



**CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS  
DEL INSTITUTO POLITÉCNICO NACIONAL**

**DEPARTAMENTO DE INGENIERÍA ELÉCTRICA  
SECCIÓN DE COMUNICACIONES**

**“Análisis de teletráfico de sistemas celulares con usuarios  
VoIP persistentes y calendarizador de máximo SIR basado  
en estadísticas de orden”**

**TESIS**

Que presenta

**M. en C. LUIS ALBERTO VASQUEZ TOLEDO**

Para obtener el grado de  
**DOCTOR EN CIENCIAS**

En la Especialidad de  
**INGENIERÍA ELÉCTRICA**

Director de Tesis:  
**Dr. DOMINGO LARA RODRÍGUEZ**

**CIUDAD DE MÉXICO**

**AGOSTO 2018**

---

# Agradecimientos

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico brindado durante la realización de este proyecto.

Agradezco al Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV), unidad Zacatenco, esta institución de enorme calidad, que me brindó todo el apoyo durante mi estancia.

Al departamento de Ingeniería Eléctrica, sección de Comunicaciones, por haberme dado la oportunidad de aprender, conocer, formarme y por la infraestructura brindada.

Quiero agradecer a mi asesor de tesis, el Dr. Domingo Lara Rodríguez, por contribuir a mi formación aceptando dirigir esta tesis, por los conocimientos invaluable que me brindo para llevar a cabo esta investigación, y sobretodo por su confianza, paciencia y amistad brindada.

Agradezco a los miembros del jurado, Dr. Hildeberto Jardón Aguilar, Dr. Raúl García Ruiz, Dr. Héctor Hugo García Compeán y Dr. Ricardo Marcelín Jiménez, por sus valiosos comentarios que hicieron al trabajo final y por el tiempo que dedicaron para revisarlo.

A mis padres, Julián Vasquez Linares y Lucila Toledo Valdivieso, quienes me dieron la vida, que me han formado y educado. A quienes nunca podré pagar con las riquezas más grandes de mundo. A ellos los seres universalmente más queridos sinceramente Gracias.

A mis hermanos, Florentina Vasquez Toledo, por ser mi ejemplo de fortaleza y el apoyo brindado, Julián Vasquez Toledo, por ser mi compañero y ejemplo de vida y estar siempre a mi disposición.

A mi esposa Berenice Borja Benítez, por su amor, su apoyo incondicional, sus

palabras de aliento y ánimo que me brinda día con día para alcanzar nuevas metas, tanto profesionales como personales.

A mi Tía Florentina Vasquez Linares, por ser la guía y mi apoyo incondicional.

A mi primo, Roberto Vazquez Aquino, por todas las vivencias y por siempre animarme.

A mi cuñada Georgina Ruvalcaba Ortiz por el apoyo brindado a mi familia.

---

# Dedicatoria

A mi adorado hijo Matías Vasquez Borja, posiblemente en este momento no entiendas mis palabras, pero para cuando seas capaz, quiero que te des cuenta de lo que significas para mí. Eres la razón de que me levante cada día para esforzarme por el presente y el mañana, eres mi principal motivación.



---

# Índice general

<b>Agradecimientos</b>	<b>1</b>
<b>Dedicatoria</b>	<b>3</b>
<b>Resumen</b>	<b>11</b>
<b>Abstract</b>	<b>13</b>
<b>Introducción</b>	<b>15</b>
<b>Objetivos</b>	<b>17</b>
<b>1. Generalidades</b>	<b>19</b>
1.1. Acceso múltiple por división de frecuencia ortogonal (OFDMA) . . . .	20
1.2. Perspectiva sobre la próxima generación de sistemas de comunicacio- nes móviles . . . . .	22
1.3. Acceso múltiple no ortogonal (NOMA) . . . . .	26
1.4. Esquemas de calendarización para las nuevas generaciones de sistemas celulares . . . . .	31
<b>2. Estado del arte</b>	<b>35</b>
2.1. Estado del arte de los esquemas de calendarización . . . . .	35
2.2. Distribución de usuarios no uniforme . . . . .	41
2.3. Tráfico Skype . . . . .	46
<b>3. Calendarización persistente</b>	<b>51</b>
3.1. Asignación dinámica . . . . .	53

---

3.2. Asignación persistente . . . . .	55
3.3. Número de bloques de recursos para distintos esquemas de modulación y codificación . . . . .	57
3.4. Análisis de teletráfico para usuarios persistentes. . . . .	63
3.4.1. Diagrama de estado . . . . .	64
3.4.2. Diagrama de transición de estados . . . . .	66
3.4.3. Tasas de cambio de estado y probabilidad de estado . . . . .	66
<b>4. Análisis de SIR y estadísticas de orden</b>	<b>69</b>
4.1. Cálculo de la SIR . . . . .	70
4.2. Esquemas de asignación de recursos. . . . .	72
4.3. Análisis de estadísticas de orden. . . . .	73
4.4. Resultados numéricos . . . . .	79
<b>5. Esquema de acceso coordinado para sistemas celulares 5G</b>	<b>87</b>
5.1. Tendencias para 5G . . . . .	89
5.2. Características de NOMA . . . . .	91
5.3. NOMA básico con receptor SIC . . . . .	92
5.4. NOMA Red . . . . .	95
5.5. Tasa de transmisión promedio . . . . .	101
5.6. Ganancia de canal . . . . .	103
<b>Conclusiones generales y trabajo a futuro</b>	<b>109</b>
<b>Apéndice A</b>	<b>113</b>
<b>Rerefencias</b>	<b>119</b>

---

# Índice de figuras

1.1. Prefijo cíclico en OFDM. . . . .	21
1.2. Esquema OFDMA. . . . .	22
1.3. Esquema NOMA. . . . .	28
1.4. Uso de SIC en NOMA. . . . .	29
1.5. Comparación de capacidad de canal OMA vs NOMA: a) UL; b) DL. . . . .	31
2.1. Ejemplo de los tres modelos de densidad de usuarios ( $G_{Tun} = 1$ ). . . . .	42
2.2. Patrón de distribución de usuarios para diferentes modelos de densidad basados en la distancia. . . . .	45
2.3. Resumen de método analítico. . . . .	50
3.1. Asignación dinámica de usuarios VoIP. . . . .	53
3.2. Asignación persistente de usuarios VoIP. . . . .	55
3.3. Diagrama de transición de posibles estados del sistema. . . . .	65
3.4. Diagrama de transición de estados. . . . .	66
4.1. Sistema de 19 celdas. . . . .	71
4.2. Probabilidad de áreas para cada MCS. . . . .	75
4.3. Probabilidad de áreas considerando estadísticas de orden. . . . .	76
4.4. Modelo de distribución de usuarios. . . . .	78
4.5. Estados válidos para 5 MHz. . . . .	80
4.6. Un caso específico de estados válidos para 5 MHz. . . . .	81
4.7. Probabilidad de terminación forzada. . . . .	82
4.8. Probabilidad de asignación de la n-ésima mejor banda, $C=10$ . . . . .	82
4.9. SIR vs asignación de la n-ésima mejor banda. . . . .	83



---

4.10. Tasa promedio de datos vs asignación de la n-ésima mejor banda. . . . .	84
4.11. Probabilidad de bloqueo para tráfico skype. . . . .	84
4.12. Probabilidad de terminación forzada para tráfico skype. . . . .	85
5.1. Propuestas de tendencias para 5G. . . . .	88
5.2. NOMA con receptor SIC. . . . .	93
5.3. Proceso SIC. . . . .	94
5.4. NOMA red. . . . .	96
5.5. Esquema coordinado NOMA-OMA. . . . .	98
5.6. <i>Throughput</i> para diferentes usuarios con NOMA y OMA. . . . .	106
5.7. <i>Throughput</i> NOMA red vs esquema coordinado. . . . .	107

---

# Índice de tablas

2.1. Ganancia $G_{Tun}$ . . . . .	46
3.1. Parámetros del modelo de tráfico VoIP. . . . .	58
3.2. Configuración de ancho de banda de transmisión. . . . .	59
3.3. Eficiencia relativa de distintos anchos de banda. . . . .	60
3.4. Número de RB y MCS para tráfico VoIP. . . . .	60
3.5. Número de elementos de recurso. . . . .	62
3.6. Número de RB y MCS para tráfico skype. . . . .	63
4.1. Modulación y codificación vs SIR. . . . .	75
4.2. Probabilidades para cada MCS considerando estadísticas de orden. . .	76
4.3. Parámetro $G_{Tun}$ . . . . .	77
4.4. Probabilidades para cada MCS considerando estadísticas de orden. . .	78
4.5. Parámetros del sistema. . . . .	79
4.6. Número de estados válidos. . . . .	79

---

# Resumen

En la presente tesis doctoral se desarrolla una metodología de evaluación del desempeño de sistemas celulares con tráfico VoIP y tráfico Skype y calendarización persistente basados en estadísticas de orden. Trabajamos en este tipo de tráfico dado que VoIP es el servicio multimedia con el crecimiento más rápido en los sistemas celulares actuales y esta tendencia continuará en el futuro. La transmisión de voz por parte de los usuarios VoIP con esquemas de calendarización dinámico podría convertirse en un cuello de botella en los canales de señalización debido a la transmisión periódica de pequeños paquetes de señalización. Para resolver este problema, se ha propuesto una calendarización persistente. Dado que la información de asignación de recursos se indica en la primera ráfaga de datos, no se requiere la asignación de recursos de información para la duración de las ráfagas de datos. Esta tesis doctoral presenta un estudio del rendimiento de la calendarización persistente basada en la transmisión de tráfico VoIP y tráfico Skype, utilizando como herramientas matemáticas las estadísticas de orden y el análisis de teletráfico. Se realiza una combinación de análisis de interferencia y procesos de nacimiento y muerte para obtener la tasa de transmisión promedio en el sistema, utilizando la ecuación de Shannon. Obtenemos las estadísticas de orden de las SIR de cada banda (relación de señal a interferencia) considerando la interferencia cocanal. Las estadísticas de orden se utilizan luego para analizar un esquema de retroalimentación de mejor banda, en el que cada usuario informa el índice de sus mejores bandas a la estación base. Si varios usuarios presentan una banda como su mejor banda, se analizan dos enfoques diferentes: asignación aleatoria y asignación de máxima SIR. Se obtiene la función de densidad de probabilidad (pdf) de la SIR, se observa que la asignación de máxima SIR tiene una ganancia significativa de hasta 20 % con respecto a la asignación aleatoria. Después,

se obtiene la probabilidad de usar un esquema de modulación y codificación específico. Para obtener la probabilidad de terminación forzada y la tasa de transmisión promedio de datos, estas probabilidades se introducen en el proceso de nacimiento y muerte como diferentes tasas de arribos y tiempos de servicio. Se observa que las estadísticas de orden modelan los sistemas celulares con mayor precisión. Como ejemplo para la distribución de usuario uniforme y tráfico de voz sobre VoIP, los resultados numéricos muestran que la probabilidad de terminación forzada se corrige en más del 70% en comparación sin estadísticas de orden. Además, utilizando la metodología desarrollada con estadísticas de orden, se realiza una evaluación para una distribución de usuarios no uniforme con tráfico Skype, considerado como un servicio muy importante para los sistemas celulares 5G.

Como propuesta final, nos centramos en la red de acceso múltiple no ortogonal (NOMA) con multiplexación de usuario en el dominio de la potencia en el lado del transmisor y cancelación de interferencia sucesiva (SIC) en el lado del receptor. El principal problema presentado en NOMA Red es el diseño del predecodificador. Tiene que aplicarse a los usuarios con condiciones de canal débiles para mitigar eficazmente la interferencia entre las celdas. Proponemos un esquema de operación coordinado entre OMA y NOMA que elimina el predecodificador y aumenta el rendimiento del sistema. La ganancia del canal se analiza mediante estadísticas de orden obteniendo su función de densidad de probabilidad. Se proporcionan resultados de simulación. Mostramos, bajo múltiples configuraciones, que la velocidad de transmisión lograda por nuestro esquema es mayor en más del 42% en comparación con NOMA en el borde de la celda.

---

# Abstract

In this thesis, a methodology for evaluating the performance of cellular systems with VoIP and Skype traffic and persistent scheduling based on order statistics is developed. We work in this type of traffic because VoIP is the multimedia service with the fastest growth on current cellular systems and this trend will continue in the future. Voice transmission by VoIP users with traditional dynamic scheduling schemes could become a bottleneck in the signaling channels due to the periodic transmission of signaling small packets. To solve this problem, persistent scheduling has been proposed. Since the resource allocation information is indicated in the first burst of persistent data no information resource allocation is required for the duration of the data bursts. This thesis presents a performance study of the persistent scheduling based on the transmission of VoIP traffic, using as mathematical tools order statistics and teletraffic analysis. A combination of interference analysis and birth and death processes are performed to obtain the average transmission rate in the system using the Shannon equation. We obtain the order statistics of band SIR(signal-to-interference ratio) by considering the co-channel interference. The order statistics are then used to analyze a best-m band feedback scheme, in which every user reports the index of their best-m bands to the BS. If multiple users present a band as their best band, two different approaches are analysed: Random allocation and Maximum SIR allocation. The probability density function (pdf) of the SIR was obtained, it is observed that the Maximum SIR allocation has a significant gain of up to 20 % in regard to random allocation. After that, the probability of using a specific modulation and coding scheme is obtained. In order to obtain the probability of forced termination and the average data rate, these probabilities are introduced in the birth and death process as a different arrival rates and service times. It is

observed that the order statistics model the cellular systems more accurately. As an example for uniform user distribution and voice over VoIP traffic, the numerical results show that the probability of forced termination is corrected by more than 70 % in comparison without order statistics. More over, using the methodology developed with order statistics, an evaluation is made for a non-uniform user distribution with Skype traffic, considered as a very important service for 5G cellular systems.

As a final proposal, we focus on non-orthogonal multiple access (NOMA) network with power-domain user multiplexing at the transmitter side and successive interference cancellation (SIC) on the receiver side. The main problem presented in NOMA network is the design of the precoder. It has to be applied to users with weak channel conditions to effectively mitigate interference between cells. We propose a coordinated operation access scheme between OMA and NOMA that removes the precoder and increases the performance of the system. The channel gain is analyzed by order statistics obtaining its probability density function. System level simulation results are provided. We show, under multiple configurations, that the transmission rate achieved by our scheme is higher by more than 42 % compared to NOMA on the edge of the cell.

---

# Introducción

El presente trabajo tiene como tema de investigación el Análisis de Teletráfico de Sistemas Celulares con Usuarios VoIP Persistentes y Calendarizador de Máximo SIR Basado en Estadísticas de Orden.

