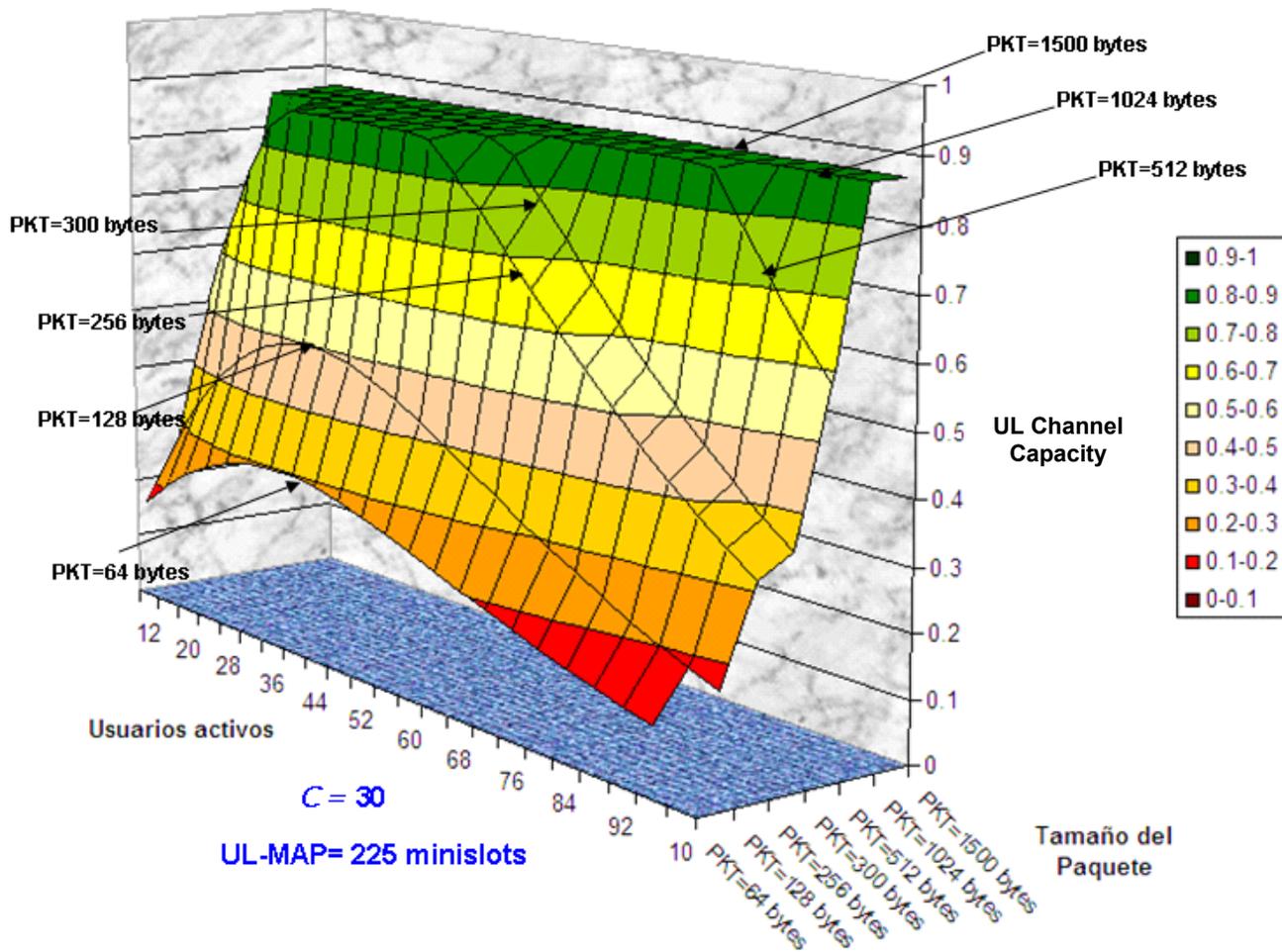


Figura 5.5 –Eficiencia del sistema IEEE 802.16 utilizando el modelo propuesto variando el tamaño de los paquetes transmitidos.

En primera instancia podríamos concluir que el escenario ideal se daría cuando las ESs transmitan únicamente paquetes grandes, dado que así el sistema se comporta de una manera altamente eficiente sin importar el número de ESs activas.

Pero después de analizar el impacto al retardo del sistema, podemos también concluir que se deberán transmitir paquetes grandes sólo en el caso de que la aplicación que estén utilizando las ESs soporte un retardo de al menos 60 ms.

5.1.1.3 Resultados variando el tamaño de los paquetes y también C

De los análisis previos, podemos a estas alturas concluir contundentemente que la eficiencia de la red es directamente proporcional al tamaño de los paquetes que transmitan las ESs, ahora vamos a experimentar adicionalmente con el número de *slots* de contención. De la Figura 5.6 podemos deducir que si administramos adecuadamente los *slots* de contención podremos obtener mejores rendimientos con un tamaño de paquete mediano (entre 128 y 512 bytes). Recordemos que ahora los *slots* de contención son determinados en base a la cantidad de bloques codificados requeridos para transmitir el paquete (según su tamaño).

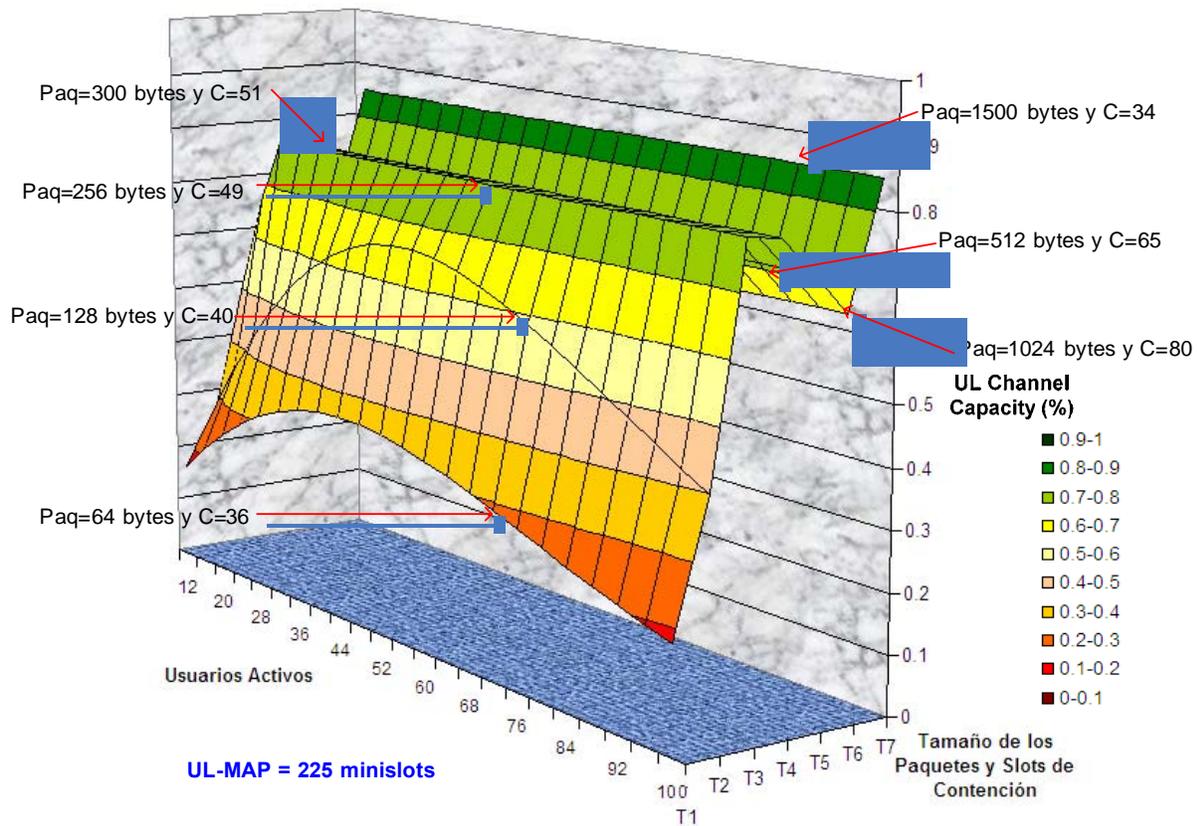


Figura 5.6 – Eficiencia del sistema IEEE 802.16 utilizando el modelo propuesto variando el tamaño de los paquetes transmitidos y los slots de contención en función a cada tamaño de paquete.

T1= Paquetes de 64 bytes y 36 slots de contención.
T2= Paquetes de 128 bytes y 40 slots de contención.
T3= Paquetes de 256 bytes y 49 slots de contención.
T4= Paquetes de 300 bytes y 51 slots de contención.
T5= Paquetes de 512 bytes y 65 slots de contención.
T6= Paquetes de 1024 bytes y 80 slots de contención.
T7= Paquetes de 1500 bytes y 34 slots de contención.

Tabla 5.4 –Equivalencias para el eje “y” de la Figura 5.6.

Dado que visualmente es mucho mejor analizar el comportamiento de la eficiencia del sistema, con todas sus variantes en una sola gráfica, continuaremos mostrando los escenarios experimentales propuestos en las siguientes imágenes en 3ª dimensión.

Ahora también es importante hacer notar que no por tener una mayor cantidad de *slots* de contención, con el fin de atender a la mayor cantidad de usuarios posibles (recepción exitosa de mensajes REQ), vamos a tener una mejor eficiencia del canal de transmisión.

Observamos en la Figura 5.7, que la línea que representa al comportamiento del sistema cuando se transmiten paquetes de 512 *bytes* y se utilizan 100 *slots* de contención (ver detalle en color rojo indicando la línea) nos describe una baja eficiencia (aproximadamente del 50 %). Inclusive si lo comparamos con la misma línea en las Figuras 5.5 y 5.6, y el único parámetro diferente fue la cantidad de *slots* de contención.

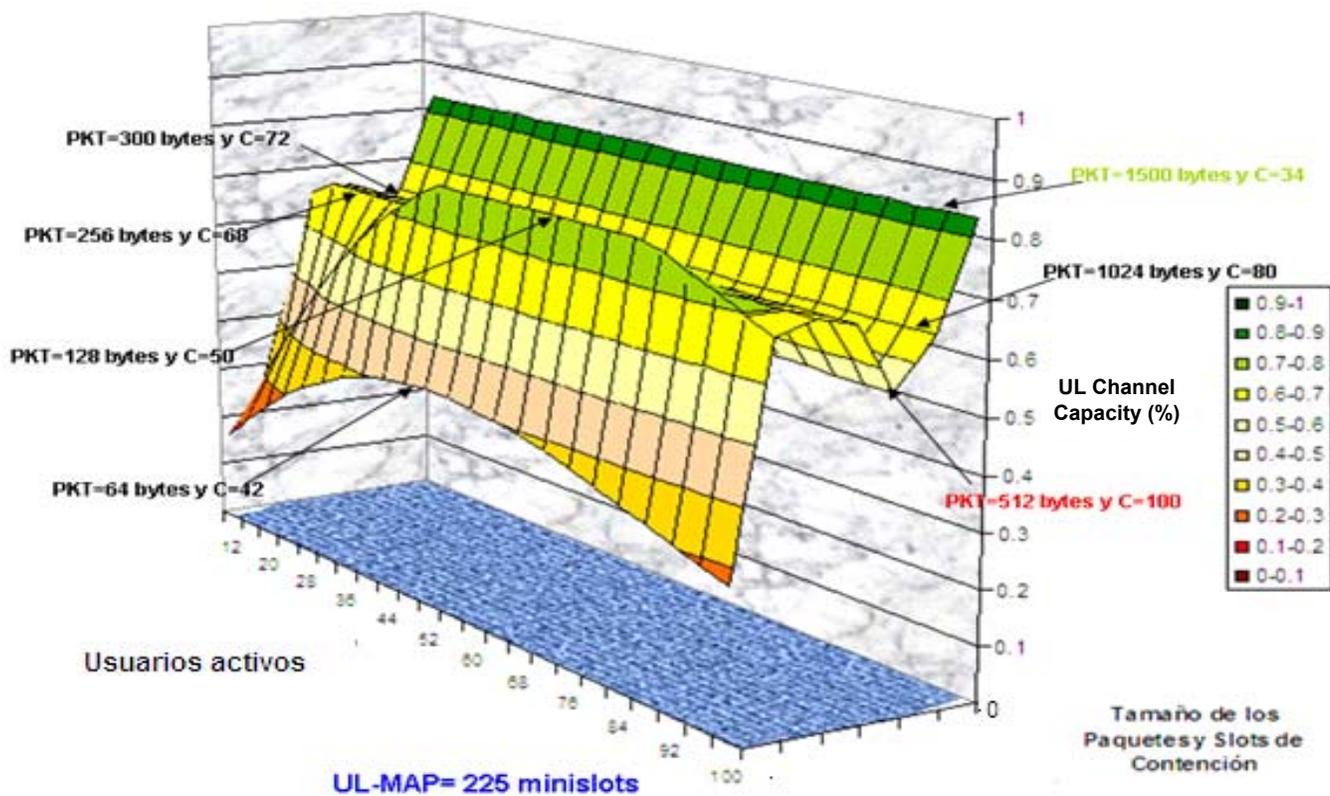


Figura 5.7 – Eficiencia del sistema IEEE 802.16 utilizando el modelo propuesto variando el tamaño de los paquetes transmitidos y los slots de contención en función a cada tamaño.

Aquí se ejemplifica claramente el comportamiento del sistema cuando se incrementa demasiado el número de *slots* de contención, lo que nos lleva a concluir que el comportamiento de la red va a depender de 3 parámetros principales:

- El número de ESs activas
- El tamaño de los paquetes que estén transmitiendo dichas ESs
- Los *slots* de contención que la EB envíe en los UL-MAPs.

5.1.2 ANÁLISIS DE RESULTADOS TEÓRICOS PARA UN UL-FRAME = 2 ms

Para los siguientes resultados el único parámetro que cambiará en los escenarios propuestos será el tamaño del UL-Frame, tal y como se muestra en la Tabla 5.5.

PARÁMETRO	VALOR
Slots de contención (C)	Variable en función al tamaño del paquete a transmitir por las ESs
Duración del UL-Frame	2 ms / 450 minislots

Tabla 5.5- Parámetros que cambiarán en relación a la Tabla 5.1 para obtener los resultados teóricos con un UL-Frame de 2ms.

Una vez que reproducimos todas las gráficas que se obtuvieron en la sección 5.1.1 y con el fin de poder compararlas con las nuevas (con el nuevo tamaño del UL-Frame), mezclaremos los valores obtenidos para ambos tamaños de UL-Frame

Con esto tendremos suficiente información para dar una buena recomendación sobre el manejo de los *slots* de contención y el manejo dinámico del tamaño de los paquetes para el caso de una red basada en el estándar IEEE 802.16-2004 con hasta 100 usuarios activos.