Para realizar una transmisión exitosa las interferencias en un sistema de comunicaciones celular deben ser eliminadas o al menos controladas. Debido a que en el canal los recursos se comparten, la asignación es una parte crítica para un funcionamiento apropiado de la red. En este trabajo nos enfocamos en los esquemas de acceso al medio OFDMA y NOMA.

Los esquemas de acceso múltiple no solo son necesarios en los sistemas de comunicaciones, también en otros sistemas tales como un sistema de computadoras, sistemas de almacenamiento o un servicio de algún tipo, donde un recurso se comparte por un número de usuarios independientes. En este trabajo nos centramos en un canal de comunicaciones compartido.

La tecnología de acceso múltiple es muy importante para los sistemas celulares. Mediante las técnicas de acceso múltiple, los usuarios pueden acceder simultáneamente al medio físico y compartir los recursos finitos del sistema, tales como espectro, tiempo y potencia.

Es bien conocido que se utilizaron FDMA para la primera generación de sistemas celulares (1G), TDMA sobre todo para la segunda generación de sistemas celulares (2G), CDMA para la tercera generación de sistemas celulares (3G) y OFDMA para la

cuarta generación de sistemas celulares (4G). En estos esquemas de acceso múltiple convencionales, a los diferentes usuarios se les asignan los recursos ortogonales, ya sea en el tiempo, frecuencia o en el dominio del código con el fin de evitar o mitigar la interferencia interusuario. Sin embargo, el rápido crecimiento de Internet móvil impulsará a 1000 veces el aumento del tráfico de datos en 2020 para 5G. Por lo tanto, la eficiencia espectral se convierte en uno de los principales retos para manejar este tipo de tráfico de datos. 5G tiene que soportar la conectividad masiva de los usuarios y/o dispositivos para satisfacer la demanda de baja latencia, dispositivos de bajo costo, y diversos tipos de servicios. Para satisfacer estos requisitos, son necesarias las tecnologías mejoradas. Hasta ahora, se han propuesto algunos candidatos potenciales para hacer frente a los retos de 5G, tales como MIMO masivo, las comunicaciones de ondas milimétricas, la red ultra densa, y acceso múltiple no ortogonal (NOMA).

En este trabajo, presentamos un análisis de la calendarización persistente basado en la transmisión de tráfico VoIP y tráfico Skype. El desempeño del sistema se evalúa por una combinación de procesos de nacimiento y muerte y análisis de interferencia.

Para el análisis se utilizan estadísticas de orden como herramienta matemática para obtener la función de densidad de probabilidad de la relación señal a interferencia considerando la interferencia cocanal. Las estadísticas de orden se utilizan para obtener la mejor banda para tener una mejor transmisión.

Por último se desarrolla una propuesta de evaluación para un esquema NOMA-Red utilizando el modelo desarrollado con estadísticas de orden. Finalmente, se presentan las conclusiones de la tesis y el trabajo a futuro.



---

# Objetivos

El objetivo general del presente trabajo doctoral es desarrollar una metodología de evaluación del desempeño de sistemas celulares OMA y NOMA con tráfico VoIP y tráfico Skype y calendarización persistente basados en estadísticas de orden para diferentes distribuciones de usuarios.

Para su consecución, se establecen los siguientes objetivos particulares:

1. Realizar un estudio detallado de los esquemas de calendarización en los sistemas celulares.
2. Estudiar los modelos matemáticos para las distribuciones de los usuarios.
3. Realizar un estudio detallado del tráfico VoIP y tráfico Skype.
4. Desarrollar una metodología de análisis de interferencia, para obtener la SIR, la tasa de transmisión alcanzada por cada usuario, las estadísticas de orden, la probabilidad de usar un esquema de modulación y codificación específica para dos esquemas de asignación: Asignación aleatoria y Asignación de máximo SIR.
5. Desarrollar un esquema de asignación de recursos para usuarios VoIP, mediante un proceso de nacimiento y muerte en función del esquema de modulación y codificación que corresponde a cada usuario que nos permita evaluar la probabilidad de terminación forzada, en función de las diferentes tasas de arribos y tiempos de servicio.

6. Realizar un estudio detallado de los sistemas celulares 5G, enfocándose en los esquemas de acceso para desarrollar una propuesta de operación coordinado OMA-NOMA.

---

# Capítulo 1

## Generalidades

Para que una transmisión sea exitosa en los sistemas celulares, se debe eliminar o al menos controlar la interferencia. El canal se convierte entonces en un recurso compartido cuya asignación es crítica para el funcionamiento correcto de la red. La tecnología de acceso múltiple (MA) es de gran importancia para el sistema celular. A través de las técnicas de acceso múltiple, los usuarios pueden acceder simultáneamente al medio físico y compartir los recursos finitos del sistema, como el espectro, el tiempo o potencia [1].

En la historia de las comunicaciones inalámbricas de la primera generación (1G) a la cuarta generación (4G), el esquema de acceso múltiple ha sido la tecnología clave para distinguir los distintos sistemas celulares. En estos esquemas de acceso múltiple convencionales, a los diferentes usuarios se les asignan los recursos ortogonales, ya sea en el tiempo, frecuencia o en el dominio del código con el fin de evitar o mitigar la interferencia interusuario. De esta manera, la ganancia de multiplexación se puede lograr con una complejidad razonable. Sin embargo, el crecimiento rápido de Internet móvil impulsará a 1000 veces el aumento del tráfico de datos en el 2020 para 5G. Por lo tanto, la eficiencia espectral se convierte en uno de los principales retos para manejar este tipo de tráfico de datos. Por otra parte, debido al desarrollo rápido de Internet de las Cosas (IOT), 5G tiene que soportar la conectividad masiva de los usuarios y/o dispositivos para satisfacer la demanda de baja latencia, dispositivos de bajo costo, y diversos tipos de servicios. Para satisfacer estos requisitos, son

necesarias las tecnologías mejoradas. Hasta ahora, se han propuesto algunos candidatos potenciales para hacer frente a los retos de 5G, tales como MIMO masivo, las comunicaciones de ondas milimétricas, la red de ultra densa, y acceso múltiple no ortogonal (NOMA) [2].

## 1.1. Acceso múltiple por división de frecuencia ortogonal (OFDMA)

Es una forma de modulación por portadora múltiple, donde el espacio entre portadoras se selecciona de tal forma que las portadoras sean ortogonales entre sí [3]. OFDM (Multiplexación por División de Frecuencia Ortogonal) fue propuesto por primera vez, hace más de 4 décadas por R. W. Chang. El principio básico de OFDM es dividir el espectro disponible en bandas pequeñas de canales paralelos que se conocen como subportadoras, y se transmite la información en estos canales paralelos. El nombre viene del hecho de que las respuestas en frecuencia de los canales se traslapan y son ortogonales.

Cuando la modulación se aplica a una portadora, las bandas laterales se dispersan hacia afuera a cada lado. Esto es necesario para que el receptor pueda recibir la señal y demodular exitosamente el dato. Como resultado de ello las señales se transmiten cerca una de otra, deben ser separadas de modo que el receptor pueda separarlas utilizando un filtro, por lo tanto hay una banda de guarda entre ellas. Esto no es el caso de OFDM, aunque las bandas laterales de cada portadora se superponen, se pueden recibir sin la interferencia debido a que son ortogonales entre sí. Esto se logra haciendo que la separación entre portadora sea igual o recíproco del periodo de símbolo. Una señal OFDM consiste de un número de portadoras espaciadas.

La ortogonalidad se puede perder cuando la señal pasa por un canal disperso, debido a interferencia inter símbolo. Para evitar esta interferencia se utiliza un prefijo cíclico (CP), el cual consiste en agregar la parte final de la señal OFDM al inicio de dicha señal, como se muestra en la Figura 1.1 [4].

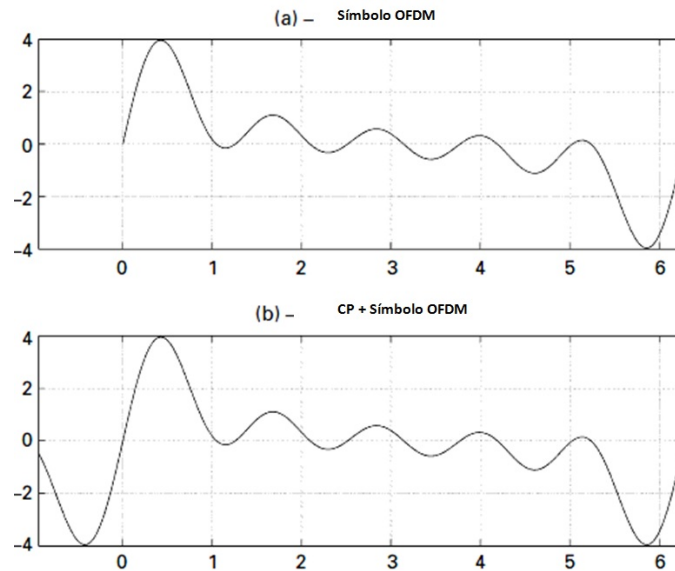


Figura 1.1: Prefijo cíclico en OFDM.

El tamaño del prefijo cíclico se busca generalmente mayor que el retardo de dispersión máximo. El prefijo cíclico hace que se transmitan las señales OFDM periódicamente y ayuda a evitar interferencia inter símbolo e interferencia entre subportadoras.

OFDM también utiliza múltiples subportadoras, pero las subportadoras están muy próximas la una a la otra sin causar interferencias, quitando las bandas de guarda entre subportadoras adyacentes. Esto es posible porque las frecuencias (subportadoras) son ortogonales.

Al igual que OFDM, OFDMA utiliza también subportadoras múltiples, pero las subportadoras se dividen en grupos de subportadoras. Cada grupo se llama subcanal. Las subportadoras de un subcanal no necesariamente deben ser adyacentes. En el enlace descendente, un subcanal puede ser destinado a diferentes receptores. En el enlace ascendente, a un transmisor se le pueden asignar uno o más subcanales.

La subcanalización define los subcanales que pueden ser asignados a los abonados, dependiendo de sus condiciones de canal y los requerimientos de datos. Usando

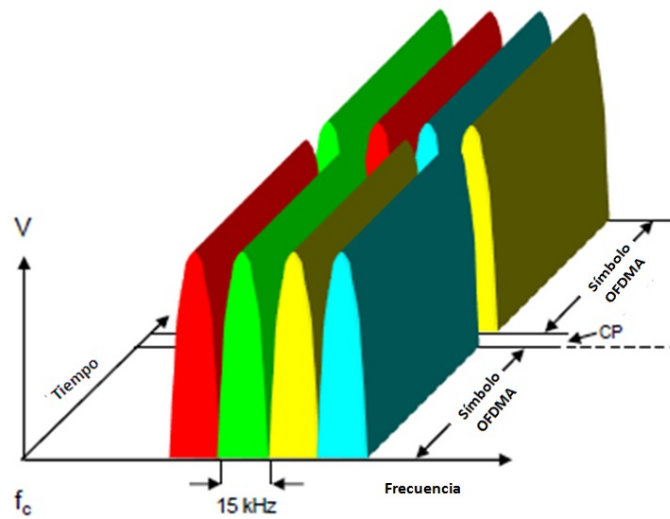


Figura 1.2: Esquema OFDMA.

la subcanalización las estaciones base pueden asignar más potencia de transmisión a usuarios con baja relación señal a ruido y menos potencia a los usuarios con mayor relación señal a ruido.

Se puede definir OFDMA como OFDM multiusuario que permite accesos múltiples en el mismo canal. OFDMA distribuye subportadoras entre todos los usuarios, para que puedan transmitir y recibir al mismo tiempo dentro de un solo canal. Este esquema se muestra en la Figura 1.2. Los sistemas celulares de cuarta generación utilizan acceso OFDMA

## 1.2. Perspectiva sobre la próxima generación de sistemas de comunicaciones móviles

Para cumplir con los requisitos de 5G que permiten una mayor capacidad, velocidades más rápidas, mayor conectividad, mayor confiabilidad, menor latencia, mayor versatilidad y topologías específicas del dominio de aplicación, los nuevos conceptos y enfoques de diseño son muy necesarios. Los estándares actuales para 4G pueden

influir en la introducción de características de radio y soluciones de red para sistemas 5G. Nuevas arquitecturas de red que explotan nuevos espectros de frecuencia (por ejemplo, ondas milimétricas) están emergiendo de laboratorios de investigación de todo el mundo. Además se están desarrollando terminales y receptores avanzados para optimizar el rendimiento de la red. La división de los planos de control y datos (actualmente estudiados en 3GPP) es un paradigma interesante para 5G, junto con sistemas MIMO avanzados, redes definidas por software, Internet de las Cosas y computación en la nube. Además, se requieren nuevos protocolos de radio que permitan tráfico heterogéneos [5].

Como se discute en [6]-[8], podemos esperar que el tráfico de datos móviles crezca en un factor de 1000 para el año 2020. No es posible soportar este tipo de crecimiento del tráfico en las redes de comunicación inalámbricas basándose en una única solución [9]. Se requieren varias mejoras y modificaciones a los sistemas existentes, nuevos conceptos y mejoras para abordar el problema del crecimiento exponencial del tráfico de datos móviles. Las posibles soluciones son, por ejemplo, nuevas bandas de frecuencia, nuevos diseños de capas físicas, esquemas masivos de múltiples entradas de múltiples salidas (MIMO) [10], redes autoorganizadas, etc. Todos estos candidatos a la solución deben estudiarse para alcanzar los objetivos exhaustivos de rendimiento que exige el crecimiento proyectado. Hay varias iniciativas actuales que se centran en la investigación relacionada con 5G teniendo en cuenta las soluciones antes mencionadas [11].

Es un momento crucial en la investigación de telecomunicaciones, ya que la conectividad se convierte en el aspecto importante de los nuevos mercados, como la automoción, las ciudades inteligentes, la red inteligente y la salud. La infraestructura y las tecnologías de la red de la quinta generación (5G) deben soportar los requisitos de estos mercados, así como a los servicios de banda ancha móvil. Las redes 5G serán desafiadas por la conectividad “en todas partes”, “todo el tiempo” con tráfico procedente de “Internet de las cosas” (IoT). Las demandas de velocidad de datos, latencia y movilidad son extremadamente variables y los mecanismos de comunicación presentan nuevos desafíos para el diseño de la red. Las redes 5G no son simplemente

consideradas como una evolución de las redes 4G existentes, sino que el énfasis en una variedad de paradigmas de comunicación emergentes como la comunicación tipo máquina y el IoT ha cambiado la filosofía de diseño de la futura comunicación. El diseño de la red 5G será impulsado por el servicio con miles de millones de dispositivos nuevos que generan tráfico heterogéneo. Los requisitos de velocidad y QoS van a ser muy estrictos y heterogéneos en comparación con las redes 4G [12].