En la Figura 5.8 se muestran escenarios con los valores teóricos obtenidos en condiciones semejantes para ambos tamaños de UL-Frame(1 y 2 ms) y transmitiendo paquetes de 300 bytes.

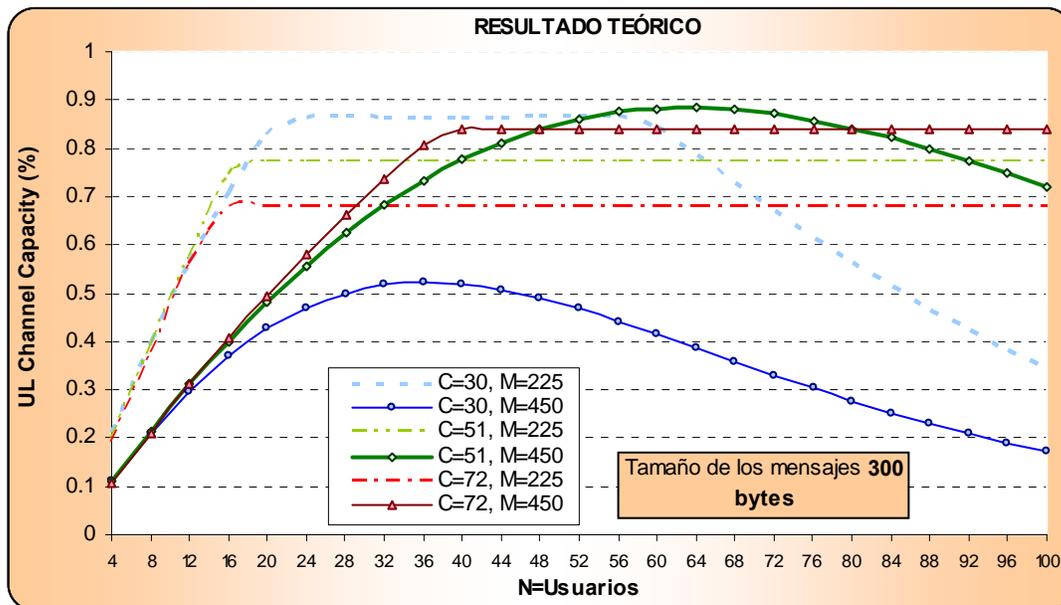


Figura 5.8 – Comparación de escenarios con UL-MAPs de 225 y 450 minislots.

De todas las gráficas mostradas podemos concluir lo siguiente:

- Mejor desempeño (pico de casi 89 %) entre 52 y 80 ESs.
 - UL-Frame= 2 ms y C= 51.
- Más rápida estabilización con el mejor desempeño
 - UL-Frame=1 ms y C=51 (77 % continuo a partir de 20 ESs)
 - UL-Frame= 1 ms y C= 72 con 68 % continuo a partir de 16 ESs).

Pero recordemos que otro valor muy importante a considerar en una red con muchos usuarios demandando recursos continuamente es, el retardo. Procedamos a hacer un análisis sobre el mismo, para ambos escenarios.

En las Figuras 5.9 (a) y 5.9 (b) se muestra el retardo del sistema en función del número de usuarios, expresado en minislots y milisegundos respectivamente. Se presenta en ambas gráficas una comparación al utilizar los dos tamaños para el UL-MAP.

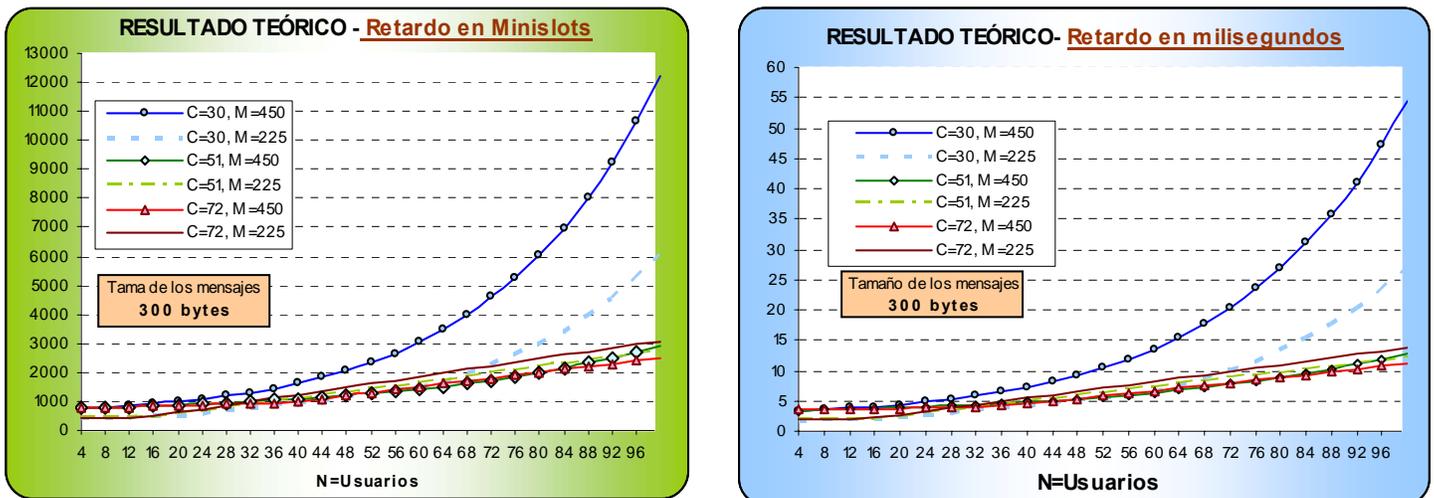


Figura 5.9 – Retardo (a -izquierda) en Minislots y (b-derecha) en milisegundos entre cada ciclo de transmisión en función del número de ESs activas. Comparación con el UL-MAP de 1 y 2 ms, usando los mismos slots de contención para ambos casos, se fija el tamaño de los paquetes en 300 bytes.

Ambas gráficas describen un comportamiento similar para todos los casos (no importando tamaño del UL-MAP y slots de contención) para pocos usuarios (20). Y continúa aumentando similarmente en función a la cantidad de usuarios activos, con excepción del escenario de 30 slots de contención para un UL-MAP de 2 ms, en donde se observa un aumento mucho más considerable en el retardo.

Aunque el retardo aumenta más rápidamente cuando $C=30$, podemos decir que el retardo no afectaría la eficiencia de una aplicación en tiempo real, aún para el caso de tener 100 usuarios activos en forma simultánea dado que el retardo máximo alcanzado por el sistema es de 55 milisegundos.

La Figura 5.10 (de igual forma que la Figura 5.3) describe la eficiencia del ancho de banda del sistema en función de la cantidad de usuarios activos con 30 *slots* de contención, variando el tamaño de los mensajes que se transmiten.

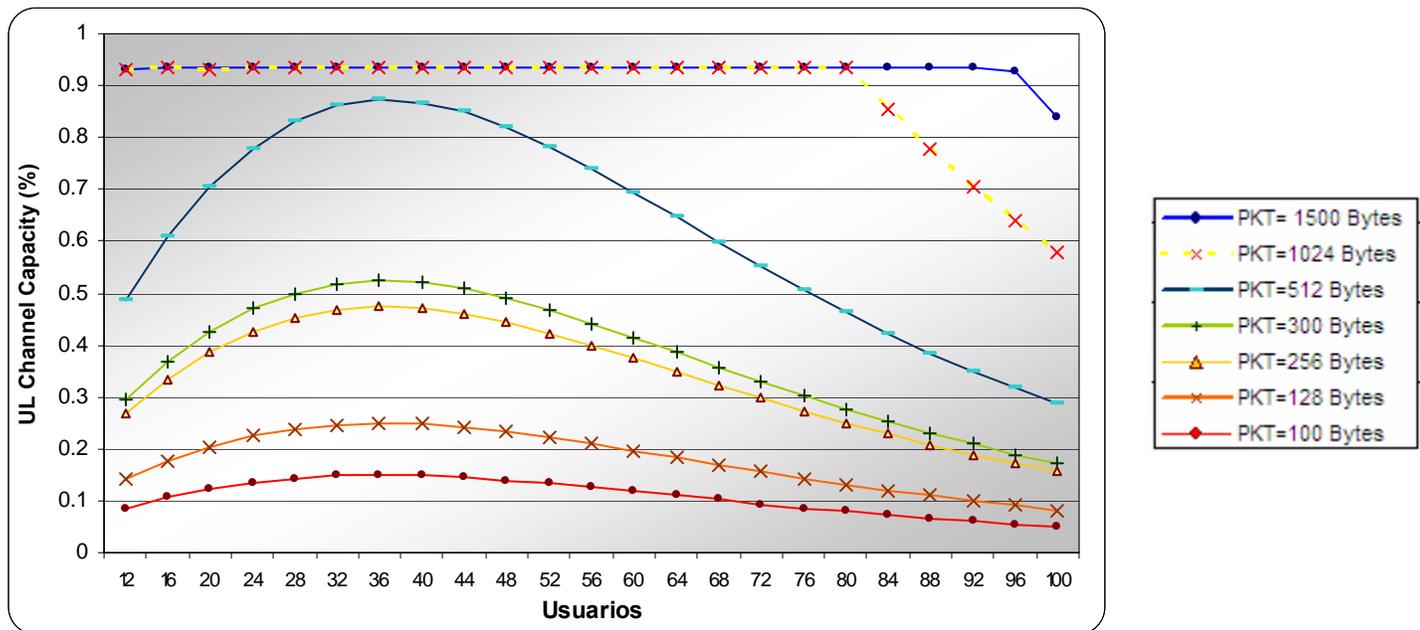


Figura 5.10 – Eficiencia del sistema IEEE 802.16 utilizando el modelo propuesto variando el tamaño de los paquetes transmitidos, usando un UL-MAP de 450 minislots.

Cuando las ESs transmiten paquetes de mayor tamaño, con un UL-MAP más grande se puede observar que la eficiencia del ancho de banda del sistema aumenta, aunque este incremento se ve reflejado negativamente en el retardo.

Y esto se debe a que la EB puede otorgar prácticamente todo el ancho de banda a pocos usuarios, dado que con pocos paquetes se usa todo el ancho de banda disponible dado su tamaño. En este caso el sistema alcanza casi un 95 de eficiencia.

5.1.3 ANÁLISIS DE RESULTADOS DE SIMULACIÓN PARA UN UL-FRAME = 2 ms

Después de haber obtenido algunos resultados teóricos interesantes, el siguiente paso será el probar el modelo matemático propuesto en el capítulo 4. Esto a través de un software que nos ayude a simular el comportamiento de una red basada en el protocolo IEEE 802.16, con esto extrapolaremos los resultados obtenidos.

5.1.3.1 Modelo de simulación

En esta sección describiremos un modelo de simulación detallado de la capa MAC para una red basada en IEEE 802.16 que fue implementado sobre OPNET Modeler® Versión 11 por un equipo multidisciplinario de la Facultad de Ingeniería, tal y como se describe en la referencia [5]. Durante las simulaciones emplearon un modelo jerárquico como se muestra en la Figura 5.11.

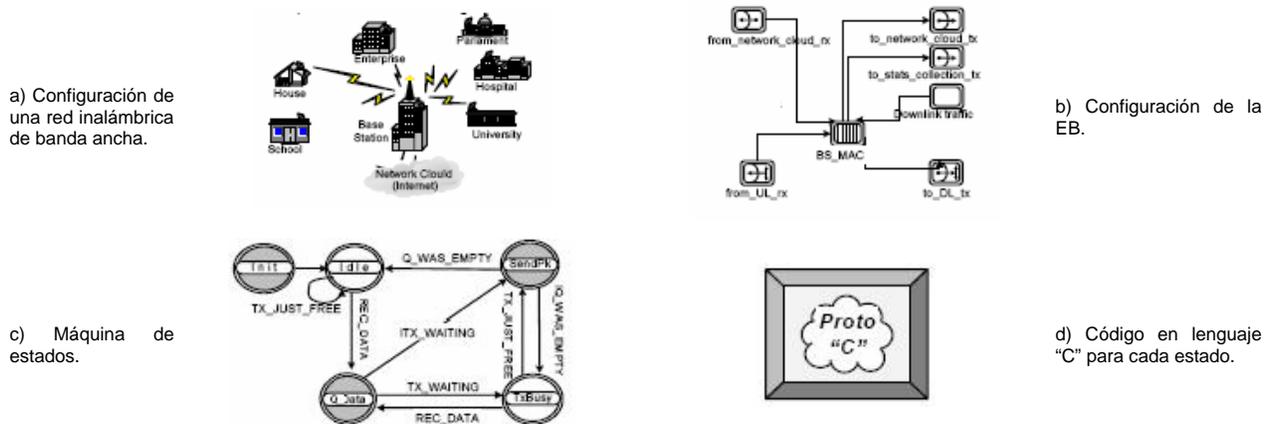


Figura 5.11 – Modelo jerárquico de simulación.

En el nivel más alto (Figura 5.11a) tenemos la topología de una red inalámbrica de banda ancha, en donde se muestran los componentes de la red físicamente. El siguiente nivel (Figura 5.11b) define la funcionalidad de una EB en términos de sus componentes, por ejemplo, fuentes de tráfico, su capa MAC, interfaces, etc.

En el siguiente nivel del modelo (Figura 5.11c) se define la operación de cada componente a través de una máquina de estados. Ahora, las acciones de un componente en particular de cada máquina de estados están escritos en código "Proto C".

Este enfoque nos permite modificar la operación del protocolo IEEE 802.16-2004 más fácilmente, así distintas optimizaciones y mejoras pueden ser probadas.

Los parámetros usados en las simulaciones en OPNET™ son detallados en la Tabla 5.6.

PARÁMETRO	VALOR
Slots de contención (C)	30, 51 y 72
Tamaño de un minislot en el UL-Frame	16 bytes
Duración del UL-Frame	2 ms / 450 minislots
Reed Solomon banda de guarda entre CodeWords.	6 bytes
Paridad Reed Solomon	10 bytes
Reed Solomon preámbulo	4 bytes
Codeword Reed Solomon (grants)	245 bytes
Encabezado de Capa MAC para 802.16	6 bytes
Minislots que demora la transmisión de un paquete desde una ES (x)	21 minislots
Número máximo de Ess activas (N)	100
Tipo de Tráfico	CBR
Ancho de banda del Canal	28.8 Mbps
Tamaño del paquete transmitido por las ESs	300 bytes
Distancia entre la ES más lejana y la EB.	5 Km
Valor inicial P _{SC} (Prob de contender con éxito)	0.01
Retardo inicial (w ₀) en ser asignado un grant a una ES	1 minislot

Tabla 5.6– Parámetros de simulación utilizados en OPNET™.

5.1.3.2 Comparación y análisis de eficiencia

Procederemos a realizar un análisis de la eficiencia del ancho de banda de nuestro sistema basado en el protocolo IEEE 802.16-2004 usando los modelos analíticos y de simulación descritos anteriormente.

Examinaremos el ancho de banda en función al número de ESs activas, las cuales usan la técnica *Best Effort* para obtener asignaciones de ancho de banda para poder transmitir sus datos.

Para extrapolar el modelo matemático en la simulación de OPNET™ asumiremos que las ESs activas envían flujos de datos de igual duración (paquetes de 300 bytes) en la capa de acceso al medio (MAC). Cuando dichos paquetes son codificados para ser transportados a la capa física del sistema basado en el protocolo IEEE 802.16, entonces cada ES usará 21 minislots para transmitir dicho paquete, usando una *Codeword* de 245 bytes, una paridad Reed Solomon de 10 bytes, un preámbulo de 4 bytes y una banda de guarda de 6 bytes.

La Figura 5.12 muestra el comportamiento de la eficiencia del ancho de banda del sistema en función del número de ESs activas y encoladas. Se incluyen los resultados analíticos y de simulación con el número de *slots* de contención (C) como parámetro.

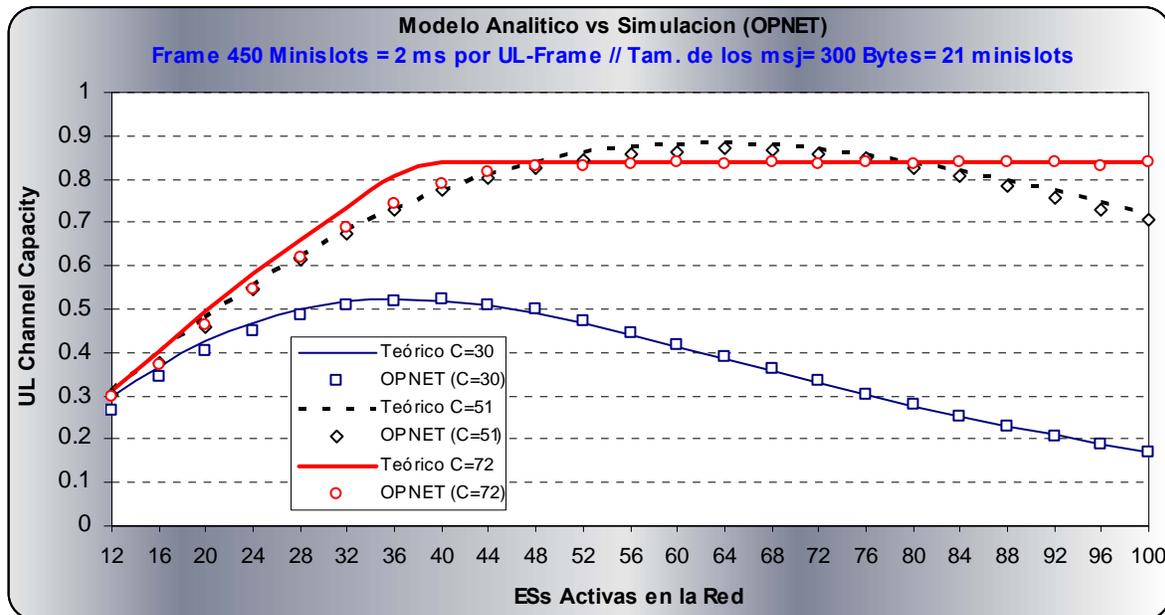


Figura 5.12 – Resultados de la eficiencia del ancho de banda de sistema usando el modelo teórico y la simulación.

Podemos observar que cuando el número de *slots* de contención por UL-MAP es pequeño ($C=30$), la máxima eficiencia del sistema es de aproximadamente 52% (para el modelo teórico y para la simulación), esto ocurre cuando tenemos aproximadamente 40 ESs activas. A partir de ese máximo, la eficiencia disminuye proporcionalmente al incremento de las ESs activas en la red hasta llegar a un punto mínimo del 18 %, cuando hay 100 ESs. Casualmente este valor corresponde a la eficiencia máxima del canal del primer protocolo de acceso a un medio compartido, **ALOHA**¹.

Al incrementar C a 51 *slots* se pone en evidencia un incremento considerable en la eficiencia del sistema con una mayor cantidad de usuarios, es decir, se alcanza una eficiencia máxima de 88 % cuando se tienen entre 60 y 70 ESs activas.

Aunque, nuevamente a partir de ese punto máximo, la eficiencia comienza a disminuir en función del incremento de ESs activas hasta llegar a un mínimo, nada despreciable de 72 %.

Un comportamiento muy interesante es observado cuando incrementamos todavía más el número de *slots* de contención, ahora $C=72$. Ahora la eficiencia alcanza un máximo de casi un 84 %, el cuál se mantiene constante inclusive para cuando el número de ESs aumenta.

¹ **ALOHA** – Protocolo de acceso a un medio inalámbrico compartido desarrollado en la Universidad de Hawaii en 1970, bajo la dirección de Norman Abramson.

Pero como se menciona en la secciones 5.1.1 y 5.1.2 se paga un cierto precio con un mayor incremento en el retardo. Dado que aunque la eficiencia del sistema es alta porque se asignan una gran cantidad de “grants” para que sean utilizados por la ESs, el tiempo el cual pasan las ESs (sin un grant asignado) encoladas es mayor.

Al analizar más a detalle la Figura 5.12 podemos darnos cuenta que al simular el nuevo modelo matemático sobre OPNET, los resultados son muy cercanos a los obtenidos teóricamente, mismos que se detallan en la sección 5.1.2.

5.2 COMPARACIÓN DE MODELOS (GPRS vs IEEE 802.16-2004)

La idea principal de esta comparación es la de mostrar evidencias que nos permitan demostrar que el uso de nuestro modelo (ecuaciones en diferencias) es más eficiente que el modelo usado por V.A Chitre y J.N Daigle sobre una red GPRS en 1999. Para ello, realizaremos un análisis cualitativo y cuantitativo en donde mostraremos los resultados arrojados por ambas propuestas técnicas, en donde evaluaremos tópicos como, el tiempo de ejecución, tiempo de análisis, manejo de información, exactitud, base matemática, dificultad en la implementación informática, entre otras cosas.

5.2.1 DESCRIBIENDO EL MODELO PARA GPRS

Los autores de [1] y [2] describen su modelo como una aproximación iterativa (de dos niveles) a través de la cual se calcula el tiempo de espera de un usuario para transmitir y la eficiencia de un sistema GPRS.

Inicialmente se asumen como conocidos los valores de espera promedio (para transmitir) y la probabilidad de contención exitosa de un usuario aleatorio. Basados en este par de premisas, realizan ciertas simplificaciones y suposiciones sobre el comportamiento del sistema, entonces se utilizan dichas suposiciones para iterar con la probabilidad de contención exitosa hasta que el parámetro converge con un tiempo de espera constante. Hacen algunas evaluaciones de factibilidad y se usan los resultados de esas revisiones para ajustar su aproximación de la media del tiempo de espera. Entonces repiten el proceso. De ese modo, en el nivel interior (primer nivel), iteran sobre la probabilidad de contención exitosa y en el nivel exterior (segundo nivel) iteran sobre el valor del tiempo promedio de espera (para transmitir) de un usuario aleatorio.

En los dos modelos se usan prácticamente los mismos parámetros para estimar la eficiencia y el retardo del sistema pero la diferencia más significativa es que el modelo usado para GPRS define una variable adicional, la probabilidad de captura para cada slot dado.

Los usuarios que están en modo de contención, contienden en cada bloque asignado para ese fin hasta que un usuario tiene éxito, entonces en cualquier periodo de contención dado, el número de usuarios conteniendo es una variable aleatoria, \bar{u} (el equivalente es nuestro modelo es \bar{b}). Cada uno de los usuarios conteniendo elige un slot de contención aleatoria e independientemente del resto de los demás del conjunto de R slots de contención. Definen \bar{n} como el número de usuarios eligiendo el slot r , en donde r es un slot elegido arbitrariamente. Entonces, $\bar{n}|\bar{u}=u$ es una variable de probabilidad binomial aleatoria con u y $1/R$ como parámetros, esto es:

$$P\{\bar{n} = n | \bar{u} = u\} = \binom{u}{n} \frac{1}{R} \left(1 - \frac{1}{R}\right)^{u-n}$$

En algún slot de contención dado, sólo un usuario puede contender con éxito y la probabilidad de que algún usuario contienda con éxito es especificado por el modelo de captura descrito.

5.2.2 DIFERENCIA ANALÍTICA EN LOS MODELOS

La probabilidad de captura es la principal diferencia entre los dos modelos, en la sección 4.2.2.1 definimos la probabilidad de éxito en la contención (la misma descrita en la ecuación arriba mencionada del modelo para GPRS) en un slot aleatorio (P_s) como la probabilidad de que en un número promedio de usuarios solicitando ancho de banda \bar{n} , sólo uno de ellos contienda en un slot dado.

Es decir, en nuestro modelo no se toma como variable la probabilidad de captura. Lo cual simplifica drásticamente el algoritmo, dado que determinar la probabilidad binomial (de captura en un slot dado) implica mayor desarrollo matemático y por ende más recursos informáticos y complejidad.

En nuestro modelo seguimos usando probabilidad binomial, con el número de ESs contendientes y el tamaño del bloque de contención como parámetros, pero con la diferencia de que sólo un usuario contienda en un slot dado ($n=1$). La expresión que nos representa dicha probabilidad (mostrada arriba) se ve simplificada considerablemente dado que, sólo nos interesa por el momento, conocer la probabilidad de contención de un grupo de contendientes en un solo *slot* de contención dado, quedando expresada como:

$$P_s = \binom{\bar{n}}{1} \frac{1}{C} \left(1 - \frac{1}{C}\right)^{\bar{n}-1} \quad (13)$$

Nota: No se considera la probabilidad de captura, lo cual nos permite hacer que el número de usuarios conteniendo en un slot sea igual a 1.

Con la ecuación anterior podemos determinar acertadamente el número esperado de ESs que transmiten un REQ exitoso, es decir, contienden exitosamente en el actual UL-Frame.

Para el modelo GPRS, propusieron un procedimiento iterativo para determinar el valor de P_{sc} . Sin embargo, y aunque el modelo matemático que usaremos es distinto en mucho sentidos, decidimos probar la solución propuesta en dichos trabajos en nuestro contexto con la red basada en IEEE 802.16-2004.

El hecho de que en nuestro modelo no se permitiera la captura de un paquete para calcular la probabilidad de tener una contención exitosa en un *slot* dado (P_s) es una diferencia fundamental, la cual nos permite dar un paso adicional en el desarrollo.

Aunque ambos modelos determinan la probabilidad de que un usuario contienda en un periodo de contención dado (P_c) y la probabilidad de contención exitosa en un UL-Frame dado (P_{sc}) son determinados con los mismos criterios, el hecho de no considerar la probabilidad de captura nos permitió determinar la probabilidad de contención exitosa en un slot determinado (P_s) de una forma totalmente distinta (usando nuestra ecuación 13).

En el modelo de V.A Chitre y J.N Daigle al igual que el propuesto en este documento se hace uso de probabilidades previamente obtenidas con el fin de determinar iterando el valor de la probabilidad de contención exitosa en un UL-Frame (P_{sc}). El procedimiento cuando se asume que el retardo de asignación de ancho de a una ES (\bar{w}) es conocido y suministrando un valor inicial estimado para la probabilidad de contención exitosa en un UL-Frame dado (P_{sc}).

Como se mostrará, el cambio en el criterio para determinar o no la probabilidad de captura no implica ninguna diferencia en los resultados después de los análisis teóricos y las comparaciones gráficas.

5.2.2 DIFERENCIA EXPERIMENTAL EN LOS MODELOS

Para esta fase del análisis se utilizaron los siguientes recursos y los parámetros de la tabla 5.7 con el fin de tener una comparación congruente de ambos enfoques.

PARÁMETRO	VALOR
Slots de contención (C)	30, 51 y 72
Tamaño de un minislot en el UL-Frame	16 bytes
Duración del UL-Frame	450 minislots
Número máximo de ESs activas (N)	100
Tipo de Tráfico	CBR (paquetes de 300 Bytes)
Ancho de banda del Canal	28.8 Mbps
Valor inicial P_{sc0} (Probabilidad de contender con éxito)	0.01
Retardo inicial (w_0) en ser asignado un grant a una ES	2 minislot

Tabla 5.7– Parámetros utilizados para obtener los resultados teóricos con un UL-Frame de 2 ms.

Los programas desarrollados en C++ que se utilizaron para generar la Figura 5.13 se incluyen en los apéndices B (ecuaciones en diferencias) y E (probabilidad binomial). Ambos programas fueron ejecutados en una computadora portátil con las siguientes características técnicas:

HP nx6120 con Procesador Intel Pentium M a 1.86 GHz con 1 Gb de RAM
Sistema operativo: Windows XP, Service Pack 2 Version 2002

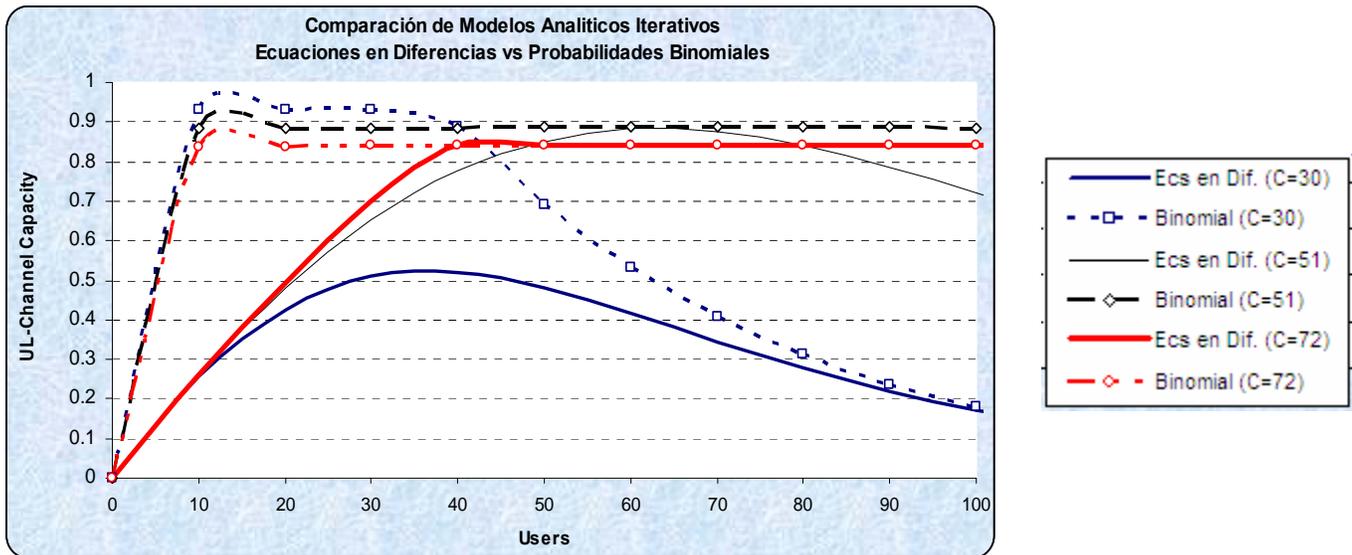


Figura 5.13 – Comparación de eficiencia entre modelos matemáticos iterativos (Ecs. en Diferencias vs. Probabilidad Binomial)

Se puede observar que en los tres casos el modelo usado para GPRS es considerablemente más eficiente cuando hay pocos usuarios en la red (menos de 20) pero cuando se van incrementando los usuarios y para los casos de 51 y 72 slots de contención, las gráficas van convergiendo aproximadamente al mismo valor (90%) para los dos modelos.