En relación con los sistemas de comunicación móvil actuales, se estima que 5G deberá tener un volumen de datos móviles 1000 veces mayor por área. Si se pretende lograr un aumento de la capacidad de red de 1000 veces en los años siguientes para satisfacer las experiencias de los usuarios y ofrecer latencia ultrabaja, será necesario ir más y más denso con las implementaciones celulares para aprovechar la reutilización del espectro. Los marcos arquitectónicos emergentes, como la red de acceso de radio en la nube, las redes definidas por software y la virtualización de redes desempeñarán un papel importante para proporcionar soluciones eficientes para mejorar la eficiencia espectral en redes de celdas pequeñas ultradensas y manejar la escasez de espectro prevista para 2020 [12].

Anticipando que se intercambiarán muchos más datos inalámbricos, es importante darse cuenta de la necesidad de comunicación en otras bandas. Además de las bandas de frecuencia de microondas, se han propuesto bandas de frecuencia de ondas milimétricas (*mmWave*) como medio para satisfacer la escasez de espectro. Sin embargo, la comunicación en las bandas *mmWave* es desafiante debido a la alta pérdida por atenuación resultante de los deterioros atmosféricos. La investigación sobre el modelado de canales, las soluciones a las limitaciones de hardware debido a más requisitos de procesamiento y el desarrollo de tecnología para utilizar con eficacia las bandas de *mmWave* desempeñará un papel clave en la realización de redes 5G.

MIMO masivo es otra tecnología para alcanzar los objetivos de las redes inalámbricas 5G. Además de lograr la diversidad mediante la tecnología MIMO, el concepto MIMO masivo, donde las matrices de antenas con decenas o cientos de elementos se despliegan en la estación base (y receptores) para lograr alta eficiencia espectral y de



energía. Las técnicas de procesamiento de la señal y la estimación de la información del estado del canal para un número tan grande de antenas requieren nuevas técnicas de capa física para realizar de manera efectiva las ganancias de los grandes conjuntos de antenas, lo que abre nuevos desafíos de investigación. En el contexto de las redes 5G, existen numerosos desafíos de investigación para diseñar sistemas de capas cruzadas para una mejor calidad de servicio (QoS). Las tecnologías y técnicas de capas física y MAC deben desarrollarse más e integrarse bien en los marcos emergentes para cumplir con los desafíos mencionados anteriormente.

La demanda cada vez mayor de servicios inalámbricos y las continuas mejoras en la tecnología inalámbrica han llevado a la aparición de diferentes tipos de redes inalámbricas. Estas redes emergentes incluyen redes 5G [13], Internet de las cosas (IoT) [14], redes de comunicación para redes inteligentes, etc. Se prevé que las redes 5G adoptarán una serie de conceptos y tecnologías como celdas pequeñas densas y redes heterogéneas, comunicaciones de dispositivo a dispositivo (D2D), algoritmos y protocolos de eficiencia energética, transmisión multibanda y full dúplex, y nuevas técnicas de acceso múltiple como la técnica de acceso múltiple no ortogonal (NON-OMA), etc.

El IoT es un concepto emergente en el que una variedad de objetos inteligentes o cosas que nos rodean, como teléfonos móviles, sensores, actuadores, etiquetas de identificación por radiofrecuencia, etc., se integrarán perfectamente y se comunicarán e interactuarán entre sí para lograr objetivos comunes [14]. En general, estas redes tienen como objetivo proporcionar servicios inalámbricos a todos los usuarios de la red con buena calidad de servicio y conectividad ubicua y de alta velocidad de datos. Junto con los beneficios previstos, estas redes emergentes presentan numerosos desafíos, como la asignación y gestión del espectro de radio, la coexistencia de diferentes redes y el aumento del consumo de energía, etc. El uso eficaz de estas redes y el afrontamiento de los desafíos asociados sobre el modelado y diseño óptimo de las redes, así como sobre el método de optimización y los algoritmos para una gestión y utilización óptimas del espectro radioeléctrico y el consumo de recursos energéticos.

Recientemente, NOMA se ha revelado como una técnica prometedora de acceso múltiple para 5G y más allá de redes inalámbricas 5G [15]. La idea básica de NOMA reside en el uso simultáneo del mismo espectro de radio por parte de múltiples usuarios a costa de pequeñas interferencias entre los usuarios. Las redes que emplean NOMA pueden atender a un mayor número de usuarios en las bandas de espectro disponibles y pueden servir a los usuarios individuales con mayor ancho de banda. Una de las ideas clave para una operación exitosa de NOMA es formar grupos de usuarios (explotando las diferencias en las ganancias de sus canales) y permitirles transmitir en el mismo recurso de radio con la potencia apropiada, y luego usar la cancelación de interferencia sucesiva (SIC) en el (los) receptor (es) para decodificar la señal de mensaje de diferentes usuarios [16].

NOMA es una solución prometedora para acomodar a más usuarios mediante la asignación de recursos de NOMA en dominios de potencia o código, en comparación con el acceso múltiple ortogonal. En [12] investigan la aplicación de NOMA con cancelación de interferencia sucesiva en sistemas celulares MIMO multiusuario de enlace descendente (DL), donde el número total de antenas receptoras en una celda en el equipo de usuario es mayor que el número de antenas de transmisión en la estación base.

### 1.3. Acceso múltiple no ortogonal (NOMA)

En la historia de las comunicaciones inalámbricas de la primera generación (1G) a la cuarta generación (4G), el esquema de acceso múltiple ha sido la tecnología clave para distinguir los distintos sistemas celulares. En estos esquemas de acceso múltiple convencionales, a los diferentes usuarios se les asignan los recursos ortogonales, ya sea en el tiempo, frecuencia o en el dominio del código con el fin de evitar o mitigar la interferencia interusuario. De esta manera, la ganancia de multiplexación se puede lograr con una complejidad razonable. Sin embargo, el crecimiento rápido de Internet móvil impulsará a 1000 veces el aumento del tráfico de datos en el 2020 para 5G. Por lo tanto, la eficiencia espectral se convierte en uno de los principales retos para mane-

jar este tipo de tráfico de datos. Por otra parte, debido al desarrollo rápido de Internet de las Cosas (IOT), 5G tiene que soportar la conectividad masiva de los usuarios y/o dispositivos para satisfacer la demanda de baja latencia, dispositivos de bajo costo, y diversos tipos de servicios. Para satisfacer estos requisitos, son necesarias las tecnologías mejoradas. Hasta ahora, se han propuesto algunos candidatos potenciales para hacer frente a los retos de 5G, tales como MIMO masivo, las comunicaciones de ondas milimétricas, la red ultra densa, y el acceso múltiple no ortogonal (NOMA) [2].

En los esquemas convencionales de OMA, varios usuarios son asignados con recursos de radio que son ortogonales en el dominio tiempo, la frecuencia o del código. Idealmente, no hay interferencias entre múltiples usuarios, debido a la asignación de recursos ortogonales, la detección de un solo usuario se puede utilizar para separar las señales de los diferentes usuarios. En teoría, se sabe que el OMA no siempre puede alcanzar la capacidad de la suma de tasas de los sistemas inalámbricos multiusuario [17]. Aparte de eso, en los sistemas de OMA convencionales, el número máximo de usuarios admitidos está limitada por la cantidad total y la granularidad de calendarización de los recursos ortogonales.

Recientemente, NOMA se ha investigado para hacer frente a los problemas de OMA como se mencionó anteriormente. Básicamente, NOMA permite interferencias controlables para la asignación de recursos no ortogonales con el aumento tolerable en complejidad del receptor, usando cancelación de interferencia sucesiva (SIC, por sus siglas en inglés, Successive Interference Cancellation). El esquema NOMA se muestra en la Figura 1.3.

En el receptor, se utiliza SIC comúnmente para realizar la detección multiusuario. Debido al efecto cerca-lejos, las condiciones del canal pueden variar significativamente entre los usuarios. SIC se aplica a los usuarios con una relación señal a ruido más interferencia (SINR) relativamente alta.

Se supone que cualquier usuario puede decodificar correctamente las señales de otros usuarios cuya orden de decodificación viene antes que el usuario correspon-

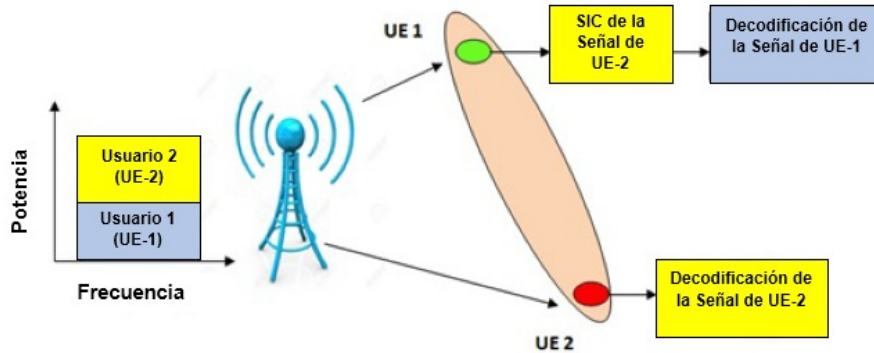


Figura 1.3: Esquema NOMA.

diente. Por lo tanto, UE- $i$  puede eliminar la interferencia entre los usuarios  $j$ , cuya ganancia de canal es menor que la del  $i$ -ésimo usuario ( $j < i$ ), es decir, las ganancias de canal se clasifican de mayor a menor.

En la Figura 1.3, el UE-2 no realiza la cancelación de interferencias puesto que viene primero en el orden de decodificación. El UE-1 decodifica primero la señal del usuario 2 y resta su componente a la señal recibida, luego, decodifica la señal del usuario 1 sin interferencia del usuario 2.

La estación base transmite una señal para UE- $i$  ( $i = 1, 2$ ), con potencia de transmisión,  $P_i$ , y la suma de  $P_i$  es igual a  $P$  ( $\sum P_i = P$ ).

La señal recibida por el UE- $i$  se representa como la ecuación (1.1).

$$y_i = h_i x + w_i \tag{1.1}$$

donde  $x = \sqrt{P_1}x_1 + \sqrt{P_2}x_2$ ,  $h_i$  es el coeficiente del canal complejo entre el usuario UE- $i$  y la estación base (BS). El término  $w_i$  denota el ruido blanco gaussiano y aditivo (AWGN) incluyendo la interferencia intercelda. La densidad espectral de potencia de  $w_i$  es  $N_0$ .

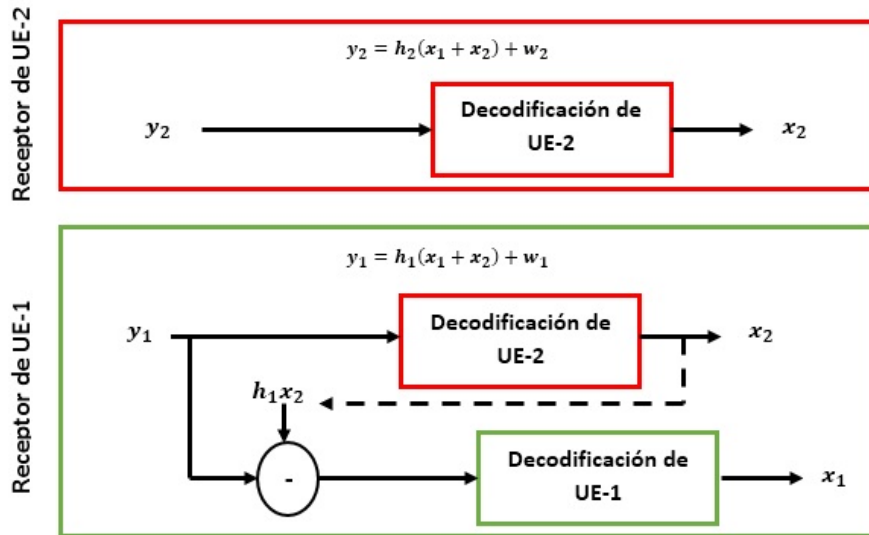


Figura 1.4: Uso de SIC en NOMA.

El proceso SIC se implementa en el receptor del UE. El orden óptimo para la decodificación es el orden decreciente de la ganancia del canal. Entonces el usuario  $i$  puede eliminar la interferencia inter-usuario del  $j$ -ésimo usuario cuya ganancia de canal  $|h_j|^2$  es menor que la ganancia de canal  $|h_i|^2$ .

En el caso de 2 usuarios mostrado en la Figura 1.3, se asume que  $|h_1|^2 > |h_2|^2$ , el usuario 2 no realiza la cancelación de interferencias ya que viene primero en el orden de decodificación. El usuario 1 decodifica primero  $x_2$  y resta su componente a la señal recibida  $y_1$ , luego, descodifica  $x_1$  sin interferencia de  $x_2$ . Este proceso se muestra en la Figura 1.4.

En comparación con OMA, las principales ventajas de NOMA incluyen las siguientes.

- Mejora de la eficiencia espectral: De acuerdo con el análisis de la capacidad multi-usuario en el trabajo pionero [18]. En el enlace ascendente NOMA es capaz de alcanzar la capacidad límite, mientras que los sistemas de OMA son en general subóptimos, como se muestra en la Figura 1.5.a. Sin embargo, en este

punto óptimo, el rendimiento de la equidad de usuario es bastante pobre cuando la diferencia de las potencias recibidas de los dos usuarios es significativo, ya que la tasa del usuario débil es mucho menor que la del usuario fuerte. En el enlace descendente, la Figura 1.5.b muestra que la frontera de tasas de NOMA está fuera de la región de tasa de OMA en general. NOMA es óptimo mientras OMA es estrictamente subóptimo si la información del estado del canal (CSI) sólo se conoce en el receptor móvil [18].

- Conectividad masiva: La asignación de recursos no ortogonal en NOMA indica que el número de usuarios o dispositivos compatibles no se limita estrictamente por la cantidad de recursos disponibles y su calendarización. Por lo tanto, NOMA tiene significativamente una capacidad de usuarios mayor que OMA mediante el uso de la asignación de recursos no ortogonal [19].
  
- Baja latencia de transmisión y el costo de señalización: En OMA con transmisión basada en concesión, un usuario tiene que enviar una petición de calendarización a la estación de base (BS). Luego, basándose en la petición recibida, la BS realiza la calendarización para la transmisión del enlace ascendente y envía una concesión sobre el canal de enlace descendente. Este procedimiento da como resultado una gran latencia y alto costo de señalización, que empeora o incluso es inaceptable en el escenario de la conectividad masiva prevista para 5G. Por el contrario, dicha calendarización dinámica no se requiere en algunos esquemas de enlace ascendente de NOMA, haciendo una transmisión de enlace ascendente libre de concesión que puede reducir significativamente la latencia de transmisión y la sobrecarga de señalización.