En la tabla 5.8 se incluyen los porcentajes de la eficiencia del canal de subida para ambos modelos.

Users	Modelo para IEEE 802.16-2004			Modelo para GPRS		
	C=30	C=51	C=72	C=30	C=51	C=72
0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
10	25.55%	26.33%	26.19%	93.20%	88.30%	83.90%
20	42.61%	48.24%	49.56%	93.20%	88.60%	83.90%
30	51.01%	65.40%	69.92%	93.30%	88.60%	84.00%
40	51.98%	77.67%	83.96%	89.00%	88.60%	84.00%
50	47.93%	85.13%	83.94%	69.20%	88.70%	84.00%
60	41.42%	88.21%	83.95%	53.30%	88.70%	84.00%
70	34.32%	87.58%	83.99%	40.80%	88.70%	84.00%
80	27.69%	84.11%	83.97%	31.20%	88.70%	84.00%
90	21.94%	78.72%	83.99%	23.80%	88.70%	84.00%
100	17.17%	72.19%	83.98%	18.20%	88.60%	84.00%

Tabla 5.8 – Eficiencia (expresada en porcentaje) del canal de subida para ambos modelos.

Además la falta de eficiencia de nuestro algoritmo con pocos usuarios en la red puede ser compensada fácilmente por medio de un manejo eficiente y dinámico de los slots asignados para la contención y del tamaño de los paquetes transmitidos por las ESs. Ver las gráficas en tercera dimensión descritas en la sección 5.1.1.2 y 5.1.1.3.

Ahora, por qué afirmamos que nuestro modelo es mejor, nos basamos en el principio de que el tener un sistema eficiente en todo momento no es lo más importante, lo que se persigue es que el sistema sea eficiente sólo cuando deba serlo, es decir, cuando existan muchos usuarios en la red.

En los casos de cuando hay pocos usuarios en la red inclusive hay slots de contención que ni siquiera pueden ser utilizados (por falta de usuarios).

Asímismo, es importante mencionar que la pérdida de un porcentaje de la eficiencia al tener pocos usuarios en la red, se compensa abiertamente con la simplicidad del nuevo algoritmo y de su facilidad en la implementación.

Sólo basta ver la diferencia en tamaño de los programas informáticos desarrollados para cada modelo, en el caso de GPRS el desarrollo contiene más de 200 líneas de código (apéndice E) mientras que el modelo con ecuaciones en diferencias usado para IEEE 802.16-2004 y descrito a detalle en este documento sólo se elaboró en 37 líneas de código (apéndice B). En la tabla 5.9 se incluyen algunos criterios adicionales de comparación entre ambos modelos.

	Modelo para IEEE 802.16-2004	Modelo para GPRS
Líneas de programación	36	Más de 200
Tiempo en converger	4 segundos máximo	7 minutos máximo
Iteraciones para converger	4	Más de 20
Máxima Eficiencia del UL	92%	93.20%
Eficiencia Promedio	83.97%	84%
Número de Niveles de Iteración	1	2

Tabla 5.9 – Cuadro comparativo entre los modelos.

Otra ventaja que fue explotada desde el punto de vista informático fue que para considerar la parte de la codificación (Reed Solomon) que usa el estándar IEEE 802.16 se creó un programa separado (apéndice C) con el fin de obtener el tamaño real de los paquetes transmitidos por las ESs. En el modelo para GPRS este proceso se obtenía dentro del mismo programa, lo cual lo hace mucho más complejo.



Además, otros factores muy importantes que deben ser tomados en cuenta al comparar ambos modelos son la complejidad, la eficiencia matemática así como la capacidad de cómputo requerido para ejecutarlos rápidamente.

El algoritmo para GPRS es un desarrollo con aproximación iterativa de dos niveles, el nuestro es sólo de un nivel, este hecho reduce la complejidad considerablemente y como podemos ver en la Figura 5.13 sin sacrificar eficiencia en momentos de alta demanda.

Recordemos que este tipo de algoritmos están siendo desarrollados con el fin de diferenciar la calidad técnica del estándar IEEE 802.16 aplicada a una red inalámbrica de banda ancha real, dado que no son parte del mismo. Lo cual podría implicar que una empresa proveedora de servicios inalámbricos de banda ancha opte por comprar equipo de cierto fabricante en función a los beneficios técnico-comerciales.

La idea de una empresa proveedora de servicios de red, es la de tener muchos usuarios usando la misma infraestructura en forma eficiente, por esta razón es despreciable que nuestro modelo no ofrezca un alto desempeño cuando hay pocos usuarios (menos de 20) simultáneamente.

CONCLUSIONES

Dando un puntual seguimiento a las hipótesis propuestas y después de haber analizado a fondo, como antecedente, los trabajos anteriores sobre la eficiencia de una red GSM (usando GPRS). El primer objetivo fue el utilizar un modelo matemático más sencillo que nos llevara a obtener resultados teóricos satisfactorios y a su vez, nos permitiera requerir menos recursos informáticos al aplicarlo a una red inalámbrica de banda ancha basada en el estándar IEEE 802.16-2004.

Contundentemente podemos afirmar que el tiempo que toma el nuevo modelo (con ecuaciones en diferencias) en converger en relación con el de la red GSM (considerando la probabilidad de captura y manejo de probabilidades binomiales), el cual es casi 8 veces menor.

En la mayoría de los casos, inclusive para el caso de cuando hay 100 usuarios simultáneos en la red, el nuevo modelo itera solo 4 veces antes de encontrar el componente del retardo total del ciclo de transmisión de un paquetes que corresponde al tiempo que el paquete pasa en contención (“C”). El encontrar el valor de este dato depende de parámetros probabilísticos, motivo por el cual, determinar este componente del retardo (contención) en el proceso de transmisión de un paquete implica la parte más complicada del algoritmo.

Una vez cumplido el primer objetivo, procedimos a adaptar el nuevo modelo matemático con las características de la capa de acceso al medio de una red inalámbrica de banda ancha basada en el estándar IEEE 802.16-2004 a realizar un sin número de pruebas teóricas variando todos los parámetros posibles.

Los parámetros que fueron variando y con los que obtuvimos los resultados teóricos fueron:

- *Slots* de contención.
- Tamaño de los paquetes.
- Número de usuarios activos en la red.
- Tamaño del UL-MAP.

La idea era contar con suficiente información para poder dar una opinión técnica sobre cuales son los mejores parámetros de configuración del sistema en un momento dado para construir los escenarios y así la red funcione de manera eficiente para aquellos usuarios que tengan una calidad de servicio (QoS) del tipo BE.

Para poder tener los escenarios teóricos, y como se menciona en el Capítulo 5, se desarrollaron 2 programas con la finalidad de:

- Implementar el nuevo modelo matemático. Este programa fue elaborado en Visual C++) y se incluye la versión con comentarios detallados en el Apéndice B.
- Obtener el número de bloques (*codewords* usando codificación Reed-Solomon) requeridos para transmitir paquetes de distinto tamaño. Este programa fue elaborado con Visual C++ y se incluye la versión con comentarios en el Apéndice C.

Con el fin de mostrar claramente los resultados teóricos se utilizaron gráficas (se incluyeron tridimensionales) en donde se muestra en una lámina todos los escenarios posibles.

De dichas gráficas podemos concluir que para que un sistema inalámbrico basado en el estándar IEEE 802.16-2004 tenga la eficiencia suficiente para poderlo hacer comercial se deben tener un escenario en el que la EB pueda manejar todos los parámetros analizados en forma dinámica.

Al diseñar el modelo del sistema inalámbrico de banda ancha se debe tener en cuenta la capacidad de adaptación dinámica a los requerimientos (ancho de banda, QoS, MAC, etc) de todos los usuarios en un momento dado.

Una de las hipótesis iniciales se basaba en la eficiencia de la red inalámbrica podría depender de la cantidad de *slots* de contención, pero en realidad la eficiencia depende también de otros parámetros. Por ejemplo, el tamaño de los paquetes codificados que se están transmitiendo.

Lo que sí es muy importante hacer notar es que la cantidad adecuada de *slots* de contención es directamente proporcional al tamaño de los paquetes que los usuarios transmiten en forma simultánea, recordemos que los paquetes son codificados antes de ser transmitidos. Entonces el número de *slots* de contención debe ser múltiplo (idealmente) del número de bloques codificados que se requieren para transmitir un paquete de cierto tamaño.

Es crítico para la eficiencia del sistema que se ajuste el número de *slots* de contención en función del tamaño de los paquetes en forma adecuada, ya que de no hacerlo entonces tendremos un impacto negativo, adicional al aumento del número de usuarios simultáneos.

Como hemos ya concluido, el manejo de los *slots* de contención es delicado para la eficiencia de la red, porque tan malo es tener en exceso como tener pocos.



Cuando tenemos pocos, debido a que los paquetes que transmiten los usuarios son grandes (más de 512 *bytes*) tenemos una eficiencia muy buena pero tenemos como consecuencia negativa un incremento en la demora en el procesamiento y transmisión de los paquetes. La EB sólo puede atender a pocos usuarios contendientes a la vez, generando más colisiones en la zona de contención.

En nuestro trabajo se presentó una extensión a lo anteriormente propuesto para el manejo de la subcapa de acceso al medio de una red inalámbrica basada en el estándar IEEE 802.16-2004. Los resultados obtenidos de citado modelo analítico han sido verificados vía simulación.

Encontramos además que el modelo matemático nuevo es lo suficiente preciso como para analizar las tendencias en la eficiencia del sistema, dado que en todos los casos se obtienen resultados numéricos dentro de los intervalos de confianza.

Con este modelo, se deja un escenario abierto a realizar investigaciones adicionales que tengan que ver con la inclusión de un algoritmo de *backoff* con el fin de optimizar las colisiones de los usuarios que contienen por ancho de banda, y así disminuir el retardo de espera para transmitir en periodos de alta utilización.

También queda pendiente la experimentación con ciertos tipos de tráfico específicos, adicional al tamaño de los paquetes, ciertas aplicaciones como HTTP o VoIP presentan comportamiento en el manejo de la transmisión de sus datos totalmente diferentes. Hay que generar escenarios en los que el sistema maneje simultáneamente distintas aplicaciones y generar las mismas gráficas de este análisis para acercarnos cada vez más a un verdadero punto de equilibrio al momento de diseñar una red con las características del estándar IEEE 802.16-2004.



REFERENCIAS BIBLIOGRÁFICAS

- [1] **Performance of IP-Based Services over GPRS**, V.A. Chitre and J.N. Daigle. Fellow, IEEE official transactions. June 2003.
- [2] **IP Traffic over GPRS – An Internet Service Oriented Analysis**, V.A. Chitre and J.N. Daigle. Center on Wireless Communications-University of Mississippi. 1999.
- [3] **Performance Analysis of QoS Scheduling in Broadband IEEE 802.16 Based Networks**, V. Rangel, J Gomez, J Ortiz. Wireless Network Research Lab, Telecommunications Department, UNAM (National University of Mexico).
- [4] **QoS Management in Broadband IEEE 802.16 Based Networks**, V. Rangel, J Gomez, J Ortiz. Wireless Network Research Lab in UNAM (National University of Mexico).
- [5] **Performance Analysis of Best Effort Support in Broadband IEEE 802.16 networks**. V. Rangel, J Gomez, J Chapa from Wireless Network Research Lab in UNAM (National University of Mexico), M. Lopez Guerrero from Department of Electrical Engineering, UAM-I, Mexico D.F and Raul Aquino from Telematic Faculty UCOL, Colima, Mexico.
- [6] **A Collision Resolution Paradigm for Random Access Channel Control in Multi-Access Wireless Networks**, M.C. Yuang, B.C. Lo, J.Y. Chen, Computer and Communications Research Laboratories, Department of Computer Science and Information Engineering National Chiao Tung University, Taiwan, IEEE official transactions, February 2002.
- [7] **Access Delay Analysis in Reservation Multiple Access Protocols for Broadband Local and Cellular Network**, Ahmed Doha, Hossam Hassanein, Telecommunications Research Lab in Queen's University. From Annual IEEE International Conference LCN 2004, March 2004.
- [8] **Performance of the IEEE 802.16 MAC Protocol with Varying Request Backoff Start**, B.N. Bhandari, Ratnam V. Raja Kumar, S.L. Maskara. Dept of Electronics and Electrical Communication Engineering, Indian Institute of Technology Kharagpur, West Bengal, INDIA, IEEE official transactions. June 2006.
- [9] **Uplink Performance of the IEEE802.16 Medium Access Control (MAC) Layer Protocol**, B.N. Bhandari, Ratnam V. Raja Kumar, S.L. Maskara. Dept of Electronics and Electrical Communication Engineering, Indian Institute of Technology Kharagpur, West Bengal, INDIA, IEEE official transactions. June 2005.
- [10] **Adaptive Contention Resolution for VoIP Services in the IEEE 802.16 Networks**, A. Sayenko, O. Alanen, T. Hämäläinen. Telecommunication laboratory, MIT department, University of Jyvaskyla, Finland. June 2007.
- [11] **Mac Performance of IEEE 802.16e**, A.E. Xhafa, S. Kangude, X. Lu. Texas Instruments Inc. 2005.
- [12] **Análisis del Desempeño de las Redes Celulares GSM-GPRS**, Patricia Sánchez Gómez, Tesis Profesional, Facultad de Ingeniería, UNAM. Septiembre de 2005.
- [13] **WiMAX Technology for Broadband Wireless Access** by Loufi Nuaymi, ENST Bretagne, France, (Hardcover – 2007). Wiley and Sons Editorial USA.