Debido a las anteriores ventajas potenciales, NOMA se ha investigado activamente como una tecnología prometedora para 5G.

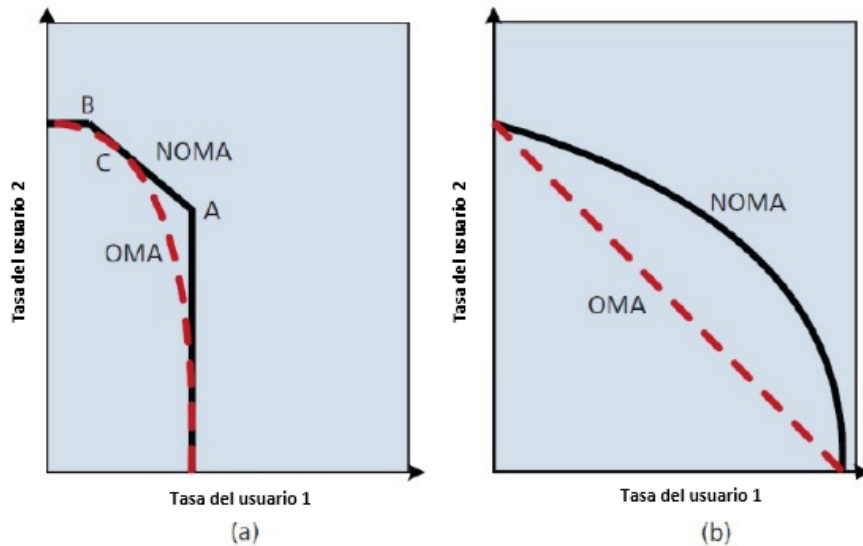


Figura 1.5: Comparación de capacidad de canal OMA vs NOMA: a) UL; b) DL.

## 1.4. Esquemas de calendarización para las nuevas generaciones de sistemas celulares

La red celular de la quinta generación (5G) exige nuevas soluciones para satisfacer, de manera eficiente, los objetivos rigurosos para la comunicación ultra-fiable y de baja latencia (URLLC, por sus siglas en inglés, Ultra-Reliable and Low Latency Communication), como la fiabilidad  $1 \times 10^{-5}$  en 1 ms.

En un sistema inalámbrico, el control de señalización del proceso de calendarización es también una fuente de errores y retardos. La calendarización semi-persistente (SPS) es una opción para reducir la señalización, llevando a una menor latencia y una mayor fiabilidad de la transmisión.

Las posibilidades están abiertas para la comunicación ultra-fiable y de baja latencia (URLLC) en las redes de quinta generación (5G), pueden traer una gran cantidad de aplicaciones novedosas para nuevos mercados. Algunos ejemplos son la automatización de la industria inalámbrica, comunicación en todo el vehículo (V2X) y control

remoto táctil [20]. Al mismo tiempo, surgen grandes desafíos para lograr los estrictos requisitos necesarios en estos contextos, fiabilidad  $1 \times 10^{-5}$  dentro de 1 ms y promedio de la latencia del plano de usuario de 0,5 ms [21].

Muchas aplicaciones requieren baja latencia y transmisiones confiables de tráfico. Por ejemplo, máquinas controladas remotamente a través de Internet táctil con retroalimentación en tiempo real [22]; y V2X, con difusión de información periódica de sensibilización en forma de mensajes de sensibilización cooperativa [23]. Dicha comunicación del tipo máquina puede generar una cantidad significativa de paquetes pequeños por un gran número de equipos de usuario (UE). La calendarización dinámica de este tipo de datos en cada intervalo de tiempo de transmisión (TTI) causaría una sobrecarga de control de señalización excesiva. Y esto, además de ser un cuello de botella en términos de capacidad, es también una fuente de errores y retardos.

En el estándar LTE se introdujo la calendarización semi-persistente (SPS) para soportar los servicios VoIP, resolviendo el problema de la exigencia de retardo para pequeños tráficos periódicos y la escasez de recursos del canal de control. En la SPS, los recursos se pre-programan con una cierta periodicidad, para evitar la sobrecarga causada por múltiples mensajes de asignación/concesión.

Recientemente, la SPS ha ganado más atención en el contexto de la reducción de la latencia considerando TTIs y periodicidades acortados. Puede beneficiar especialmente al enlace ascendente, ya que la solicitud de calendarización y el proceso de concesión pueden omitirse [24]. Para URLLC, los errores en los datos y en los canales de control deben ser evitados estrictamente con el fin de cumplir con los requisitos estrictos. En este sentido, la SPS puede traer beneficios adicionales, no sólo reduciendo la latencia, sino también el papel del canal de control como una fuente de error [25].

El inconveniente de la pre-calendarización es que, típicamente, los recursos reservados no pueden ser utilizados por otros UE, limitando la utilización de recursos.



Para URLLC, que requiere una transmisión muy robusta, el costo en términos de recursos puede ser muy alto, especialmente en malas condiciones de cobertura. Por lo tanto, emplear un esquema de retransmisión de datos como la solicitud de repetición automática híbrida (HARQ) es importante para mejorar la eficiencia de recursos [26]. De lo contrario, una gran cantidad de recursos debe reservarse para cada ciclo preestablecido, para una transmisión conservadora.

El SPS convencional incluye una calendarización persistente para la primera transmisión y una calendarización dinámica para las retransmisiones (recalendarización) [27]. Para URLLC puede ser deseable evitar también la señalización para la re-calendarización debido a los posibles errores en el canal de control. Además, la extralatenencia puede ser causada por la recalendarización tardía en escenarios de alta carga y por el propio proceso de concesión.



---

# Capítulo 2

## Estado del arte

La red de acceso LTE, basada en OFDMA, admite una amplia gama de servicios multimedia e Internet, incluso en escenarios de alta movilidad. Por lo tanto, ha sido diseñado para proporcionar altas velocidades de datos, baja latencia y una eficacia espectral mejorada con respecto a las redes anteriores. Para lograr estos objetivos, el administrador de bloques de recursos de radio explota una combinación de funciones MAC avanzadas y físicas, como compartir recursos, informes de indicadores de calidad de canal (CQI), adaptación de enlaces mediante modulación y codificación adaptables (AMC) y solicitud de retransmisión automática híbrida (HARQ) [28].

En este capítulo se describe el estado del arte de los esquemas de calendarización para los sistemas celulares. Se describe también algunas funciones de distribución de usuarios utilizadas en los sistemas celulares y distintos tipos de tráfico. Por último se muestra un resumen de las aportaciones del trabajo y de la metodología desarrollada.

### 2.1. Estado del arte de los esquemas de calendarización

En [28] los autores proporcionan una visión general de los problemas claves que surgen en el diseño de un algoritmo de asignación de recursos para redes LTE. Proporcionan un estudio amplio sobre las estrategias de asignación de paquetes de enlace

descendente propuestas recientemente para redes LTE, destacando al mismo tiempo los problemas claves que deben tenerse en cuenta al diseñar una nueva solución.

La calendarización persistente se ha estudiado mediante simulación en [29], [30] y [31]. En [29], los autores obtienen la capacidad del sistema que garantiza los requisitos de QoS, como el retardo de paquetes y la tasa de pérdida de paquetes VoIP. Comparan las capacidades de la calendarización dinámica y la calendarización semipersistente sin restricción del número de canales de control. Muestran que la capacidad de la calendarización dinámica (234 usuarios por sector) es aproximadamente un 5% más alta que la capacidad de la calendarización semipersistente (222 usuarios por sector). La ganancia de la capacidad resulta mayor debido a la multiplexación ofrecidas por HARQ. Los retardos de la calendarización semipersistente son mayores que los de la calendarización dinámica. Teniendo en cuenta la gran cantidad de señalización de control requerida en la calendarización dinámica, el precio pagado por esta ganancia es bastante alto y, por lo tanto, se puede considerar la calendarización semipersistente. Con la restricción del canal de control la calendarización semipersistente sufre mucho menos de la restricción del canal de control, mientras que el rendimiento de la calendarización dinámica depende en gran medida del número de canales de control por TTI (intervalo de tiempo de transmisión). Con la disminución del número de canales de control, la capacidad de la calendarización dinámica cae bruscamente; con 6 canales de control, puede admitir sólo 161 usuarios por sector, mientras que la calendarización semipersistente en este caso puede transportar a 218 usuarios presentando una ganancia del 35%. En [30] se estudia el impacto de diferentes tamaños de paquetes VoIP, diferentes escenarios de uso y diferentes mecanismos de calendarización. Las tramas SID no se consideran. Alrededor de 41 y 44 usuarios por sector se pueden admitir con un umbral de interrupción del 95%. La razón de tal capacidad baja es que la mayoría de los recursos que están reservados para las retransmisiones (debido a la característica de la calendarización persistente) se desperdician debido a la terminación temprana de HARQ. Muestran resultados de la capacidad VoIP para un codificador AMR 7.95 kbps. Para estas simulaciones, el paquete SID se modela y simula pero, para simplificar, usan el mismo número de bloques de recursos (RB) y diferentes MCS. Se pueden admitir aproximadamente 86

y 82 usuarios por sector para la asignación dinámica y semipersistente, respectivamente. La asignación semipersistente con una sobrecarga reducida de señalización proporciona casi la misma capacidad que la asignación dinámica. También muestran resultados de capacidad VoIP para AMR 12.2 kbps, 73 y 71 usuarios por sector son compatibles para la asignación dinámica y semipersistente, respectivamente. Al comparar las diferentes tasas de AMR bajo los mismos casos de simulación, se muestra que AMR 7.95kbps puede alcanzar una capacidad de 15-20% más alta que AMR 12.2kbps. En [31] se muestra la capacidad de usuarios VoIP en función del canal de control de enlace descendente físico (PDCCH) con diferentes calendarizadores de paquetes. Los autores muestran que con la calendarización dinámica, la capacidad de usuarios VoIP depende de manera muy lineal de la capacidad del PDCCH. Sin la adaptación del enlace, la capacidad VoIP se duplica aproximadamente cuando se duplica la cantidad del PDCCH. Con adaptación del enlace, la ganancia de un número mayor de PDCCH es aún mayor (incluso se triplica), pero con 8-10 canales de control, la capacidad de VoIP se aproxima al máximo. Por lo tanto, con calendarización dinámica, el PDCCH restringe claramente la capacidad de VoIP si la cantidad de PDCCH es inferior a 8. Con la calendarización persistente, la capacidad de VoIP no depende de la capacidad de PDCCH. La capacidad de VoIP alcanza el máximo con 4 canales de control con calendarización persistente (200 usuarios por celda) e incluso con 2 canales de control con calendarización semipersistente (185 usuarios por celda). Esto se debe a que las primeras transmisiones de paquetes VoIP se transmiten con una asignación persistente, por lo que no requieren ningún PDCCH. Sólo la primera transmisión consume un PDCCH. Las retransmisiones con HARQ están calendarizadas dinámicamente. Asumiendo un 20% de BLER para la primera transmisión, el 80% de los paquetes VoIP no necesitan recursos PDCCH. Los autores muestran que cuando se tiene menos de 9-10 canales de control, entonces no tiene sentido usar la calendarización dinámica sin adaptación al enlace. La calendarización dinámica con adaptación al enlace comienza a proporcionar una mejor capacidad VoIP que la calendarización persistente cuando ya se cuenta con 5 canales de control. Muestran que la calendarización dinámica proporciona una ganancia de SNRR de aproximadamente 78 dB en comparación con la calendarización persistente y la calendarización semipersistente proporciona una ganancia de SINR de aproximadamente 1-2 dB en

comparación con la calendarización persistente. Como se puede observar en estos trabajos el desempeño se analiza mediante simulaciones, considerando sólo un tipo de tráfico (tráfico VoIP) y distribuciones de usuarios uniforme.

En la literatura, se han publicado pocos trabajos que se centran en el modelado matemático de calendarizadores persistentes, entre los cuales se encuentran [32], [33] y [34]. En [32] el retardo en el acceso aleatorio de la calendarización semipersistente para usuarios VoIP se evalúa mediante la construcción de un modelo aproximado. Para verificar su análisis, construyen un modelo de simulación en MATLAB, examinan la viabilidad de la calendarización semipersistente con acceso aleatorio para el tráfico VoIP en LTE mediante la construcción de un modelo aproximado basado en análisis de puntos de equilibrio para analizar el rendimiento. En particular, muestran la capacidad VoIP como el número de terminales VoIP que pueden acomodarse sin exceder alguna probabilidad de pérdida de paquetes, cuando se utiliza el acceso aleatorio inicial. De este trabajo, se puede conjeturar que la calendarización semipersistente con acceso aleatorio inicial se podría utilizar alternativamente o proporcionar una mejora adicional de la capacidad VoIP. En [33] los autores estudian el rendimiento de los esquemas de asignación de frecuencias en sistemas basados en OFDMA. Desarrollan un modelo analítico para las colisiones para un número arbitrario de usuarios en las diferentes celdas, luego calculan la capacidad del sistema usando un modelo de Markov y toman en cuenta la interferencia entre celdas y su impacto en la modulación adaptativa por medio de áreas de cobertura sin considerar cambios instantáneos y estadísticas de orden. En [34] los autores derivan un modelo estadístico para el tamaño de la cola y el retardo del paquete en asignaciones persistentes estáticas y adaptativas que se pueden utilizar para cualquier proceso de arribo. Derivan la función de masa de probabilidad del tamaño de la cola y la función de densidad de probabilidad del retardo para los esquemas de asignación de recursos persistentes estáticos y adaptativos del enlace ascendente para un proceso de arribo de paquete general/arbitrario para aplicaciones M2M orientadas a usuarios y redes. Se puede observar que en estos trabajos hace falta un análisis de interferencia en conjunto con el análisis de teletráfico. También se observa que estos trabajos no consideran las estadísticas de orden.

Las versiones futuras de LTE admitirán la comunicación de dispositivo a dispositivo (D2D) asistida por red sin involucrar directamente a la estación base (BS) en el intercambio de datos entre dispositivos. El paradigma de comunicación D2D no sólo permite reducir la latencia de comunicación entre dispositivos, sino que también proporciona una posible solución para aumentar la utilización de recursos en el caso de las aplicaciones de Internet de las cosas (IoT). En la versión 13 de LTE, se ha propuesto el concepto de acceso rápido de enlace ascendente. En el acceso rápido de enlace ascendente, la sobrecarga de señalización de la solicitud de calendarización y la petición de calendarización se elimina [35].

Aunque el tráfico de datos está aumentando, el servicio de voz aún contribuye con una cantidad considerable de los ingresos de los operadores, todavía afecta mucho en la vida cotidiana de las personas. Hoy en día, las aplicaciones de Voz sobre IP (VoIP) como Skype, Gtalk y Asterisk se han vuelto muy populares en Internet. Como el sistema LTE solo es compatible con servicios de paquete, el servicio de voz utilizará VoIP que no es por conmutación de circuitos. Se espera que LTE admita una gran cantidad de usuarios VoIP al tiempo que satisface la calidad de servicio (QoS) requerida. Un gran número de usuarios VoIP consume tantos recursos, en consecuencia el rendimiento de otros servicios de datos se verá afectado. La calendarización dinámica tiene la ventaja de explotar la información instantánea del estado del canal, pero da como resultado una gran sobrecarga de señalización de control con una gran cantidad de usuarios VoIP para paquetes pequeños que aparecen con frecuencia. Para resolver este problema y admitir una gran capacidad de usuarios VoIP, en [36] proponen una calendarización semipersistente para señalar la asignación de recursos para usuarios VoIP. SPS incluye dos principios: calendarización persistente para transmisiones iniciales y calendarización dinámica para retransmisiones. Aunque SPS puede reducir de manera eficiente la demanda de señalización de control, la movilidad de los usuarios y las variaciones del estado del canal se ignoran durante el estado activo [36].

Las aplicaciones de máquina a máquina (M2M) se encuentran ampliamente distri-

buidas en una amplia variedad de mercados, incluidos los de automoción, asistencia médica y servicios públicos. Por ejemplo, una red inteligente generalmente admite múltiples aplicaciones M2M orientadas a redes y usuarios como Infraestructura de Medición Avanzada (AMI), Respuesta Automática de Demanda (DR), Detección de Fallas, Aislamiento y Restauración (FDIR) y Sistema de Medición de Área Amplia (WAMS). Estas aplicaciones pueden tener diversos requisitos de calidad de servicio (QoS). Una de las principales preocupaciones ha sido la probabilidad de un canal de acceso aleatorio sobrecargado causado por una gran cantidad de dispositivos M2M ubicados en cada celda. Uno de los principales problemas en el soporte de las aplicaciones M2M es que el canal de control de enlace descendente en el cual las concesiones de enlace ascendente son emitidas por la estación base a los dispositivos típicamente se satura antes del canal de datos de enlace ascendente. Este es un resultado directo de una gran cantidad de dispositivos, cada uno enviando un paquete pequeño. Una solución a este problema es proporcionar dispositivos M2M con una asignación persistente de recursos de enlace ascendente una vez que hayan solicitado recursos de enlace ascendente [34].

La evolución o las aplicaciones de datos de alta velocidad, como la transmisión de videos inalámbricos de alta calidad, las redes sociales y las comunicaciones de máquina a máquina, han llevado al crecimiento explosivo de la demanda de tráfico de datos para las redes inalámbricas 5G. Se prevé que el requisito de capacidad promedio de las redes inalámbricas 5G podrá alcanzar los  $25\text{Gbps}/\text{km}^2$ , que es 100 veces mayor en comparación con la capacidad del sistema actual. Mientras tanto, los requisitos de retardo para diferentes tipos de tráfico se vuelven más diversificados debido a la aparición de nuevos tipos de aplicaciones, como aplicaciones de latencia crítica como comando y control de drones. El aumento del tráfico y sus variaciones dinámicas juegan un papel cada vez más crucial en la afectación del rendimiento de retardo de las redes inalámbricas 5G [37].

Algunas aplicaciones requieren una baja latencia y transmisiones confiables de tráfico. Por ejemplo, máquinas controladas de forma remota a través de Internet con retroalimentación síncrona en tiempo real. Dicha comunicación del tipo de máquina



puede generar una cantidad significativa de pequeños paquetes mediante una gran cantidad de equipos de usuario (UE). La calendarización dinámica de este tipo de datos en cada intervalo de tiempo de transmisión (TTI) causaría una sobrecarga de señalización de control excesiva. Y esto, además de ser un cuello de botella en términos de capacidad, también es una fuente de errores y retardos. La calendarización semipersistente (SPS) se introdujo en el estándar LTE para admitir servicios de usuarios VoIP, resolviendo el problema del requisito de retardo para tráficos pequeños periódicamente y la escasez de recursos del canal de control [38].

Por lo tanto, consideramos que la calendarización persistente se utilizará para la próxima generación de sistemas celulares.

## 2.2. Distribución de usuarios no uniforme

Aplicaciones tales como descargas de videos, carga de fotos, el uso de servicios de datos en la nube perjudican severamente las redes celulares debido a su alta velocidad de datos y requerimientos de capacidad. En vez de asumir que las localizaciones de los equipos de usuarios (UE) y la estación base (BS) son independientes, en [39] introducen un modelo más realista con la dependencia entre la localización de los UE y las BS, resultando de esto una distribución de usuarios no uniforme. Por lo tanto, consideramos que la distribución de usuarios no uniforme se utilizará para las redes celulares 5G y por eso vemos la necesidad de realizar un estudio detallado de las distintas distribuciones no uniformes utilizadas y reportadas en la literatura.

Uno de los requerimientos de las redes celulares de quinta generación (5G) es incrementar la capacidad para poder operar con las futuras demandas de tráfico. Las redes heterogéneas (HetNets) se desarrollan con estaciones bases que cambian las áreas de cobertura para aumentar la capacidad y mejorar la experiencia de usuario. La distribución espacial de los usuarios es uniforme y se modela como procesos de puntos de Poisson independientes (PPP). En las HetNets reales, los usuarios tienden a ser no uniformes. En [40] introducen un modelo de densidad de usuarios no uniforme dependiendo de la distancia y la cobertura. Los autores hacen una combinación

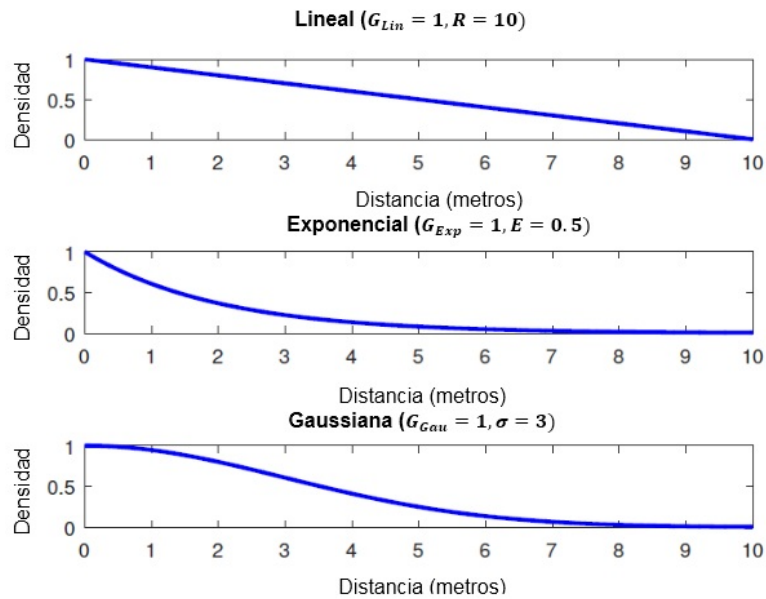


Figura 2.1: Ejemplo de los tres modelos de densidad de usuarios ( $G_{Tun} = 1$ ).

de diferentes densidades de usuarios en diferentes niveles y analizan el desempeño sobre la probabilidad de cobertura en el enlace de bajada.

En [40] los autores definen un modelo de densidad de usuarios no uniforme, donde la densidad  $f_{den}(d)$  depende de la distancia a la estación base de interés.

$f_{den}(d)$  puede ser cualquier función monóticamente decreciente con el incremento de la distancia. Los tres modelos de densidades no uniformes que se discuten en [40] son: los modelos lineal, exponencial y Gaussiano. Los ejemplos de  $f_{den}(d)$  se muestran en la Figura (2.1).

Dado  $G_{Lin}$ ,  $G_{Exp}$  y  $G_{Gau}$  representan la densidad máxima de cada modelo.  $R$ ,  $E$  y  $\sigma$  son los parámetros para controlar el nivel de la no uniformidad. Las definiciones matemáticas de las funciones de densidad de los tres modelos están dados como sigue.

El modelo de densidad Lineal de usuarios está dado por (2.1).

$$f_{denLin}(d) = G_{Tun} \cdot G_{Lin} \cdot (1 - (d/R)) \quad 0 \leq d \leq R \quad (2.1)$$

$R$  representa la pendiente de la densidad lineal de usuarios. Si  $R$  es grande (por ejemplo  $R > 800m$ ) el modelo de densidad lineal de usuarios tiene baja atracción (significa que los usuarios tienen baja correlación espacial con la estación base asociada. La densidad de la distribución de usuarios sobre un area dada es casi la misma) y se aproxima a la distribución uniforme. Si  $R$  es pequeño (por ejemplo  $R < 100m$ ), esto representa una atracción alta y los usuarios se distribuyen cerca del centro de la celda.

El modelo de densidad Exponencial de usuarios está dado por (2.2).

$$f_{denExp}(d) = G_{Tun} \cdot G_{Exp} \cdot e^{-Ed} \quad 0 \leq d \leq \infty \quad (2.2)$$

$E$  refleja la pendiente de la densidad exponencial de usuarios. Si  $E$  es grande, tiene alta densidad y alta atracción cerca del centro de la celda. Si  $E$  es pequeño, tiene una densidad de usuario casi uniforme y baja atracción. Cuando  $E = 0$ , es exactamente lo mismo que la densidad de usuario uniforme.

El modelo de densidad Gaussiano de usuarios está dado por (2.3).

$$f_{denGau}(d) = G_{Tun} \cdot G_{Gau} \cdot e^{-d^2/\sigma^2} \quad 0 \leq d \leq \infty \quad (2.3)$$

$\sigma$  refleja la pendiente de la densidad Gaussiana de usuarios. Si  $\sigma$  es grande, tiene una densidad de usuarios casi uniforme y baja atracción. Si  $\sigma$  es pequeño, tiene alta densidad de usuarios y alta atracción cerca del centro de la celda.

La Figura (2.2) muestra ejemplos de patrones de distribución de usuarios para redes heterogéneas. En estos ejemplos, hay dos niveles, como BS macro (triángulos rojos) y BS pequeños (triángulos negros). De (a) a (i), los dos niveles tienen los mismos modelos de densidad de usuario. (a) - (c) son modelos de densidad lineal

$G_{Lin} = 1$ , con parámetro  $R = 100$  para (a),  $R = 400$  para (b) y  $R = 800$  para (c). (d) - (f) son modelos de densidad exponencial  $G_{Exp} = 1$  con el parámetro  $E = 0,05$  para (d),  $E = 0,01$  para (e) y  $E = 0,001$  para (f). (g) - (i) son modelos de densidad Gaussianos  $G_{Gau} = 1$  con el parámetro  $\sigma = 40$  para (g),  $\sigma = 100$  para (h) y  $\sigma = 200$  para (i). De (j) hasta (k), los dos niveles tienen diferentes modelos de densidad de usuarios. Para (j), las macro celdas son para el modelo de densidad lineal  $G_{Lin} = 1$  con el parámetro  $R = 800$  y las celdas pequeñas son para el modelo de densidad exponencial  $G_{Exp} = 4$  con el parámetro  $E = 0,05$ . Para (k), las macro celdas son para el modelo de densidad exponencial  $G_{Exp} = 1$  con el parámetro  $E = 0,001$  y las celdas pequeñas son para modelos de densidad Gaussiana  $G_{Gau} = 4$  con el parámetro  $\sigma = 40$ . (l) da la distribución de la BS pura en la que los triángulos rojos son macro BS y los triángulos negros son BS pequeñas, sólo en caso de que los triángulos de las BS sean invisibles cuando la densidad del usuario es alta cerca de los centros de las celdas en (a) - (k). Las curvas azules son los límites de las celdas. El área está a 1 km por 1 km.

El modelo de densidad de usuario no uniforme es:

1. En cada nivel, la densidad de usuario  $f_{den}(d)$  es una función de la distancia  $d$  a su BS.
2. Diferentes niveles pueden tener modelos de densidad de usuarios diferentes.
3. La densidad del usuario en las macro celdas se acerca a la distribución uniforme (o baja atracción). La densidad del usuario en celdas pequeñas es densa alrededor de los centros de las celdas pequeñas (o atracción alta).

Existe una restricción en la densidad de usuario no uniforme  $f_{den}(d)$  de modo que el número total de usuarios debe ser el mismo o mantenerse fijo en un área determinada que cubre la BS de acuerdo con diferentes modelos de densidad de usuario en diferentes niveles. La ganancia  $G_{Tun}$  es variada, por lo que la cantidad de usuarios se mantiene en un valor específico. Si la densidad de usuario no uniforme en cada nivel se multiplica por la ganancia  $G_{Tun}$ , debería tener el mismo número de usuarios en comparación con el modelo de densidad de usuario uniforme de referencia

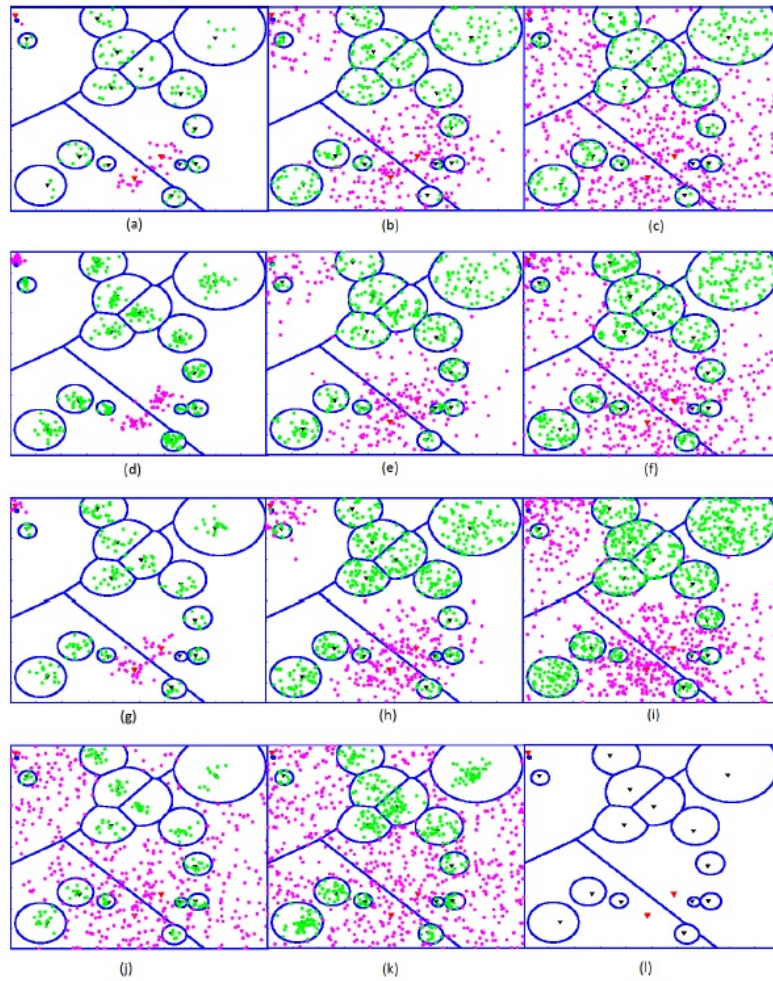


Figura 2.2: Patrón de distribución de usuarios para diferentes modelos de densidad basados en la distancia.