- [14] **IEEE 802.16-2001, "IEEE Standard for Local and Metropolitan Area Networks** - Part 16: Air Interface for Fixed Broadband Wireless Access Systems, April 2002. IEEE 3 Park Avenue, New York, NY 10016-5997, USA.
- [15] **IEEE 802.16-2004, "IEEE Standard for Local and Metropolitan Area Networks** - Part 16: Air Interface for Fixed Broadband Wireless Access Systems, October 2004. IEEE 3 Park Avenue, New York, NY 10016-5997, USA
- [16] **IEEE 802.16e-2005, "IEEE Standard for Local and Metropolitan Area Networks** - Part 16: Air Interface for Fixed Broadband Wireless Access Systems, 2005. IEEE 3 Park Avenue, New York, NY 10016-5997, USA.
- [17] **Mobile WiMAX: A Performance and Comparative Summary** by Doug Gray, September 2006. WiMAX Forum
- [18] **Mobile WiMAX – Part 1: A technical Overview and Performance Evaluation**, 2006. WiMAX Forum.
- [19] **Design of HandOff procedures for broadband access IEEE 802.16 Based Networks**, V. Rangel, J.E. Cota, J Gomez, J. Reyes. Wireless Network Research Lab, Telecommunications Department, UNAM (National University of Mexico).
- [20] **Notas de Temas Especiales de Telecomunicaciones: Modelado de Redes Inalámbricas de Banda Ancha: WiMAX IEEE802.16** – Capítulo 6: Codificación y Modulación. Autor: V. Rangel. Semestre 2007-II IIMAS, UNAM. México, D.F.
- [21] **Probability and Statistics for Engineering and the Sciences** by Jay L. Devore (Hardcover – Feb 2007). Thomson Learning Books. ISBN 0-534-37281-3. United States.
- [22] **Analysis of Packet Network Having Contention-Based Reservation with aplicaciones to GPRS**, J.N. Daigle. Fellow IEEE and Marcos Nascimento Magalhanes. IEEE/ACM Transactions on Networking, Vol. 11, No. 4, August 2003.
- [23] **WirelessMAN: Inside the IEEE 802.16 Standard for Wireless Metropolitan Area Networks** by Carl Eklund, Roger B. Marks, Subbu Ponnuswamy, and Kenneth L. Stanwood (Paperback - May 15, 2006)
- ** **Mobile and PSTN Communication Services: Competition or Complementary ? - OCDE/GD (95)96** - Head of Publications Service, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France. (Paris 1995)
- ** **Mobile Multiple Play: New Service Pricing and Policy Implications. DSTI/ICCP/TISP(2006)1/FINAL**
By Yoshikazu Okamoto, OECD's Directorate for Science, Technology and Industry. OCDE 15-Jan-2007
- *** **Managing the Risk of Mobile Banking Technologies - OCDE** - This report was commissioned by FinMark Trust. Bankable Frontier Associates.
- *** **The Implications of WiMAX for Competition and Regulation –OCDE DSTI/ICCP/TISP(2005)4/FINAL.**
By Taylor Reynolds of the OECD's Directorate for Science, Technology and Industry. OCDE 02-Mar-2006.



APÉNDICE A

GLOSARIO DE TÉRMINOS

3G	Tercera generación de los sistemas telefónicos celulares. Ejemplos: UMTS y cdma2000.
AAA	Authentication Authorisation and Accounting. Protocolo que realiza las funciones de autenticación, autorización y asignación de una cuenta. Frecuentemente asociado a un servidor AAA.
AAS	Adaptive Antenna Systems. Sistema de Antenas Adaptivo. La subcapa MAC del estándar IEEE 802.16 tiene incluidas funcionalidades que le permiten usar AAS.
ACK	ACKnowledge or ACKnowledgement. Aviso de recibido, mensaje de control utilizado en el mecanismo de ARQ para indicar que un paquete fue recibido exitosamente.
AMC	Adaptive Modulation and Coding. Modulación adaptiva y codificación.
ARCEP	(French telecommunications regulation authority) Autorité de Régulation des Communications Electroniques et des Postes. Siglas de la autoridad regulatoria de las Telecomunicaciones en Francia.
ARQ	Automatic Repeat reQuest. Solicitud Automática de Retransmisión. Es un protocolo de transmisión de capa 2.
ATM	Asynchronous Transfer Mode. Modo de Transmisión Asíncrona.
BE	Best Effort. BE es una de las cinco clases de servicio soportadas por el estándar IEEE 802.16, es usada para los datos con más baja prioridad de transmisión como el correo electrónico.
BER	Bit Error Rate. Tasa de bit en error.
BPSK	Binary Phase Shift Keying. Es un tipo de modulación digital binaria.
BR	Bandwidth Request. Requerimiento de ancho de banda.
BS	Base Station. Estación Base de una red inalámbrica.
BSID	Base Station IDentifier. Identificador de la Estación Base.
BW	Bandwidth. Ancho de banda.
BWA	Broadband Wireless Access. Red inalámbrica de banda ancha.
CBR	Constant Bit Rate. Tipo de servicio usado para transmitir datos en una red a una tasa constante.
CDMA	Code Division Multiple Access. Múltiple Acceso por División de Código.
CPE	Consumer Premises Equipment. Equipo instalado del lado del usuario.
CPS	Common Part Sublayer. Subcapa común de la capa MAC del estándar IEEE 802.16.
CRC	Cyclic Redundancy Check. Código de redundancia cíclica.
CS	Convergence Sublayer. Subcapa de Convergencia, es la parte superior de la capa MAC de una red basada en el estándar IEEE 802.16. Esta capa realiza el servicio específico de transformar o mapear los datos de redes externas antes de su transmisión a través del medio inalámbrico.
DAMA	Demand Assigned Multiple Access. Múltiple Acceso por Demanda Asignada.
DC	Direct Current. Corriente Directa.
DCD	Downlink Channel Descriptor. Descriptor del canal de bajada. Es un mensaje de administración que describe la composición del canal de bajada en una red IEEE 802.16.
DIUC	Downlink Interval Usage Code. Código que identifica a cada perfil de ráfaga de transmisión, el cual acompaña cada ráfaga de datos de bajada.
DL	DownLink. Canal de bajada. Canal para que la EB transmita a la ES.



DL-MAP	DownLink MAP. Mapa del canal bajada. Es una mensaje de administración, transmitido al inicio de capa DL-Frame indicando su contenido.
EIRP	Equivalent Isotropic Radiated Power. Potencia isotrópica radiada equivalente.
ertPS	Extended real-time Polling Service. Nueva clase de servicio introducida en la enmienda IEEE 802.16e y que se encuentra entre las clase rtPS y UGS.
FCH	Frame Control Header. Encabezado de control de trama
FDD	Frequency Division Duplexing.
FEC	Forward Error Correction. Corrección de errores adelantada es una forma de codificación en un canal de transmisión.
FSN	Fragment Sequence Number. Número de secuencia de fragmento.
FTP	File Transfer Protocol. Protocolo para transferencia de archivos.
GMH	Generic MAC Header. Encabezado MAC genérico.
GSM	Global System for Mobile communication. Ejemplo de la segunda generación de un sistema telefónico celular.
HARQ	Hybrid Automatic Repeat reQuest. Evolución del protocolo ARQ.
HCS	Header Check Sequence. Verificador de secuencia de encabezado.
H-FDD	Half-duplex FDD.
HO	HandOver. Término utilizado para expresar el cambio de una ES de una célula a otra sin perder la conexión activa.
HTTP	Hyper Text Transfer Protocol.
HUMAN	High-speed Unlicensed Metropolitan Area Network. Free license 802.16 specification.
IE	Information Element. Element of a MAC message. For example, a DL-MAP_IE describes one burst profile.
IEEE	Institute of Electrical and Electronics Engineers. Instituto de Ingenieros Eléctrico-Electrónicos.
IETF	Internet Engineering Task Force. Fuerza de Tarea de Ingeniería de Internet.
IMS	IP Multimedia Subsystem. Subsistema Multimedia IP.
IP	Internet Protocol.
ISM	Industrial, Scientific and Medical. Appellation of the unlicensed 2.4 GHz frequency bandwidth.
IUC	Interval Usage Code. See DIUC and UIUC. Código de uso de Intervalo.
LDPC	Low-Density Parity Check code. Codificación de canal usando un código de verificación de paridad de baja densidad.
LEN	LENGth. Longitud en bytes de un MAC PDU. Incluye el encabezado MAC y el CRC, si es que está presente.
LoS	Line-of-Sight. Línea de vista.
LTE	Long-Term Evolution. Evolución del sistema de telefonía 3G.
MAC	Media Access Control Layer. Parte de la capa 2 de modelo de referencia OSI.
MAC	Message Authentication Code. Código de autenticación de mensaje.
MAN	Metropolitan Area Network. Red de área metropolitana, el IEEE 802.16 en un sistema inalámbrico MAN.
MBS	Multicast and Broadcast Services.
MCS	Modulation and Coding Scheme.
MDHO	Macro Diversity HandOver. En un estado en donde la comunicación móvil se efectúa con más de una estación base.
MIB	Management Information Base. The BS and SS managed nodes collect and store the managed objects in an 802.16 MIB format
MIMO	Multiple-Input Multiple-Output. Tipo de antena de última generación.
MIP	Mobile IP.



MMDS	Multichannel Multipoint Distribution Service. Servicio de distribución multicanal multipunto.
MPDU	MAC PDU. Unidad de datos de protocolo para la capa MAC.
ES	Estación Subscriptora.
EBS	Estación Base.
MSDU	MAC SDU. Unidad de datos de Servicio para la capa MAC.
NACK	Non-ACKnowledge or Non-ACKnowledgement. Mensaje de control usado en el mecanismo de control de flujo ARQ.
NAP	Network Access Provider (cf. Architecture WiMAX). Business entity that provides a WiMAX radio access infrastructure to one or more WiMAX Network Services.
NLoS	Non-Line-of-Sight. Sin línea de vista.
nrtPS	Non-real-time Polling Services. Uno de los cinco tipos de QoS de WiMAX.
NSP	Network Service Provider (cf. Architecture WiMAX). Business entity that provides IP connectivity and WiMAX services to WiMAX subscribers.
NWG	NetWork Group. Grupo del foro WiMAX a cargo de la creación de las especificaciones de arquitectura de alto nivel.
OEM	Original Equipment Manufacturer.
OFDM	Orthogonal Frequency Division Multiplexing. Transmission technique. Su principio es transmitir la información en muchas frecuencias ortogonales portadoras.
OFDMA	Orthogonal Frequency Division Multiple Access. Es la variante de OFDM para un escenario de acceso múltiple.
OPUSC	Optional PUSC.
PAPR	Peak-to-Average Power Ratio. En una transmisión en OFDM, el PAPR es el valor pico de portadoras transmitidas para transmitir una señal promedio.
PBR	PiggyBack Request. Campo para el Subencabezado de Grant Management en el UL-MAP con la requisición de ancho de banda por parte de la ES.
PCM	Pulse Coded Modulation. Sistema clásico de transmisión de señales telefónicas. Variantes son T1 y E1.
PDU	Protocol Data Unit.
PHS	Payload Header Suppression. Proceso opcional en la subcapa CS.
PHSF	Payload Header Suppression Field.
PHSI	Payload Header Suppression Index.
PHSM	Payload Header Suppression Mask.
PHSS	Payload Header Suppression Size.
PHSV	Payload Header Suppression Valid.
PHY	PHYSical layer.
PICS	Protocol Implementation Conformance Specification document.
PM	Poll-Me bit. Cuando las ES están activas en conexiones de tipo UGS pueden habilitar el bit PM (contenido en el Grant Management subheader) dentro de un paquete MAC de la conexión UGS para indicarle a la EB que puede polear para asignar ancho de banda para usuarios con conexiones distintas a UGS.
PMP	Point-to-MultiPoint. Topología básica de WiMAX.
PN	Pseudo-Noise sequence.
PRBS	Pseudo-Random Binary Sequence. Se utiliza en el bloque de randomización.
PS	Physical Slot. Función de la capa física. Usada como una unidad de atribución de recursos.
PUSC	Partial Usage of SubChannels. Modo de permutación OFDMA.
QAM	Quadrature Amplitude Modulation. Tipo clásico de modulación.
QoS	Quality of Service. Calidad de Servicio.
QPSK	Quadrature Phase Shift Keying. Tipo clásico de modulación.



RF	Radio Frequency.
RFC	Request For Comment. Document IETF.
RRA	Radio Resource Agent.
RRC	Radio Resource Controller.
RRM	Radio Resource Management.
RS	Reed–Solomon code. Tipo de codificación de canal.
RSSI	Received Signal Strength Indicator. Indicador del nivel de potencia en la señal recibida.
RTG	Receive/transmit Transition Gap. El RTG es un intervalo entre la ráfaga del canal de subida y la subsecuente ráfaga del canal de bajada en una transmisión TDD.
RTP	Real-Time Protocol.
rtPS	Real-time Polling Services. Uno de los cinco tipos de QoS de WiMAX.
SAP	Service Access Point.
SBC	SS Basic Capability.
SC	Single Carrier.
SDU	Service Data Unit.
SFA	Service Flow Authorization.
SFID	Service Flow Identifier. Es un servicio de la capa MAC.
SFM	Service Flow Management.
SI	Slip Indicator. Campo del Grant Management subheader que indica la combinación de asignaciones de UL en relación a la cola de espera para dicho UL.
SISO	Single-Input Single-Output. Caso específico para la tecnología MIMO.
SLA	Service Level Agreements.
SM	Spatial Multiplexing. Familia de algoritmos de la tecnología MIMO.
SN	Sequence Number. Número del bloque transmitido usado en el mecanismo de control de flujo ARQ.
SNMP	Simple Network Management Protocol. IETF Network Management Reference model protocol.
SNR	Signal-to-Noise Ratio.
CINR	Carrier-to-Interference-and-Noise Ratio.
SOFDMA	Scalable OFDMA.
SPID	SubPacket Identifier. Used in the HARQ process.
SS	Subscriber Station.
STBC	Space Time Block Coding. Variante de la tecnología MIMO.
STC	Space Time Coding. Variante de la tecnología MIMO.
TCP	Transmission Control Protocol.
TCS	Transmission Convergence Sublayer. Mecanismo opcional de la capa física.
TDD	Time Division Duplexing.
TDM	Time Division Multiplexing.
TFTP	Trivial File Transfer Protocol.
TLV	Type/Length/Value.
TO	Transmission Opportunity.
TTG	Tx/Rx Transition Gap. El TTG es un intervalo entre la ráfaga del canal de bajada y la subsecuente ráfaga del canal de subida en una transmisión TDD.
UDP	User Datagram Protocol.
UDR	Usage Data Records.