Tabla 2.1: Ganancia  $G_{Tun}$ .

$G_{Tun}$	Atracción alta	Atracción media	Atracción baja
Lineal	4.76	1.6	1.23
Exponencia	17.08	2.87	1.15
Gaussiana	4.93	2.08	1.38

en la misma área. Luego se nombra a  $G_{Tun}$  como la ganancia de ajuste del modelo de densidad de usuario no uniforme. La ganancia de ajuste  $G_{Tun}$  se puede obtener experimentalmente. La tabla (2.1) muestra un ejemplo de la ganancia  $G_{Tun}$  obtenida experimentalmente.

### 2.3. Tráfico Skype

Las redes domésticas se definen como dos o más nodos que se pueden comunicar entre sí directamente o mediante nodos de retransmisión en la capa física. Esta definición aplica para redes domésticas basados en tecnologías inalámbricas. Hoy en día, aplicaciones VoIP, tales como Skype, Gtalk y Asterisk son muy populares en internet. El tráfico Skype es muy popular para servicios de voz [41].

El tráfico Skype utiliza una velocidad mínima para una sesión de video de 128 kbps, que es aproximadamente 10 veces más que una llamada de voz. Por lo tanto, podemos suponer que para tráfico Skype se utilizan 10 veces más de bloques de recursos en comparación con el tráfico VoIP, dependiendo de cada esquema de modulación y codificación.

El tráfico Skype es considerado como un servicio muy importante para sistemas celulares 5G. Skype es un servicio VoIP provisto por Microsoft y que actualmente es muy utilizado internacionalmente. Los paquetes Skype arriban cada 20 ms.

Por lo tanto, consideramos que la calendarización persistente se utilizará para la próxima generación de sistemas celulares, así como también la distribución de

usuarios no uniforme para redes 5G y diferentes tipos de servicios, tales como tráfico VoIP y tráfico Skype.

Como podemos observar en los modelos analíticos sólo desarrollan un modelo de teletráfico o un modelo de análisis de interferencia de forma separada y sin utilizar estadísticas de orden. En este trabajo de tesis, presentamos un modelo analítico para analizar la calendarización persistente para la transmisión de distintos tipos de tráfico de usuarios (tráfico VoIP y tráfico Skype), considerando un análisis conjunto de teletráfico e interferencia, basados en estadísticas de orden para distribuciones de usuarios uniforme y no uniforme. El desempeño del sistema se evalúa por una combinación de procesos de nacimiento y muerte y análisis de interferencia. Se propone un modelo analítico usando cadenas de Markov para evaluar el desempeño del sistema, este modelo es una cadena de Markov en dos dimensiones. Los usuarios se dividen en dos categorías: usuarios activos y usuarios en silencio. Los cambios de estados que ocurren en el sistema son cambios de activo a silencio y viceversa. Obtenemos las estadísticas de orden las SIR's de las bandas considerando interferencia co-canal. Las estadísticas de orden se utilizan para analizar el esquema de retroalimentación de mejor banda, en el cual cada usuario reporta a la estación base el índice de su mejor banda. Si múltiples usuarios reportan una misma banda como su mejor banda, se analizan dos mecanismos de asignación de recursos: asignación aleatoria y asignación de máximo SIR. La asignación aleatoria se utiliza cuando múltiples usuarios presentan una banda  $B_i$  como su mejor banda, entonces la estación base asigna la banda a uno de los usuarios de forma aleatoria. Por otra parte, cuando múltiples usuarios reportan la banda  $B_i$  como su mejor banda, la estación base asigna la banda al usuario que presenta la mejor SIR. La SIR obtenida se utiliza para evaluar la probabilidad de utilizar un esquema de modulación y codificación (MCS) específico. Estos esquemas se introducen en el proceso de nacimiento y muerte como diferentes tasas de arribos y tiempos de servicios. Este método analítico se resume en la Figura 2.3. El método se describe como sigue:

1. Análisis de Interferencia: Este bloque nos da el análisis de interferencia experimentado por los usuarios, causando la degradación de la SIR. Primero

obtenemos la SIR, luego obtenemos la tasa de transmisión alcanzada por cada usuario usando la ecuación de Shannon para la asignación del esquema de modulación y codificación alcanzable por el usuario. Las estadísticas de orden de las SIR's se obtienen considerando el análisis de interferencia. Las estadísticas de orden se utilizan luego para analizar un esquema de retroalimentación de mejor banda, en el que cada usuario informa el índice de sus mejores bandas a la estación base. Después de obtener las estadísticas de orden de las SIR's, calculamos la probabilidad de usar un esquema de modulación y codificación de todos los disponibles para obtener las nuevas áreas de cobertura. Con este análisis, podemos asignar recursos a través de dos esquemas: Asignación aleatoria y Asignación de máximo SIR. Con los diferentes esquemas de modulación y codificación, se obtienen diferentes tasas de arribo y tiempos de servicio, que se utilizan para estimar la capacidad del sistema con el análisis de teletráfico.

2. Análisis de Teletráfico: en este bloque, el proceso de asignación de recursos para usuarios de VoIP (usuarios activos y en silencio) se realiza mediante un proceso de nacimiento y muerte. Mediante este análisis, asignamos los recursos de manera óptima en función del esquema de modulación y codificación que corresponde a cada usuario y nos permite evaluar la probabilidad de terminación forzada, en función de las diferentes tasas de arribo y tiempos de servicio.
3. Toma de Decisiones: en este bloque, se realiza la evaluación de la probabilidad de bloqueo y se compara con la probabilidad esperada. Si la condición no se cumple, las tasas de arribo se reducen y el sistema se retroalimenta. De lo contrario, se obtienen los resultados.

El resto del trabajo se organiza como sigue: en el capítulo 3, se explica el principio de la calendarización persistente para tráfico VoIP y se calcula el número de bloques de recursos para diferentes esquemas de modulación y codificación. En el capítulo 4 se presenta el análisis de interferencia y se calculan las estadísticas de orden de la SIR de las diferentes bandas, para tráfico VoIP y tráfico Skype, para distribuciones de usuarios uniforme y no uniforme, mostrando resultados con asignación aleatoria y asignación de máximo SIR. Se calcula la probabilidad de bloqueo y probabilidad de terminación forzada para los dos tipos de tráfico y dos tipos de distribuciones de



usuarios. En el capítulo 5 se muestra una propuesta de operación coordinada para redes celulares 5G utilizando el método analítico desarrollado. Por último se presentan las conclusiones y trabajo a futuro.

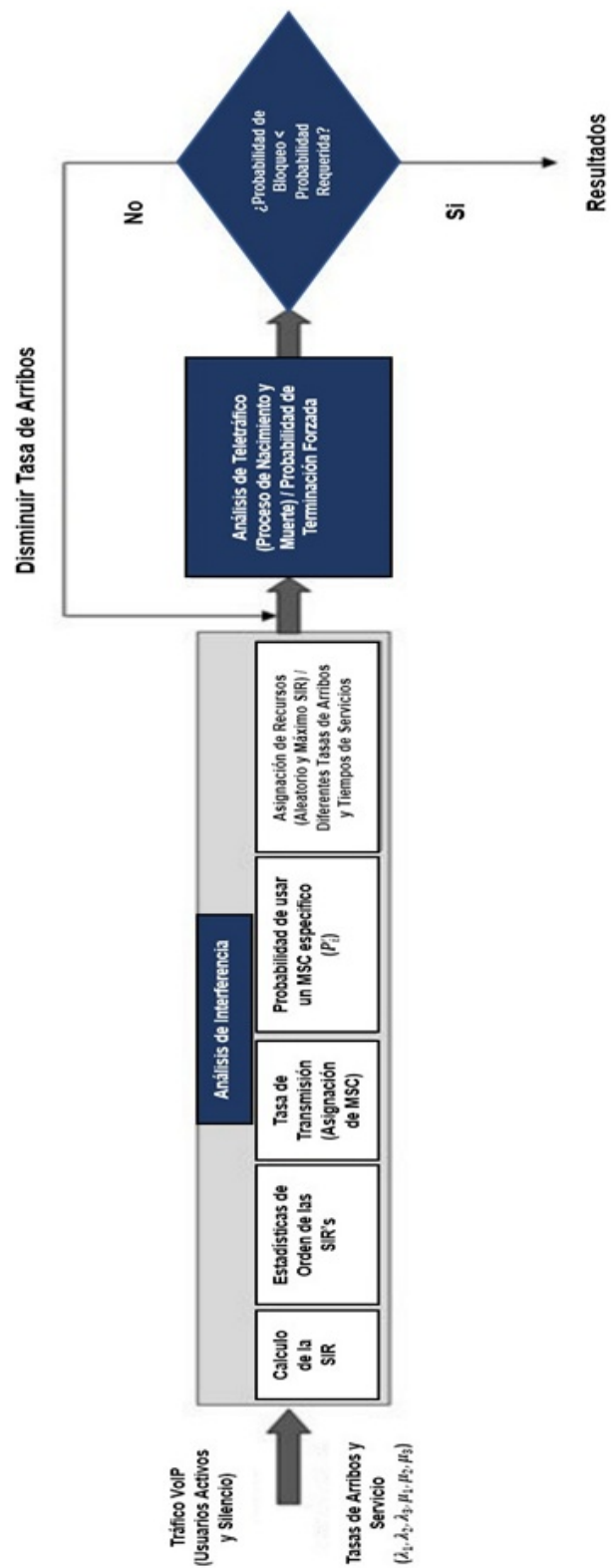


Figura 2.3: Resumen de método analítico.

---

## Capítulo 3

# Calendarización persistente

Con la popularidad creciente de los dispositivos móviles tales como los teléfonos inteligentes y las tabletas móviles, el tráfico de datos móviles ha ido aumentando constantemente. Entre las aplicaciones que muestran este crecimiento están la Voz sobre IP (VoIP) y el video en tiempo real, ambos sujetos a requisitos estrictos de Calidad de Servicio (QoS) sobre retraso y fiabilidad [42]. VoIP en las redes celulares inalámbricas ha atraído más y más interés recientemente debido a su infraestructura IP integrada. Los algoritmos de calendarización de paquetes y los métodos de asignación de recursos están destinados a desempeñar un papel fundamental entre todos los elementos del sistema.

La convergencia de redes IP agrega todos los servicios de datos en una columna vertebral. La calidad de servicio de estos servicios no se enfrenta a restricciones de capacidad en las redes de gran ancho de banda. Por otra parte, las redes de ancho de banda limitado, por ejemplo, las redes inalámbricas, están diseñados para manejar servicios de prioridad como la voz en tiempo real.

A pesar de que el tráfico de datos va en aumento, el servicio de voz sigue aportando una cantidad considerable del consumo de los recursos de los servidores, teniendo mucha influencia en la vida cotidiana de la gente. Hoy en día, las aplicaciones VoIP tales como Skype, Gtalk y Asterisk, se han vuelto muy populares en Internet. Dado que el Sistema Celular basado en la técnica de Acceso Múltiple por División de

Frecuencia Ortogonal (OFDMA) utiliza servicios de paquetes, el servicio de voz utilizará Voz sobre IP (VoIP), distinta de la Voz por conmutación de circuito (CS). A medida que más y más estaciones base LTE se están desplegando, VoIP reemplazará a la tecnología CS convirtiéndose en el principal método de servicio de voz [36].

Como se explica a continuación, un gran número de usuarios VoIP consume muchos recursos mediante los esquemas tradicionales de asignación de recursos y el rendimiento de otros servicios de datos se verá afectado en consecuencia. Por lo tanto, es importante encontrar un algoritmo de asignación de recursos eficaz para garantizar no sólo la calidad del servicio de voz, sino también el rendimiento de otros servicios de datos.

El esquema de calendarización tradicional para la transmisión de datos es la calendarización dinámica. La calendarización dinámica tiene la ventaja de explotar la información instantánea del estado del canal, pero da como resultado una alta sobrecarga de señalización de control para un gran número de usuarios VoIP para paquetes pequeños que aparecen frecuentemente.

Sin embargo, puesto que el control de señalización y los datos de usuario comparten los mismos recursos inalámbricos en sistemas OFDMA, pueden requerir una cantidad significativa de sobrecarga de señalización para soportar el tráfico de pequeño tamaño sensible al retardo, tal como servicio de VoIP. Kwon et al. [43] mostró que hasta el 60 % de los recursos está ocupado por la señalización para soportar el servicio de VoIP en el sistema IEEE 802.16e. Para hacerle frente a este inconveniente, se ha propuesto un esquema de calendarización persistente en IEEE 802.16e/m [44] [45], 3GPP LTE [46], y 3GPP2 UMB [47].

En los sistemas celulares la calendarización se realiza en la estación base y los recursos son asignados a los usuarios mediante peticiones de calendarización. En las redes actuales los servicios más importantes son los servicios VoIP, los cuales tienen características diferentes a otros servicios. El tamaño de los paquetes de datos para servicios VoIP son muy pequeños y se generan periódicamente. Considerando esta

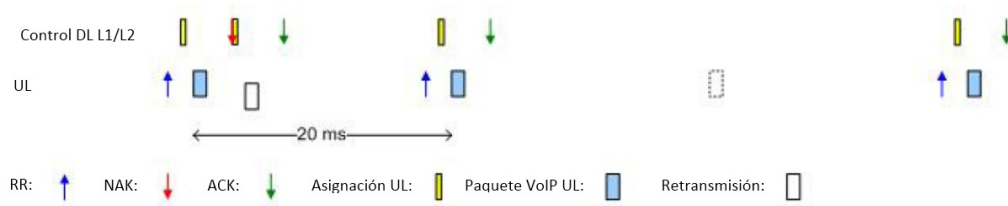


Figura 3.1: Asignación dinámica de usuarios VoIP.

característica la calendarización dinámica (peticiones y asignaciones de señalización cada intervalo de tiempo de transmisión (TTI)) no es una buena opción desde el punto de vista de señalización. Es por esto que la calendarización persistente se estudia como una opción para los servicios VoIP.