UCD	Uplink Channel Descriptor. Mensaje de administración de la capa MAC que describe el canal de subida.
UGS	Unsolicited Grant Services. Uno de los cinco tipos de QoS de WiMAX.
UIUC	Uplink Interval Usage Code. Identificador del perfil de ráfaga que se acompaña en cada ráfaga del canal de subida.
UL	UpLink. Canal de subida.
UL-MAP	UpLink MAP. Es el mensaje de administración MAC que indica el contenido del UPLINK Frame.
UTC	Universal Coordinated Time.
V-NSP	Visited NSP.
VoIP	Voice over IP.
WiFi	Wireless Fidelity. IEEE 802.11 certification consortium.
WiMAX	Worldwide Interoperability for Microwave Access Forum. El foro WiMAX provee la certificación de conformidad, compatibilidad e interoperabilidad de los productos basados en el estándar IEEE 802.16. WiMAX es también el nombre común de la tecnología basada en el estándar IEEE 802.16.
WLL	Wireless Local Loop. Sistema de comunicación inalámbrica.

APÉNDICE B

PROGRAMA PARA DETERMINAR LA EFICIENCIA (THROUGHPUT), MODELO CON ECUACIONES EN DIFERENCIAS

Programa realizado en Visual C++ y ligado a un archivo de EXCEL, las referencias técnicas del mismo se encuentran detalladamente descritas en la sección 5.2.2 de este documento.

```
Sub Analytical()  
Dim C, M, I, W, X, j, N, k As Integer  
Dim Psc_1, Psc_n, Psc_n_plus_1, q, beta As Double  
' Constantes  
M = Cells(2, 2)      'Minislots por cada UL-Frame  
W = Cells(3, 2)      'Retardo inicial (minislots) en recibir un grant  
X = Cells(5, 2)      'Bloques RLC (minislots) que conforman una Tx exitosa de un usuario  
Psc_1 = Cells(8, 2)  'Probabilidad inicial de contener exitosamente en un UL-Frame  
I = 1.5 * M          'Retardo inicial (minislots) desde q un pkt llega de una capa superior hasta que la SS genera un grant RE  
For j = 1 To 3      'Inicio del ciclo recursivo  
    C = Cells(1, 1 + j)  
    q = (1 - 1 / C) 'Variable para simplificar Ec (19) y (20)  
    For k = 1 To 25  
        N = 0 + k * 4  
        I = M * 1.5      'Estimación del retardo inicial Ec (4)  
        beta = ((I + C + X + W) / M) - 1      'Variable para simplificar Ec (17)  
        Do 'W  
            W = W + 1  
            Psc_n_plus_1 = Psc_1  
            Do 'Psc  
                Psc_n = Psc_n_plus_1  
                Psc_n_plus_1 = q ^ (N / (beta * Psc_n + 1) - 1) 'Ec (20)  
                e = Abs(Psc_n_plus_1 - Psc_n)  
            Loop Until e < 0.000001      'Umbral de convergencia del algoritmo  
            t = I + (C + M * (1 - Psc_n_plus_1) / Psc_n_plus_1) + W + X      'Ec (16)  
            d = t * (M - C) / (M * N)      ' Ec (21) slots de transmisión disponibles por ciclo  
        Loop Until d > X  
        thr = N * X / t      ' Ec (2) obteniendo la eficiencia  
        If j = 1 Then  
            Cells(10 + k, 1) = N  
        End If  
        Cells(10 + k, 1 + j) = thr  
        Cells(40 + k, 2 + j) = t  
        W = 0  
    Next  
Next  
End Sub
```

APÉNDICE C

PROGRAMA PARA DETERMINAR EL TAMAÑO DE UN BLOQUE EN LA RED IEEE 802.16

El programa realizado en Visual C++. Se incluyen comentarios con el fin de explicar la manera en la que se determina el número de bloques (incluyendo encabezados de las capas MAC y Física) en función al tamaño del paquete de datos que las estaciones transmitirán. El número de bloques es un parámetro muy importante con el fin de utilizar de manera óptima la zona de contención en el UL-Frame.

```
//Programa para determinar la duración de un bloque para transmitir un paquete de tamaño fijo
//Se consideran encabezados de la capa Física (se utiliza codificación Reed Solomon) y de la capa MAC (por defecto)

#include <stdio.h>
#include <stdlib.h>
#include <math.h>

//Definición de función que genera el número de RLCs
int computeblocksSerial(int);

//Programa principal
main()
{
    double Tdata_bloks,Eg;
    int j,U;

    printf(" U      Bloques RLC      Bloques Tot  \n\n");

    //printf("G\n\n");
    U = 10;
    Eg = computeblocksSerial(200); // Bloques RLC de 16 bytes con encabezados y codifacion para paquetes de XX Bytes de
    Tdata_bloks = Eg*U;           // Cada usuario puede transmir Eg bloques RLC (de XX bytes)

    //printf("%3.0f\t%3.0f\n",U,Eg);
    printf("%3.0i\t%3.0f\t%3.0f\n",U,Eg,Tdata_bloks);
    return 0;
}

/*****
int computeblocksSerial(int no_of_bits)
{
    int BWA_MAC_header = 6; // Header 802.16
    int FEC1 = 6; // Reed Solomon entre Code Words
    int FEC2 = 10; // Longitud de Code Words
    int preamble = 4; // Preámbulo de 802.16
    int guard_band = 6; // Banda de Guarda entre ráfagas

    int remain, length_BWA_MAC, length_Eth;
    int RLC_pkts_1, RLC_pkts_2, Total_RLC_pkts, quotient, no_PDU_pkts;

    length_Eth = no_of_bits;

    length_BWA_MAC=length_Eth + BWA_MAC_header;

    no_PDU_pkts = length_BWA_MAC/245;
    remain = (length_BWA_MAC - no_PDU_pkts*245);
    RLC_pkts_1= no_PDU_pkts*245;
    RLC_pkts_2 = remain + FEC2;
    --
    Total_RLC_pkts = FEC2* no_PDU_pkts + RLC_pkts_1 + RLC_pkts_2 + preamble + guard_band;

    quotient = Total_RLC_pkts/16;

    if ((Total_RLC_pkts - quotient*16) != 0)
        quotient += 1;

    Total_RLC_pkts = quotient;

    return (Total_RLC_pkts);
}
*****/
```

APÉNDICE D

WHITE PAPER: PERFORMANCE ANALYSIS OF BEST EFFORT SUPPORT IN BROADBAND IEEE 802.16 NETWORKS

PERFORMANCE ANALYSIS OF BEST EFFORT SUPPORT IN BROADBAND IEEE 802.16 NETWORKS

V. Rangel, J. Gomez, J. Chapa
Department of Telecommunications
UNAM, CP 04510, Mexico City
{victor, javierg, javierc}@fi-b.unam.mx

M. Lopez-Guerrero
Department of Electrical Engineering
UAM-I, CP 09340, Mexico City
milo@xanum.uam.mx

Raul Aquino
School of Telematics
UCOL, Colima, Mexico
aquinor@ucol.mx

ABSTRACT

This paper presents a performance analysis of best effort support over the MAC protocol proposed in the IEEE 802.16 standard. We focus on the uplink channel since it is the critical element for efficient delivery of services on demand to individual users (i.e., Best Effort service). We describe a modeling approach and derive the resulting analytical model. Such a model allows us to efficiently compute network throughput. We test this model by means of computer simulations which are also reported in this paper. Simulation results and theoretical computations showed a close agreement, thus validating the model.

Index Terms— Channel modeling; Wireless Broadband Networks; performance analysis; WiMAX.

1. INTRODUCTION

Broadband Wireless Access (BWA) has become the best option to meet residential and small business demand for high speed Internet access and multimedia services. However, the wide-scale adoption of BWA systems will be determined by its ability to overcome cost and performance barriers. If BWA can meet these challenges, it could easily be the next revolution in wireless systems after WLAN.

The first version of the IEEE 802.16 protocol was completed in October 2001. This version was called IEEE 802.16-2001 [1] and it defines the air interface and medium access control (MAC) protocol for a wireless metropolitan area network (WMAN). It was intended for high-bandwidth wireless voice and data for residential and enterprise use. At the beginning of its development, this protocol was oriented to fixed wireless users with line of sight (LOS) in the 11-66 GHz spectrum range.

In 2004 the aim of the 802.16 protocol was changed to support residential access and non line of sight (NLOS). The second version is called the IEEE 802.16-2004 standard [2] and it supports two operating modes of the MAC layer: 1) point to multipoint (PMP), where traffic occurs only between the Base Station (BS) and Subscriber Stations (SS), and 2) Mesh topology, where traffic can be routed through other SSs and can occur directly between SSs. In addition, IEEE 802.16-2004 also includes OFDM modulation and

supports 256 carries, which reduces considerably multipath fading effects.

Recently, the IEEE 802.16 Task Force released a new version that enables mobility in SSs. The IEEE 802-16e [3] standard promises mobility support up to speeds of 120 km/h and an asymmetrical link structure. It will enable a SS to have a handheld form factor well suited for PDAs, phones and laptops.

Due to all potential options in BWA standards, at the present time there is a significant interest by the research community, manufactures and operators about what type of performance can really be expected from IEEE 802.16-2004 [2] and IEEE 802.16e [3] certified equipment in the near future. Recent studies found in the literature focus on system performance regarding support of Quality of Service (QoS) [4-7] and cross-layer issues [8-10] for the IEEE 802.16 MAC protocol. However, only relatively few studies [11-13] explicitly approach the analysis of the Best Effort (BE) service class in comparison to other service classes (UGS, rtPS and nrtPS).

Performance of the Best Effort service is severely affected as the network gets congested due to collisions of bandwidth requests transmitted on contention access. The aim of this paper is to provide an analytical model to assess the performance of BE in IEEE 802.16 broadband wireless networks.

The proposed model takes into account that in IEEE 802.16 MAC based systems there is no packet capture. This is due to the fact that subscriber stations always transmit with the minimum transmit power to achieve a specific Bit Error Rate (BER) at the base station. This policy avoids co-channel interference with other cells. Consequently, when two or more stations transmit a bandwidth request in the same contention minislots, a collision occurs and it is not possible to recover any of the bandwidth requests. Our model is based on a few assumptions. The system is considered to be operating in steady state and it is considered that all active SSs are continuously backlogged. It also makes use of a constant value to represent the size of the backoff window of the Exponential Backoff Algorithm (EBA) as we will explain in Section 4.

The paper is structured as follows. Section 2 presents an overview of the relevant parts of the IEEE 802.16 protocol. Section 3 derives the theoretical model for the performance of the uplink channel of the protocol. In Section 4, we present the performance analysis of the IEEE 802.16 protocol comparing analytical and simulation results. We conclude the paper in Section 5.

2. IEEE 802.16 MAC LAYER OVERVIEW

The IEEE 802.16 MAC layer makes use of two access techniques, frequency division duplex (FDD) and time division duplex (TDD). Bandwidth requests and data from SSs to the BS are carried in an uplink (UL) frame. Transmissions from the BS to SSs are carried by a downlink (DL) frame. Upon receiving a bandwidth request, the BS handles bandwidth allocation by assigning uplink grants based on requests from the SSs. A typical signaling frame for TDD includes an UL-frame (see Fig. 1a) and a DL-frame (see Fig. 1b) using a single channel frequency as illustrated in Fig. 1c. In FDD these frames are transmitted at the same time using different frequencies as illustrated in Fig. 1d (see [2] for further details on the frame structure).

On power-up the SSs carry out the Initialization and Registration setup as indicated in [2]. After this setup is completed, a SS can create one or more connections over which their data will be transmitted to and from the BS. SSs request transmission opportunities using the contention access period (or contention block) of the current UL-frame. The BS collects these requests and determines the number of minislots (grant size) that each SS will be allowed to transmit in the next UL-frame, using a UL_MAP subframe, as shown in Fig. 1b. The UL-MAP frame contains Information Elements (IE) which describe the use of the UL-frame (maintenance, contention and reservation access). The UL-MAP is broadcasted in the DL channel by the BS in each DL-Frame. After receiving the UL-MAP, a SS will transmit data in the predefined reserved minislots indicated in the IE. These reserved minislots are transmission opportunities assigned by a scheduling algorithm using a service class, such as UGS (Unsolicited Grant Service) for CBR (Constant Bit Rate) traffic, rtPS (real-time Polling Service) for VBR (Variable Bit Rate), nrtPS (non real-time Polling Service) for non real-time bursty traffic, and BE (Best Effort) for traffic such as Internet, email and all other non real time traffic. Regarding the UL-frame structure depicted in Fig. 1a, in this work we assume that only contention and reservation minislots are considered in the UL-frame, since maintenance minislots are included scarcely.

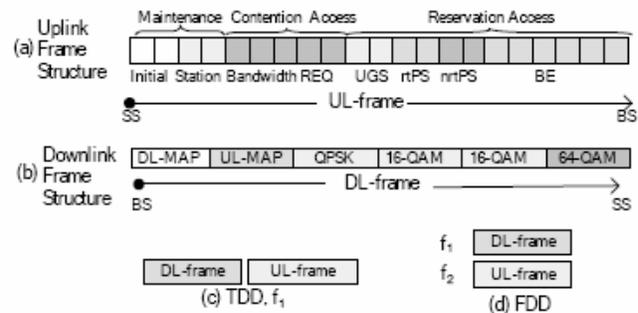


Fig. 1. Frame structure for TDD and FDD access.

3. ANALYTICAL MODEL

Since this work focuses on the BE service, let us start by describing the sequence of actions that take place when a SS makes use of it for data transfer. When a SS is active, (let us say SS_x), it forms a continuous loop with the sequence of actions depicted in Fig. 2. Upon a packet arrival from an upper layer protocol, SS_x waits for the next UL-MAP containing a contention period. Then, SS_x randomly chooses one of the available contention minislots and transmits a bandwidth request (REQ) indicating the packet length. In case some other SS (let us say SS_y) selects the same contention minislot, a collision occurs and the subscriber stations (SS_x and SS_y) will receive neither a grant nor an acknowledgement (ACK) in the following UL-MAP. Thus, SS_x retransmits its REQ until it is successfully transmitted. Upon a successful reception of a REQ from SS_x , the BS converts the packet size to a number of minislots that should be reserved in subsequent UL-frames. In case the REQ from SS_x does not fit in the next UL-frame, the BS sends a null grant to SS_x in order to acknowledge the REQ.

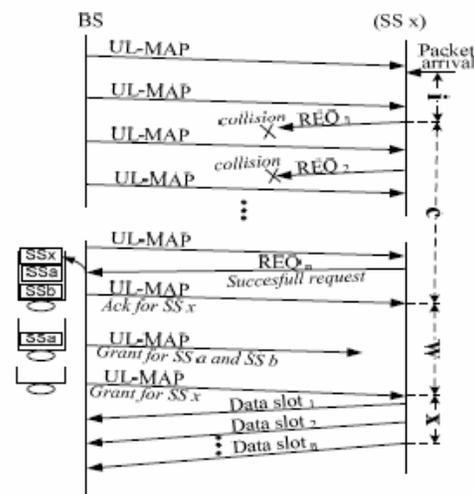


Fig. 2. Access delay components.

Modeling of such sequence of events can be carried out by breaking down a single packet transmission in its delay components. Let us denote by i the time delay, measured in minislots, from the time a packet arrives from upper layers until the beginning of the contention block where SS_x transmits its REQ. Let c represent the total time in minislots spent during contention, which starts with the beginning of the contention block where SS_x transmits the first REQ, until SS_x receives an ACK (i.e., a null grant in the IEEE 802.16 protocol). Let us define by w the time in minislots that the scheduler takes in order to grant the REQ of SS_x . It is measured from the ACK reception to the grant reception at SS_x . At the BS, the scheduler serves REQs using a FIFO discipline. If the REQ from SS_x cannot be granted in the next UL-frame, it waits until previous REQs from other SSs are served. Note that in case a REQ can be immediately served, instead of returning an ACK, the BS returns a grant indicating the number of minislots that were reserved. In this case, reception of such a grant also signals the end of the contention time c and therefore w is zero.

Finally, let us denote by x the delay component that represents the actual number of minislots spent during packet transmission from SS_x . Fig. 2 depicts the relation between the events described above and the delay components of the model. Therefore, the time to transmit a single packet (t) can be directly obtained by adding all delay components, i.e., $t = i + c + w + x$. By taking expectation on both sides of this equation we obtain

$$\bar{t} = \bar{i} + \bar{c} + \bar{w} + \bar{x}. \quad (1)$$

From these parameters, the normalized network throughput (γ) of the UL channel can be computed as the fraction of minislots actually spent in data transmission per station, scaled to the total number of SSs (N), thus,

$$\gamma = N\bar{x} / \bar{t}. \quad (2)$$

The aim of this paper is to provide a method for the computation of this metric. The notation used in this derivation is presented in Table I.

Regarding Eq. (2) it is necessary to compute \bar{x} and \bar{t} . The mean delay for packet transmission in minislots (i.e., \bar{x}) can be directly obtained from the packet size distribution, and taking into consideration that minislots can be of size 2, 4, 8, 16, 32 or 64 bytes. Thus,

$$\bar{x} = \left[\frac{\bar{m} + MAC_OV + PHY_OV}{slot_size} \right] \quad (3)$$

where \bar{m} is the mean packet size at the LLC layer, MAC_OV is the MAC overhead of the IEEE 802.16 protocol (the default value for MAC_OV is 6 bytes). The overhead at the physical layer, PHY_OV , depends on the coding techniques involved (see [2] Chapter 8, for further information on coding rates and modulation techniques).

Computation of \bar{t} implies computation of other delay components as described below. Estimation of the mean

initial delay \bar{i} has to consider the following three delay factors. First, a packet arrival from an upper layer protocol may occur anywhere within the current UL-frame, thus the SS must wait until the next available contention block in order to transmit a REQ. The mean value for this delay can be approximated by $M/2$ minislots, where M is the total length of the UL-frame. Second, the SS has to wait a complete frame of size M minislots in order to transmit a REQ in the available contention block of the next UL-frame. Third, once the REQ is received, the BS may grant this REQ anywhere in the following UL-frame, which can also be approximated by $M/2$ minislots. Thus the initial delay, can be approximated as

$$\bar{i} \cong 2M. \quad (4)$$

Whereas computation of \bar{x} and \bar{i} is straightforward, computation of \bar{c} and \bar{w} is far more complicated. This procedure is described in the following sections.

3.1 Computation of the contention delay

For simplicity we assume that the number of failed contentions that a SS needs to succeed follows a geometric distribution. Thus, given the probability of a successful contention in a contention block (P_{SC}), the mean number of minislots used for contention \bar{c} would be given by

$$\bar{c} = C + M \frac{(1 - P_{SC})}{P_{SC}} \quad (5)$$

where C is the size of the contention block and M is the total length of a UL-frame (both measured in minislots).

Let \bar{n} and \bar{s} be the expected number of the total and successful contenders in a contention block, respectively. It is clear that the probability P_{SC} can be estimated as

TABLE I. IEEE 802.16 MAC MODEL NOTATION

Symbol	Definition
C	Size of the contention block in a UL-frame
R	Size of the reservation period in a UL-frame
M	Total size of a UL-frame, $M=C+R$
i	Initial delay
c	Contention delay to transmit a REQ successfully
w	Time to grant a REQ for a SS
x	Time spent in a packet transmission
t	Total time in a transmission cycle, $t = i + c + w + x$
$\bar{i}, \bar{c}, \bar{w}, \bar{x}, \bar{t}$	Expectations of $i, c, w, x,$ and t , respectively
P_{SC}	Probability of a successful contention in a UL-frame
n	Number of SSs contending in the current UL-frame
s	Number of SSs that transmitted a REQ successfully in the current UL-frame
\bar{n}, \bar{s}	Expectations of n and s , respectively
\bar{b}	Expected number of contention periods per cycle, \bar{t}
N	Maximum number of active SSs
P_C	Probability that a SS contends in a contention period
P_S	Probability of a successful contention in a minislot
\bar{d}	Available transmission minislots for a SS per cycle
γ	System throughput

$$P_{SC} = \frac{\bar{s}}{N}. \quad (6)$$

Let us now turn our attention to the estimation of \bar{n} and \bar{s} . Let us denote by P_C the probability that a SS decides to contend in a contention block. Therefore, the expected number of contenders in a contention block is given by

$$\bar{n} = NP_C \quad (7)$$

and the number of successful contenders per contention block is

$$\bar{s} = (NP_C)P_{SC}. \quad (8)$$

Note that there are \bar{i} minislots per transmission cycle and one contention block every M minislots. Therefore, the mean number of contention blocks (\bar{b}) per transmission cycle is given by

$$\bar{b} = \frac{\bar{i}}{M}. \quad (9)$$

Let us assume that the system is operating in steady state. Under this assumption, each active SS gets a chance to transmit every \bar{i} minislots and parameter \bar{s} can be estimated dividing the total number of SS among the mean number of contention blocks per transmission cycle as follows

$$\bar{s} = \frac{N}{\bar{b}}. \quad (10)$$

Substituting (9) in (10) we obtain

$$\bar{s} = \frac{MN}{\bar{i}}. \quad (11)$$

From (7) and (11) in (6) we can obtain

$$P_C = \frac{M}{P_{SC}\bar{i}}. \quad (12)$$

Let us define the probability of successful contention in an arbitrary slot P_S as the probability that, from the mean number of contenders in a contention block \bar{n} , only *one* contends in a minislot (packet capture is not possible),

$$P_S = \binom{\bar{n}}{1} \frac{1}{C} \left(1 - \frac{1}{C}\right)^{\bar{n}-1}. \quad (13)$$

Therefore, parameter \bar{s} can be computed as

$$\bar{s} = CP_S. \quad (14)$$

Finally, from (14) and (7) P_{SC} can be computed as

$$P_{SC} = \frac{CP_S}{NP_C}. \quad (15)$$

Note that computation of \bar{c} in (5) implies computation of P_{SC} which in turn depends on computation of P_C and P_S as indicated by (15). At this point it is fair to mention that the authors in [14] made use of these three probabilities in their analysis of IP-based services over GPRS networks. Although P_C and P_{SC} are computed here in the same way, the fact that we do not consider packet capture allows us to compute P_S in a fundamentally different way. Computation of this probability as defined by (13) is not equivalent to the

method presented in [14]. This difference allows us to go one step further as it will be shown.

The authors in [14] make use of the previously derived probabilities in order to iteratively compute P_{SC} . Although our model and theirs differ in several ways, we also made use of this procedure with good results. The procedure starts by assuming that \bar{w} is known and providing an initial estimate for P_{SC} . Then, from (5), (1), (12) and (13) we compute P_C and P_S . These values are used in (15) in order to obtain the following value for P_{SC} . With this value the process can be repeated starting from (5). This loop can be ended when the difference between two consecutive values for P_{SC} is below a certain threshold. The final value for P_{SC} allows us to compute the value for the contention delay \bar{c} using (5). Further adjustments on \bar{w} would be needed if the assumed value for \bar{w} does not satisfy all system conditions.

In what follows we derive the equation that represents the procedure just described. Let us identify with subindex n the n -th iteration. Thus, (5) and (1) become $\bar{c}_n = C + M(1 - P_{SC_n}/P_{SC_n})$ and $\bar{i}_n = \bar{i} + \bar{c}_n + \bar{w} + \bar{x}$, respectively. From these equations, we have

$$\bar{i}_n = \bar{i} + C + M \left(\frac{1 - P_{SC_n}}{P_{SC_n}} \right) + \bar{w} + \bar{x}. \quad (16)$$

From (12) we know that $P_{C_n} = M / (P_{SC_n} \bar{i}_n)$ which combined with (16) yields,

$$P_{C_n} = \frac{1}{\beta P_{SC_n} + 1} \quad (17)$$

where $\beta = ((\bar{i} + C + \bar{w} + \bar{x})/M) - 1$.

From (7) and (17) the estimated mean number of users at the n -th iteration is

$$\bar{n}_n = \frac{1}{\beta P_{SC_n} + 1} N. \quad (18)$$

From (13) and (18) the estimated value of the probability of successful contention in an arbitrary slot at the n -th iteration is

$$P_{S_n} = \frac{N}{\beta P_{SC_n} + 1} \left(\frac{1}{C} \right) \left(1 - \frac{1}{C} \right)^{\frac{N}{\beta P_{SC_n} + 1} - 1}. \quad (19)$$

Finally, from (17), (19) and (15), after some algebra, the difference equation that can be used to obtain the value of P_{SC} is

$$P_{SC_{n+1}} = \left(1 - \frac{1}{C} \right)^{\frac{N}{\beta P_{SC_n} + 1}}. \quad (20)$$

In summary, given \bar{w} and an initial estimate for P_{SC} , we iterate (20) until a consistent value for P_{SC} is obtained. With this value the contention delay \bar{c} can be computed using (5). The algorithm implied by (20) is very easy to implement and it is much simpler than the one presented in [14].

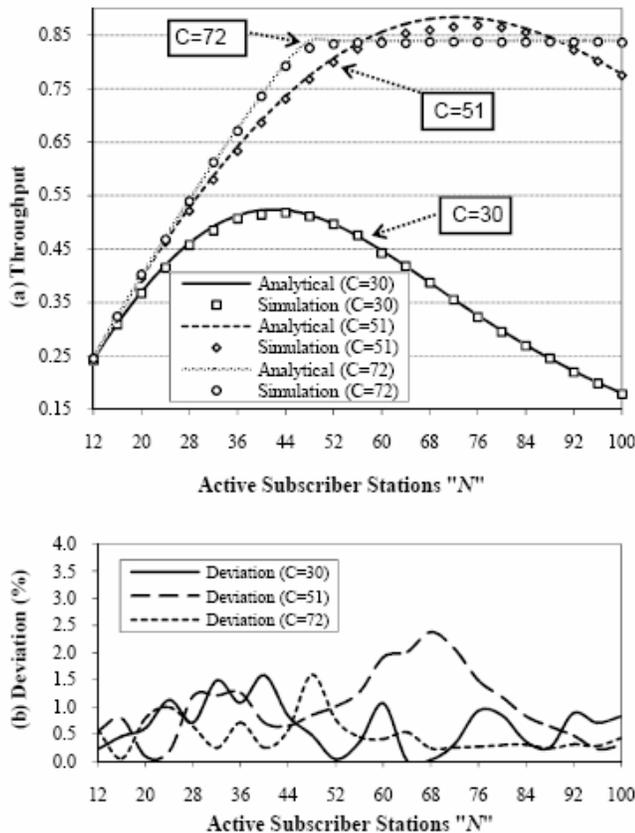


Fig.4. Performance analysis of IEEE 802.16 based Network.

each UL-frame. With 40 SSS, the average number of grants served per UL-frame was of 11.6 of a total of $(M-C)/\bar{x} = 21$ grants. The rest of the UL-frame was wasted, since just a few REQ could arrive to the BS due to collisions.

By increasing C to 51 minislots, the average number of grants served by a UL-frame increased considerably to 17, which nearly achieves 90% of system throughput when there are between 60 and 70 SSS. However, when there is a large number of SSS in the network, (i.e., more than 80 SSS) the system throughput cannot be maintained due to the large number of collisions reported.

With $C = 72$ minislots, the system throughput can be sustained even on very large networks, however the maximum system throughput is just over 80% of channel capacity.

For the three values of C shown in Fig. 4b, we observed that simulation results were in good agreement with analytical results. The maximum deviation from our simulation model was less than 3%.

5. CONCLUSION

In this paper we presented an analytical model for the MAC layer of IEEE 802.16 based networks under best effort traffic. The results obtained from the analytical model were

verified by simulation. We found that the analytical model is sufficiently accurate to study the trends in system performance and, in almost all cases, gives numerical results within the confidence intervals derived from the simulation results. The maximum deviation between analytical and simulation results was less than 3%, with a mean deviation of 0.8%. In future work will consider performance modeling at the physical layer, different traffic scenarios and we will also take into account the dynamic behavior of the Exponential Backoff Algorithm.

ACKNOWLEDGMENTS

This work was supported by DGAPA, UNAM under Grant PAPIIT IN104907 and PAPIME PE103807. Additional funding came from *Área de Investigación en Redes y Telecomunicaciones* (UAM-Iztapalapa, Mexico).

REFERENCES

- [1] IEEE 802.16-2001, "IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed Broadband Wireless Access Systems," April 2002.
- [2] IEEE 802.16-2004, "IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed Broadband Wireless Access Systems," October 2004.
- [3] IEEE 802.16e, "IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed BWA Systems", Amendment for PHY and MAC for Combined Fixed and Mobile Operation in Licensed Bands, December 2005.
- [4] C. Cicconetti, A. Erta, L. Lensini, E. Mingozzi, "Performance Evaluation of the IEEE 802.16 MAC for QoS Support," *IEEE Trans. on Mobile Computing*, vol. 6, no.1, pp. 26-38, Jan. 2007.
- [5] R. Iyengar, P. Iyer, B. Sikdar, "Delay Analysis of 802.16 based Last Mile Wireless Networks," in *2005 Proc. GLOBECOM Conf.*, St. Louis.
- [6] Y. Yu, H. Chao, "QoS Aware Link Adaptation in IEEE 802.16 Broadband Wireless Access Systems," in *2007 Proc. PIMRC Conf.*
- [7] D.Hao, S.Zou, B. Adamou, "A QoS Guaranteed Cross-Layer Scheduling Algorithm in Wireless Networks," in *2003 Proc. ICWMC Conf.*
- [8] J. Yan, G. Kuo, "Cross-Layer Design of Optimal Contention Period for IEEE 802.16 BWA Systems," in *Proc. IEEE ICC 2006*, Istanbul.
- [9] Y. Chang, F. Chien, J. Kuo, "Cross-Layer QoS Analysis of Opportunistic OFDM-TDMA and OFDMA Networks," *IEEE Journal of Selected Areas in Communications*, vol. 25, no. 4, May 2007.
- [10] Y. Mai, C. Yang, Y. Lin, "Cross-Layer QoS Framework in the IEEE 802.16 Networks," in *Proc. ICACT 2007*, Gangwon-Do, Korea
- [11] K. Gakhar M. Achir, A. Gravey, "Dynamic Resource Reservation in IEEE 802.16 Broadband Wireless Networks", in *Proc. IEEE IWQoS*, 2006. New Haven, USA.
- [12] B. Bhandari, R. Kumar, S. Maskara, "Performance of the IEEE 802.16 MAC Protocol with Varying Request Backoff Start," in *Proc. IEEE TENCON 2006*, Hong Kong.
- [13] A. Vinel, Y. Zhang, M. Lott, A. Tiurlikov, "Performance Analysis of the Random Access in IEEE 802.16," in *Proc. IEEE PMRC 2005*, Berlin.
- [14] V. A. Chitre and J. N. Daigle, "Performance of IP-Based Services over GPRS," *IEEE Trans. on Computers*, vol. 52, No. 6, pp. 727-741, June 2003.
- [15] V. Rangel, J. Gomez y J. Ortiz, "Performance Analysis of QoS Scheduling in Broadband IEEE 802.16 Based Networks," in *Proc. OPNETWORK Technology Conf.*, Washington 2006.