### 3.1. Asignación dinámica

Dado que las redes celulares actuales son redes por conmutación de paquetes donde normalmente cada paquete es calendarizado por control de señalización es natural aplicar la calendarización dinámica para paquetes VoIP. La calendarización dinámica de cada transmisión de paquete VoIP es naturalmente más flexible desde el punto de vista de la calendarización y el uso de recursos, pero también requiere mayor señalización. La calendarización completamente dinámica significa que el usuario envía una solicitud de recurso para cada paquete VoIP por separado y para cada retransmisión por separado. Con la calendarización dinámica los usuarios VoIP pueden beneficiarse de la calendarización selectiva de tiempo/frecuencia y los recursos no utilizados debido a los períodos de silencio, así como a la terminación temprana de HARQ la cual puede ser fácilmente reasignado a otros usuarios VoIP. El procedimiento de asignación dinámica se muestra en la Figura 3.1 [48]

La cantidad de señalización puede reducirse al no calendarizar las retransmisiones, sino usar el mismo (u otro recurso predefinido) para retransmisiones, sin embargo, la señalización ACK/NAK es necesaria de todos modos. Si sólo se calendarizan las primeras transmisiones, entonces HARQ tiene que ser síncrona y no adaptativa. La desventaja es que la flexibilidad de la calendarización de las retransmisiones se pierde

y resulta en agujeros en el espacio de tiempo/frecuencia cuando algunos recursos de retransmisión no se utilizan y no pueden ser reasignados.

La cantidad de señalización se puede reducir enviando la solicitud de recursos sólo al comienzo de la ráfaga de conversación. La estación base asigna el recurso de forma dinámica aproximadamente cada 20 ms. Del mismo modo, durante los períodos de silencio, la asignación se da cada 160 ms.

Los pros y contras de esta metodología pueden resumirse de la siguiente manera.

Pros:

- Calendarización flexible de VoIP y otros usuarios, especialmente si todas las transmisiones son calendarizadas dinámicamente
- Libertad de calendarización, ninguna división predeterminada de recursos entre usuarios de VoIP y otros usuarios.
- Posibilidad de calendarización selectiva en frecuencia y tiempo (puede requerir más señalización de piloto por el sondeo de canal).
- Posible adaptación del enlace rápido y lento.
- El HARQ asíncrono es posible y los recursos de retransmisión pueden asignarse libremente a usuarios VoIP o no-VoIP.
- Petición de recurso de enlace ascendente sólo al comienzo de la ráfaga de conversación lo cual ahorra recursos de enlace ascendente.

Contras:

- Requiere mucha señalización.
- Los recursos de retransmisión HARQ síncronos y no adaptativos son difíciles de asignar a otros usuarios.
- La asignación del tamaño del paquete tiene que hacerse para el tamaño máximo del paquete VoIP, es decir, hay relleno (con la trama SID hasta 30 bytes).

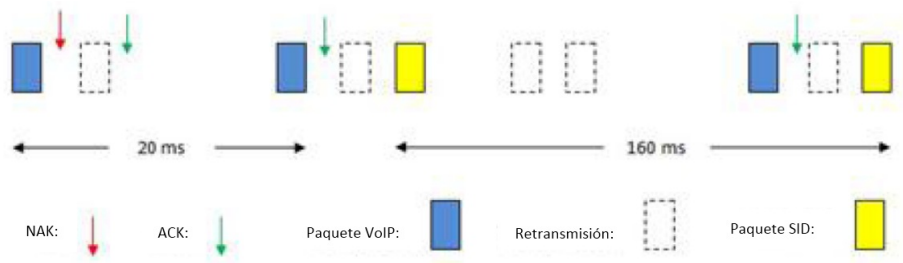


Figura 3.2: Asignación persistente de usuarios VoIP.

## 3.2. Asignación persistente

Muchas empresas han propuesto algún tipo de asignación persistente para los usuarios VoIP. El principal objetivo es reducir la cantidad de control de señalización o incluso deshacerse de ella completamente. La asignación persistente debe incluir también los recursos necesarios para las retransmisiones HARQ. En la asignación persistente puede ser que HARQ ACK/NAK no se envían, sino que cada paquete se envía un número fijo de veces.

En promedio, los usuarios VoIP están la mitad del tiempo en silencio y la otra mitad en estado activo. Si un usuario está activo, un paquete de voz se genera cada 20 ms mientras que en períodos inactivos un paquete SID (Silent Insertion Descriptor) se genera cada 160 ms. Cuando un usuario nace en el sistema y tiene información por transmitir, envía una solicitud de recursos a la estación base a través de un canal compartido. La estación base recibe el mensaje y asigna recursos al usuario de acuerdo con sus condiciones de canal, esta asignación se conserva durante la transmisión de la ráfaga de paquete. Si el usuario cambia a silencio, se informa a la estación base, que a su vez asigna recursos para los paquetes SID cada 160 ms. El procedimiento de asignación persistente se muestra en la Figura 3.2, donde el ajuste de transmisión se realiza mediante mensajes NAK/ACK [48].

Los pros y contras de esta metodología pueden resumirse de la siguiente manera.

Pros:

- Reducción de la sobrecarga de señalización.
- Simple.

Contras:

- Recursos desperdiciados:
  - o Los recursos de retransmisión HARQ no utilizados no se pueden usar por otros usuarios.
  - o Los recursos no utilizados durante los periodos en silencio no se asignan a otros usuarios.
- Se necesita un detector en el nodo B, de lo contrario el nodo B intenta decodificar y combinar las transmisiones de los recursos preasignados, incluso cuando no se envía nada.
- En la asignación persistente del tamaño del paquete tiene que ser hecha para el tamaño máximo del paquete VoIP, es decir, hay relleno (con la trama SID hasta 30 bytes).
- La máxima capacidad VoIP está limitada ya que los recursos no utilizados no pueden ser asignados a otros usuarios.
- Requiere HARQ no adaptativo.

El UE genera un paquete VoIP cada 20ms durante la ráfaga de voz y un paquete SID cada 160ms durante el período de silencio. Dado que no hay fluctuaciones en el lado de enlace ascendente, la generación de paquetes no experimentaría el retraso. El nodo B asigna diferentes recursos dependiendo del estado del UE de la siguiente manera [48]:

- Cuando el UE comienza a generar datos VoIP:
  - o El UE envía la solicitud de calendarización en el enlace ascendente.



- o El eNB asigna el recurso de frecuencia a este UE con calendarización persistente y asigna un recurso considerando el número medio de retransmisiones o el número de transmisiones. El tiempo de retransmisión debe ser preconfigurado.
  - o Después de recibir la asignación con calendarización persistente, el UE asume que el recurso se reserva para sí mismo cada 20ms y, por lo tanto, el UE envía paquete de VoIP con el recurso calendarizado.
- Cuando el UE pasa al período de silencio:
- o El UE envía un paquete SID antes de entrar en el período de silencio con el recurso asignado.
  - o El eNB libera el recurso y asigna el recurso para un paquete SID si el eNB recibe un paquete SID.
  - o Después de recibir la asignación con señalización persistente, el UE asume que el recurso está reservado para sí mismo cada 160 ms y, por tanto, el UE envía un paquete SID con el recurso calendarizado.

### 3.3. Número de bloques de recursos para distintos esquemas de modulación y codificación

Las tasas de arriba en el estado activo se pueden utilizar para calcular el número de peticiones de asignación persistente de recursos para el tráfico VoIP. En general, se espera que sólo una petición de recurso y/o una concesión de calendarización se requiera para la asignación persistente de los recursos cuando un usuario VoIP se mueve del estado inactivo al estado en activo.

Como ejemplo, tomamos los requisitos del sistema LTE, el cual se basa en OFDMA. Los parámetros para el tráfico VoIP se muestran en la Tabla 3.1 [4]. A una tasa de fuente de voz de  $12,2kb/s$ , una trama de voz generada cada 20 ms consiste de 244 bits ( $12,2kb/s * 20ms$ ). El encabezado total del protocolo por cada trama de voz incluye 10 bits de pre-encabezado de Protocolo de Transporte en Tiempo real (RTP,

Tabla 3.1: Parámetros del modelo de tráfico VoIP.

Parámetro	Valor
Codec de voz	RTP AMR 12.2, Tasa 12.2 Kb/s
Duración de la trama	20 ms
Carga SID	Paquete SID cada 160 ms Durante silencio 15 bytes (5 bytes+encabezado)
Encabezado de protocolo con compresión de encabezado	10 bit+relleno (RTP pre-encabezado) 4 byte (RTP/UDP/IP) 2 byte (RLC/seguridad) 16 bits (CRC)
Carga total de voz en la interfaz de aire	40 bytes

por sus siglas en inglés, Real-time Transport Protocol) y 2 bits de relleno, lo que da como resultado un total de 256 bits (32 bytes). Además, se adjunta al paquete un encabezado RTP comprimido /Protocolo de Datagramas de Usuario (UDP, por sus siglas en inglés, User Datagram Protocol) /Protocolo de Internet (IP, por sus siglas en inglés, Internet Protocol), que consta de 4 bytes, esto da un tamaño total de 36 bytes. Con 2 bytes para encabezados de Control de Enlace de Radio (RLC, por sus siglas en inglés, Radio Link Control) y encabezado de seguridad, y 16 bits (2 bytes) de Verificación de Redundancia Cíclica (CRC, por sus siglas en inglés, Cyclic Redundancy Check), el tamaño total de carga útil de VoIP transmitido a través de la interfaz de aire es 40 bytes. Un paquete SID consiste en un total de 15 bytes y se transmite cada 160 ms (o el equivalente a 8 tramas de voz) durante los períodos de silencio.

En LTE, un bloque de recurso (RB, por sus siglas en inglés, Resource Block) tiene dimensiones de subportadoras por símbolos. Doce subportadoras consecutivas en el dominio de la frecuencia y seis o siete símbolos en el dominio del tiempo para cada RB. La configuración del ancho de banda de transmisión es el número máximo de RB para un ancho de banda de un canal en particular. El ancho de banda máximo

Tabla 3.2: Configuración de ancho de banda de transmisión.

Ancho de banda de canal (MHz)	Número máximo de bloques de recurso	Ancho de banda máximo ocupado (MHz)
1.4	6	1.08
3	15	2.7
5	25	4.5
10	50	9.0
15	75	13.5
20	100	18

ocupado es el número de RBs multiplicado por 180 kHz, dejando el resto del ancho de banda del canal para bandas de guarda en los dos bordes. El sistema LTE admite un conjunto de seis anchos de banda de canales, como se muestra en la Tabla 3.2 [4].

El número de símbolos depende del prefijo cíclico (CP, por sus siglas en inglés, Cyclic Prefix) en uso. Cuando se utiliza un CP normal, el RB contiene siete símbolos. Cuando se utiliza un CP extendido, el RB contiene seis símbolos. Un retardo de dispersión que excede la longitud del CP normal debe utilizar un CP extendido. El ancho de banda de transmisión es el número de RB activo en una transmisión. A medida que aumenta el ancho de banda, el número de bloques de recursos aumenta.

Observamos que el ancho de banda máximo ocupado es del 90 % para 3-20 MHz. Para un ancho de banda de canal de 1.4 MHz, el ancho de banda de transmisión es sólo el 77 % del ancho de banda total del canal. Por lo tanto, para LTE en el canal de 1.4 MHz es menos eficiente, espectralmente hablando, que el canal de 3 MHz y anchos de banda de canal mayores. La configuración máxima de la banda ocupada en MHz se da en la ecuación (3.1) [4].

$$BW_{Max-Occupied} = \left( \frac{N_{RB} \cdot N_{SC}^{RB} \cdot \Delta f}{1000} \right) \quad (3.1)$$

donde  $\Delta f = 15kHz$  es el espaciamiento por subportadora.  $N_{RB}$  es el número de

Tabla 3.3: Eficiencia relativa de distintos anchos de banda.

Ancho de banda (MHz)	1.4	3	5	10	20
Bloques de recurso	6	15	25	50	100
Encabezado de banda de guarda	23 %	10 %	10 %	10 %	10 %
Encabezado BCH	2.9 %	1.1 %	0.7 %	0.3 %	0.2 %
Encabezado SCH	2.5 %	1.0 %	0.6 %	0.3 %	0.2 %

Tabla 3.4: Número de RB y MCS para tráfico VoIP.

MCS en activo	MCS en silencio	$N_{RB}$
BPSK 1/2		12
QPSK 1/2		9
QPSK 3/4		8
16QAM 1/2		6
16QAM 3/4		5
64QAM 2/3		4
64QAM 3/4	BPSK 1/2	1
QPSK 3/4 (MIMO 2X2)		2

bloques de recurso.  $N_{SC}^{RB} = 12$  es el número de subportadoras en un bloque de recurso.

Los cálculos de la eficiencia relativa de ancho de bandas se muestran en la Tabla 3.3 [4]. La eficiencia real del sistema depende del MCS utilizado, basado en la SIR. De acuerdo con los resultados de la Tabla 3.3, la eficiencia es similar para los anchos de banda de 5 a 20 MHz y existe una diferencia de 15 % para 3 MHz. La eficiencia para el ancho de banda de 1.4 MHz es aproximadamente 35-40 % menor que la de 10 MHz. Teniendo en cuenta la información de las Tablas 3.1, 3.2 y 3.3, y los procedimientos de [49], obtenemos los resultados mostrados en la Tabla 3.4, para un ancho de banda de 5 MHz.

La asignación de recursos de la trama en la interfaz de aire para cada usuario se realiza mediante la calendarización. Los  $N_{RB}$  se asignan en el ancho de banda

y se utiliza una subtrama de 1 ms para el tiempo de transmisión. Un elemento de recurso (RE, por sus siglas en ingles, Resource Element) se define como un símbolo OFDM en subportadora. Cada subtrama contiene REs piloto, Los REs de canal de control de enlace descendente físico (PDCCH) y los REs de canal compartido de enlace descendente físico (PDSCH).

Los REs del PDSCH se utilizan para la transmisión de datos del usuario. Los REs dedicados a PDCCH, piloto y símbolos de sincronización se consideran de encabezado de trama y representan un porcentaje importante de la capacidad de la trama, especialmente en los anchos de banda inferiores.