ACEPTACIÓN DE WHITE PAPER EN LA IEEE CANADIAN CONFERENCE ON ELECTRICAL AND COMPUTER ENGINEERING 2009

Property	Change Add	Value								
Conference and track		IEEE Canadian Conference on Electrical and Computer Engineering 2009 - <i>Communications and Networking</i>								
Authors	+	Name	ID	Flag	Affiliation	Email	Country	Email	Move authors	Delete
		Victor Rangel	138247		National University of Mexico	victor@fi-b.unam.mx	Mexico		▼	
		Javier Gomez	92732		National University of Mexico	javiergo@gmail.com	Mexico		▲ ▼	
		Javier Chapa	355744		Universidad Nacional Autónoma de México	jchapa@uxmcc2.iimas.unam.mx	Mexico		▲ ▼	
		Miguel Lopez-Guerrero	138549		Universidad Autónoma Metropolitana	milo@xanum.uam.mx	Mexico		▲ ▼	
		Raul Santos	86331		University of Colima	aquinator@ucol.mx	Mexico		▲	
Presenter		?								
Registration code		4085								
Title		<i>Performance Analysis of Best Effort Support in Broadband IEEE 802.16 Networks</i>								
Abstract		This paper presents a performance analysis of best effort support over the MAC protocol proposed in the IEEE 802.16 standard. We focus on the uplink channel since it is the critical element for efficient delivery of services on demand to individual users (i.e., Best Effort service). We describe a modeling approach and derive the resulting analytical model. Such a model allows us to efficiently compute network throughput. We test this model by means of computer simulations which are also reported in this paper. Simulation results and theoretical computations showed a close agreement, thus validating the model.								
Keywords		Channel modeling; Wireless Broadband Networks; performance analysis; WIMAX								
Status		accepted								
Printing problems										
Copyright form		Mar 5, 2009 5:29 PM								
Visa letter										
Final manuscript		uploaded Sun, March 8, 2009 17:51 EDT; can upload until Fri, March 13, 2009 23:59 EDT								
Files		Document (show)	Size	Changed	MD5	Check format				
			401,201	Mar 8, 2009 17:51	7ace97c22c18c458af786684d3783c0a					

You are an author of this paper.

1 Review

Review 1

Reviewer D

Novelty	Quality of Presentation	Quality of Research	Recommendation
Slightly novel (2)	Could be improved (3)	Poor quality (2)	Weak Reject (2)

Contributions (What is the major contribution(s) of this paper? Comment on the degree of novelty, creativity and technical depth in the paper.)

The paper provides the performance analysis for the throughput of the best effort in IEEE 802.16 networks. However, some main mechanisms are not considered in the performance analysis, and the novelty is not enough.



Weaknesses (What are the most important reasons NOT to accept the paper? [Please be brief and precise.])

Some assumptions are too simple. Some main mechanisms in IEEE 802.16 networks are not considered in performance analysis. In IEEE 802.16 networks, the truncated binary exponential backoff mechanism is used to resolve the contention during the resource request, which is not considered in the analysis.

Detailed comments (Provide detailed comments that will be helpful to the TPC for assessing the paper, as well as useful feedback to the authors. Please attempt to avoid generic comments.)

1. The truncated binary exponential backoff mechanism which is used in IEEE 802.16 networks to do the Contention resolution should be considered in the analysis model. <p> 2. What is the traffic model used in the analysis and simulation? The total delay t give in Eq. (1) should be related to the traffic model. <p> 3. The paper assumes that "All SSs used Best Effort technique for grant service on page 5. Does this mean that all traffic is Best Effort traffic? In practical situation, due to the low priority of BE traffic, the number of available timeslots for BE traffic should be a variable, not a constant.

Award Quality? (Would this paper be of award quality?)

No

2 TPC Reviews/Summaries

Review 1

Reviewer B

Novelty	Quality of Presentation	Quality of Research	Recommendation
Somewhat novel (3)	Well written (4)	High quality (4)	Weak Accept (4)

Review 2

Reviewer C

Novelty	Quality of Presentation	Quality of Research	Recommendation
Somewhat novel (3)	Well written (4)	High quality (4)	Strong Accept (5)

Contributions (What is the major contribution(s) of this paper? Comment on the degree of novelty, creativity and technical depth in the paper.)

The paper is well written and technically sound. The level of creativity and novelty is good.

Strengths (What are the major reasons to accept the paper? [Please be brief and precise.])

The paper addresses a relevant practical problem with sound methodology that is well explained. It provides interesting, but not surprising results.

Weaknesses (What are the most important reasons NOT to accept the paper? [Please be brief and precise.])

If space permits the authors should describe how their work and results compare with those in references [11-13].

Award Quality? (Would this paper be of award quality?)

No

APÉNDICE E

PROGRAMA PARA DETERMINAR LA EFICIENCIA (THROUGHPUT), MODELO CON PROBABILIDAD BINOMIAL (USADO EN GPRS)

```
/*Modelo para analizar teóricamente el Performance de una red basada en IEEE 802.16 con Probabilidad Binomial
#include <stdio.h>
#include <stdlib.h>
#include <math.h>

#define zero 0
#define TRUE 1
#define FALSE 0
#define numberMoments 10

//Definición de funciones
double findBinomialSuccess(int,int, double*, int);
double findProbSuccessContend(int, int, double, double*, int);

int computeblocksSerial(int);
void findBinomialProbs(int, double, double*);

//Programa principal
main()
{
double Pc_i, U, Eg, Es_i, Pc, Pu, Ec, Ebz, Ebz_i, Ed, Ez, Es, Ew, Ea;
double proportionAvailMsg, delta, successesPerCycle;
double Tdata_bloks, tThr,tThrProp, userShare, captureProb[6], upperCaptureLimit;

int CS, j, moreWait,iWait,cont_slots,M;

printf("U CS Pc E[z] E[c] E[w] E[s] Ebz Ebz_i Thr G 1/L\n");
upperCaptureLimit = 1;

for (j=0; j<= upperCaptureLimit; j++)
captureProb[j] = 0.0;
captureProb[1] = 1.0;

for (cont_slots = 30; cont_slots <=72; cont_slots+=21)
{
for (U=10; U <= 150; U+= 10) {
CS = cont_slots; // Variable con el Número de Slots de Contención por Bloque
Pc = 0.001; // Probabilidad de contención existosa en Condiciones iniciales en el UL-Fram
M = 450; // Minislots (2 ms) en el UL-Frame
Eg = computeblocksSerial(300); // Bloques RLC de 16 bytes con encabezados y codificacion para paquetes de 30
Tdata_bloks = Eg*U; // Cada usuario puede transmitir 21 bloques RLC (300 bytes)
delta = 1.0; // Retardo por Wi inicial, Time to schedule a REQ from a SS (minislots)- Inicia después de que la SS
Ew=1; // Suposición inicial w=1 Paso 0 // rec un ACK y culmina CON un grant con son slots reserv al REQ de dicha S
Ea= 675; // De Ecuación (3) Ea= 1.5*M Paso 1b. Ret en slots por estado incial, inicia cuando un pkt es recibic

moreWait = TRUE;

while (moreWait) {
while(delta > 10E-10) { // Se hace el pasol del algoritmo hasta que el margen de error sea aceptable
Pc_i = Pc; // Primero se toma el Valor Inicial de la Prob de éxito en la contención
Ec = CS + (1-Pc_i)*M/Pc_i; // Paso 1a.Se determina el nuevo delay por contención con la prob de éxito inicia
Ez = Ea + Ec + Ew + Eg; // E[z]=E[a]+E[c]+E[w]+E[g] //Paso 1c. Se determina t, el tiempo del ciclo de Tx
|
Ebz = Ez/(M); //Ecuación (7)
Es = U/Ebz; //E[s]=U*M/E[z] Ecuación(8)
Pu = Es/(Pc_i*U); //Pc=E[s]/E[u], donde E[u]=Pu*U
if (Pu > 1)
Pu = 1;
Pc = findProbSuccessContend(U, CS, Pu, captureProb,upperCaptureLimit);
delta = fabs(Pc-Pc_i);
}

//Ec = (R + (1-Pc_i)*M/Pc_i)*30; // E[c]=1+[1-Pc]*M/Pc
Ez = Ec + Ew + Eg; // E[z] = E[c] + E[w] + E[g]
//Ebz_i = 1.0 + (Ez-(double) R/32)/M;
//Ebz_i = 1.0 + (Ez-1.0)/M;
}
}
}
}
```



```

Ebz = Ez/(M); //Ebz_i = 0.25 + (Ez-0.25)/M;
//Ebz = Ebz_i;
//Ebz_real = Ez*R/(M*32);
Es = U/Ebz; //Ebz = 1 + (1-Pc_i)/Pc_i; //NO sirve para R = 16, T = .92, R=32 T = .97
Ed = (Ez - Ebz*CS); //E[s]=U*M/E[z]
//((E[z]-E[b])*8
successesPerCycle = Ebz*Es; //proportionAvailMsg = 30*Ed/Ez; //(E[z]-E[b])*8/E[z]
//E[b]*E[s]
if (Ed/U <= Eg)
    userShare = Ed/U;
else
    userShare = Eg;
    tThrProp= U*userShare/(Ez);
    delta = 1.0;
    if(userShare < Eg)
    {
        if (Eg - userShare >= 1)
            Ew+= Eg - userShare;
        else
            Ew+= 1;
            iWait = iWait+1;
    }
    else moreWait = FALSE;
printf("%3.0f %2d %3.3f %6.2f %7.2f %8.3f %8.3f %8.3f %8.3f %8.3f %3.0f %8.6f\n", U, CS, Pc, Ez, Ec, Ew, Es, Ebz,
}
}
return 0;
}

double findProbSuccessContend(int r_max, int slots, double p, double pWin[6], int UCL)
{
double tempSuccessProb, contenderProbs[161], ESslotSuccess[161];
int i, j; //contenderProbs = dvector(0,r_max);
for (j=1; j<= r_max; j++) //ESslotSuccess = dvector(1,r_max);
    contenderProbs[j] = 0;
for (j=1; j<= r_max; j++)
    ESslotSuccess[j] = 0;
findBinomialProbs(r_max, p, contenderProbs);
ESslotSuccess[1] = pWin[1]/(double) slots;
for(i=2;i<=r_max;i++)
    ESslotSuccess[i] = findBinomialSuccess(i, slots, pWin, UCL);
tempSuccessProb = 0.0;
for(i=1;i<=r_max;i++)
    tempSuccessProb += ESslotSuccess[i]*contenderProbs[i];
//free_dvector(ESslotSuccess,1,r_max);
for (j=1; j<= r_max; j++)
    ESslotSuccess[j] = 0;
//free_dvector(contenderProbs,0,r_max);
for (j=0; j<= r_max; j++)
    contenderProbs[j] = 0;
tempSuccessProb /= ((double) r_max * p); //Pc = R*Ps/(U*Pu)
//tempSuccessProb = ((double) r_max * p);
return tempSuccessProb;
}

/*****
void findBinomialProbs(int n, double pSuccess, double *probVector)
{
int i, j; double pChoose, qChoose;
pChoose = pSuccess;
qChoose = 1-pChoose;
probVector[0] = 1 -pChoose;
probVector[1] = pChoose;
for(i=2; i <= n; i++) {
    probVector[i] = probVector[i-1] * pChoose;
    for(j = i-1 ; j >0; j--)
        probVector[j] = probVector[j]*qChoose + probVector[j-1]*pChoose;
    probVector[0] *= qChoose;
}
}

/*****
double findBinomialSuccess(int r, int n, double successProb[161], int upperCaptureLimit)
{
double pChoose, qChoose, probMassChoose[161], temp, upperLimit;
int i, j;
pChoose = (double) 1/ n;
qChoose = 1-pChoose;
//probMassChoose = dvector(0,r);
for (j=0; j<= r; j++)
    probMassChoose[j] = 0;
findBinomialProbs(r, pChoose, probMassChoose);
temp = 0.0;
//upperLimit = min(upperCaptureLimit,r);
if (upperCaptureLimit <= r)
    upperLimit = upperCaptureLimit;
else
    upperLimit = r;
for(i=1;i<=upperLimit; i++) |
    temp += probMassChoose[i]*successProb[i];
//free_dvector(probMassChoose, 0, r);
}

```



```
for (j=0; j<= r; j++)
    probMassChoose[j] = 0;
temp *= (double) n;
return temp;
}

/*****/
int computeblocksSerial(int no_of_bits)
{
int BWA_MAC_header = 6;    // Header 802.16
int FEC1 = 6;             // Reed Solomon entre Code Words
}
int FEC1 = 6;             // Reed Solomon entre Code Words
int FEC2 = 10;           // Longitud de Code Words
int preamble = 4;       // Preámbulo de 802.16
int guard_band = 6;     // Banda de Guarda entre ráfagas

int remain, length_BWA_MAC, length_Eth;
int RLC_pkts_1, RLC_pkts_2, Total_RLC_pkts, quotient, no_PDU_pkts;

length_Eth = no_of_bits;

length_BWA_MAC=length_Eth + BWA_MAC_header;

no_PDU_pkts = length_BWA_MAC/245;
remain = (length_BWA_MAC - no_PDU_pkts*245);
RLC_pkts_1= no_PDU_pkts*245;
RLC_pkts_2 = remain + FEC2;

Total_RLC_pkts = FEC2* no_PDU_pkts + RLC_pkts_1 + RLC_pkts_2 + preamble + guard_band;

quotient = Total_RLC_pkts/16;

if ((Total_RLC_pkts - quotient*16) != 0)
    quotient += 1;

Total_RLC_pkts = quotient;

return (Total_RLC_pkts);
}
```