El número de bits de canal ( $N_{ch}$ ) necesarios para cada transmisión de datos de usuario se deriva estrictamente del orden de modulación ( $M$ ), la asignación de ancho de banda de radio calendarizado ( $N_{RB}$ ) y el número de REs por RB disponibles en la PDSCH,  $N_{RE/RB}$ , se muestra en la ecuación (3.2) [50].

$$N_{ch} = N_{RE/RB} \cdot N_{RB} \quad (3.2)$$

Para calcular el número de REs por RB por subtrama, se utiliza la ecuación (3.3).

$$N_{RES/RB} = 12(S_{OFDM} - C_{OFDM}) - N_{pilots} - RE_{BCH} - RE_{SCH} \quad (3.3)$$

donde:

- $C_{OFDM}$  = Número de símbolos OFDM dedicados para el campo de de control de PDCCH.
- $N_{pilots}$  = Número de REs dedicados para pilotos.
- $S_{OFDM}$  = Número de símbolos OFDM por subtrama.
- $RE_{BCH}$  = Número de REs dedicados para encabezado BCH.

Tabla 3.5: Número de elementos de recurso.

Ancho de banda (MHz)	1.4	3	5	10	20
$S_{OFDM}$ normal	14	14	14	14	14
Extendido	12	12	12	12	12
$C_{OFDM}$	1, 2, 3, 4	1, 2, 3	1, 2, 3	1, 2, 3	1, 2, 3
$N_{pilots}$	8	8	8	8	8
$RE_{BCH}$ normal	19	16	17	14	19
Extendido	17	14	14	12	16
$RE_{SCH}$ normal	17	14	14	14	19
Extendido	14	12	12	12	16

-  $RE_{SCH}$  = Número de REs dedicados para encabezado SCH.

Los canales de control ocupan los símbolos OFDM 1, 2 o 3 en una subtrama que se extiende por todo el ancho de banda del sistema. En sistemas de banda estrecha (menos de 10 RBs), los símbolos de control pueden aumentarse para incluir el cuarto símbolo OFDM. Como ejemplo de cálculo según los resultados de la Tabla 3.3, el encabezado del BCH para el ancho de banda de 1.4 MHz ocupa el 2.9% de los bloques de recursos totales, es decir,  $RB_{2,9\%} = (1,4MHz/180KHz) * 2,9)/100 = 0,225$ . Cuando se utiliza el CP normal, un bloque de recursos que consta de 12 portadoras en el dominio de la frecuencia y 7 símbolos en el dominio del tiempo, se trata de 84 elementos de recursos,  $RE_{BCH} = 0,225 * 84 \approx 19$ . Los resultados para diferentes anchos de banda se muestran en la Tabla 3.5.

Para el ancho de banda de 5 MHz, están disponibles 25 RB y, considerando la información de la Tabla 3.5, existen 109 combinaciones posibles para el uso máximo de RB para cada asignación.

Utilizando la misma metodología para tráfico Skype, que utiliza una velocidad mínima para una sesión de video de 128 kbps, es aproximadamente 10 veces mayor que una llamada de voz. Por lo tanto, podemos suponer que se utilizan 10 veces más

Tabla 3.6: Número de RB y MCS para tráfico skype.

MCS en activo	MCS en silencio	$N_{RB}$
BPSK 1/2		120
QPSK 1/2		90
QPSK 3/4		80
16QAM 1/2		60
16QAM 3/4		50
64QAM 2/3		40
64QAM 3/4	BPSK 1/2	10
QPSK 3/4 (MIMO 2X2)		20

recursos que los que se utilizan para el tráfico VoIP dependiendo de cada esquema de modulación y codificación como se muestra en la tabla 3.6. En este caso, se utiliza un ocho de banda de 20 MHz para la evaluación.

### 3.4. Análisis de teletráfico para usuarios persistentes.

El modelo propuesto se basa en la suposición de que tenemos  $S$  RBs disponibles en el sistema y que los usuarios nacen en el sistema en un estado activo con probabilidad  $p$  y en estado en silencio con probabilidad  $(1 - p)$ . Cuando los  $S$  recursos se están utilizando y un usuario desea enviar información, este usuario se considera bloqueado. Con este modelo podemos obtener la probabilidad de estar en cualquier estado del sistema, la probabilidad de bloqueo y la probabilidad de terminación forzada dependiendo de las tasas de arribo y muerte de los usuarios.

### 3.4.1. Diagrama de estado

Esta sección describe los estados posibles y estados válidos que los usuarios pueden tomar en el sistema. Las variables utilizadas en el análisis del modelo propuesto son:

- $n_1$ : El número de usuarios activos en el sistema.
- $n_2$ : El número de usuarios en silencio en el sistema.
- $S$ : El número de recursos disponibles en el sistema.
- $\lambda$ : Tasa promedio de arribos para ráfaga de voz. Los arribos se modelan como un proceso de Poisson.
- $\lambda_1$ : Tasa promedio de arribos en estado en silencio =  $(1 - p)\lambda$ .
- $\lambda_2$ : Tasa promedio de arribos en estado activo =  $p\lambda$ .
- $\lambda_3$ : Tasa de cambio de estado (activo a silencio).
- $\mu_1$ : Tasa promedio de servicio de los usuarios en silencio, el tiempo en silencio se modela como una distribución exponencial.
- $\mu_2$ : Tasa promedio de servicio de los usuarios activos, el tiempo en activo se modela como una distribución exponencial.
- $\mu_3$ : Tasa de cambio de estado (silencio a activo).

La Figura 3.3 muestra el diagrama de posibles estados en el sistema. Se observa que los usuarios pueden nacer o morir en estados activo y silencio. Además, es posible tener un cambio de estado, es decir, de activo a silencio y viceversa. La condición que cada estado debe cumplir para ser un estado válido en el sistema es dado por la ecuación (3.4).



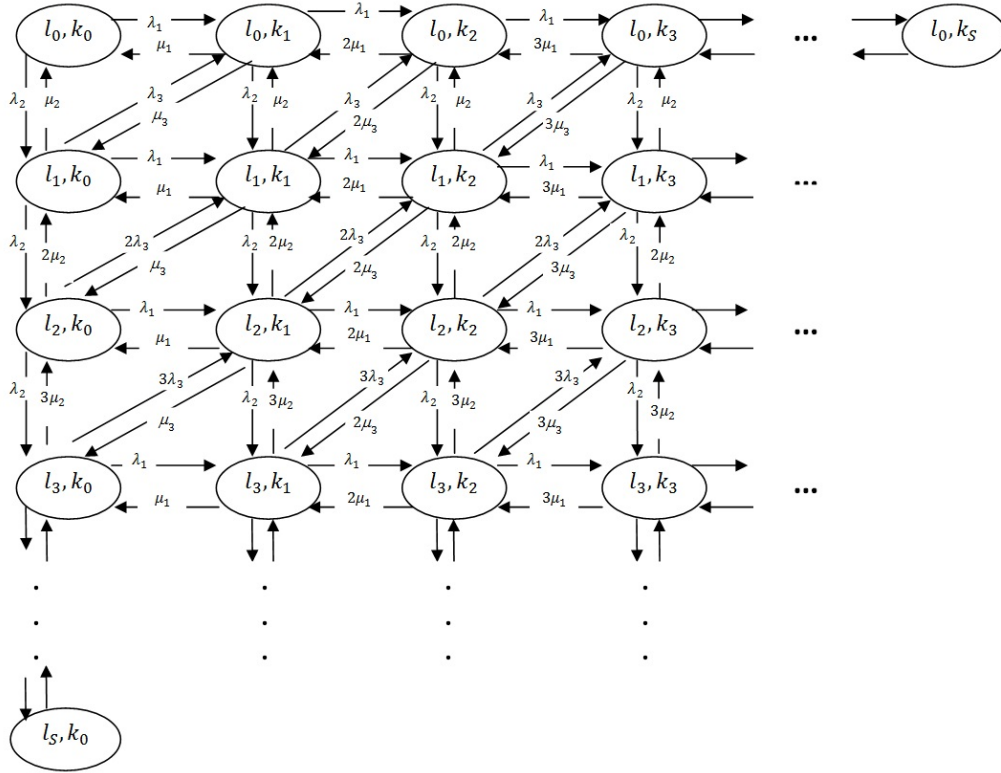


Figura 3.3: Diagrama de transición de posibles estados del sistema.

$$l_i + k_j \leq S \tag{3.4}$$

donde  $l$  es el número de bloques de recursos requeridos para el usuario activo,  $k$  es el número de bloques de recursos requeridos para el usuario en silencio,  $i$  es el número de usuarios activos y  $j$  es el número de usuarios en silencio, por lo tanto:

$$l_i = i * l; \quad k_j = j * k; \tag{3.5}$$

El modelo de nacimiento y muerte, como se muestra en la Figura 3.3, se construye para representar el cambio en el número de usuarios activos y de silencio. En este modelo, todos los usuarios se dividen en dos categorías de tráfico: usuarios activos y usuarios en silencio. Las columnas y filas de este modelo representan el cambio en el número de usuarios activos y de silencio, respectivamente.

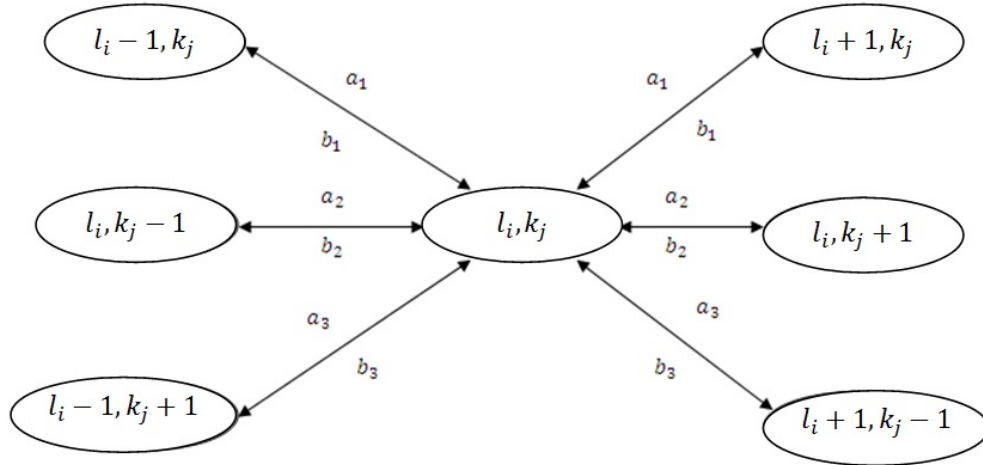


Figura 3.4: Diagrama de transición de estados.

### 3.4.2. Diagrama de transición de estados

La Figura 3.4 muestra el diagrama de transición de estados, donde  $a_1$  representa la tasa de aceptación de un usuario activo,  $a_2$  tasa de aceptación de un usuario en silencio y  $a_3$  la tasa de cambio de estado (silencio a activo).  $b_1$  representa la tasa de muerte de un usuario activo,  $b_2$  representa la tasa de muerte de un usuario en silencio y  $b_3$  es la tasa de cambio de estado (activo a silencio).

### 3.4.3. Tasas de cambio de estado y probabilidad de estado

Las tasas para los cambios de estado y la expresión para calcular la probabilidad de estado se dan en las ecuaciones (3.6-3.12).

$$a_1 = \begin{cases} \lambda_2, & \text{for } l_i, k_j \geq 0, l_i + k_j < S \\ 0, & \text{otro caso} \end{cases} \quad (3.6)$$

$$b_1 = \begin{cases} n_1\mu_2, & \text{for } n_1 > 0 \\ 0, & \text{otro caso} \end{cases} \quad (3.7)$$

$$a_2 = \begin{cases} \lambda_1, & \text{for } l_i, k_j \geq 0, l_i + k_j < S \\ 0, & \text{otro caso} \end{cases} \quad (3.8)$$

$$b_2 = \begin{cases} n_2\mu_1, & \text{for } n_2 > 0 \\ 0, & \text{otro caso} \end{cases} \quad (3.9)$$

$$a_3 = \begin{cases} n_2\mu_3, & \text{for } n_2 > 0, l_1 + k_j < S \\ 0, & \text{otro caso} \end{cases} \quad (3.10)$$

$$b_3 = \begin{cases} n_1\lambda_3, & \text{for } n_1 > 0, l_i + k_j < S \\ 0, & \text{otro caso} \end{cases} \quad (3.11)$$

En la ecuación (3.12), se muestra la expresión para calcular la probabilidad de estado del sistema analizado.

$$P(l_i, k_j) = [a_1(l_i - 1, k_j) \cdot P(l_i - 1, k_j) + b_1(l_i + 1, k_j) \cdot P(l_i + 1, k_j) + a_2(l_i, k_j - 1) \cdot P(l_i, k_j - 1) + b_2(l_i, k_j + 1) \cdot P(l_i, k_j + 1) + a_3(l_i - 1, k_j + 1) \cdot P(l_i - 1, k_j + 1) + b_3(l_i + 1, k_j - 1) \cdot P(l_i + 1, k_j - 1)] / [a_1(l_i, k_j) + b_1(l_i, k_j) + a_2(l_i, k_j) + b_2(l_i, k_j) + a_3(l_i, k_j) + b_3(l_i, k_j)] \quad (3.12)$$

Calculamos la probabilidad de terminación forzada correspondiente a un usuario en silencio que cambia al estado activo, debido a que los usuarios en silencio sólo necesitan un RB para enviar el mensaje SID, al cambiar de estado puede no existir recursos disponibles para satisfacer al usuario dependiendo del anillo donde el usuario ha nacido.



---

## Capítulo 4

# Análisis de SIR y estadísticas de orden

En este capítulo se desarrolla un modelo matemático para el análisis de la calendarización persistente basado en la transmisión de tráfico VoIP. Obtenemos las estadísticas de orden de las SIR's considerando la interferencia de co-canal. Las estadísticas de orden se utilizan para analizar un esquema de retroalimentación de la mejor banda, en el que cada usuario reporta el índice de sus mejores bandas a la estación base. Se realiza una combinación de procesos de nacimiento y muerte y análisis de interferencia para obtener la tasa de transmisión media en el sistema, y los esquemas de modulación y codificación (MCS) que debe utilizar el usuario de acuerdo con el calendarizador de Máximo SIR basado en las estadísticas de orden.

En nuestro esquema de selección de tasas propuesto, combinamos dos análisis que en otros estudios se realizan independientemente: SIR y Análisis de Teletráfico. Consideramos dos factores importantes que no se consideraron en otros trabajos relacionados: la dependencia de variables aleatorias de las señales de interferencia con las estadísticas de orden y el costo del uso de recursos para señalización. Los resultados numéricos muestran que el esquema propuesto de asignación de recursos con estadísticas de orden puede mejorar eficientemente la utilización de recursos y reducir la probabilidad de terminación forzada en comparación con los esquemas convencionales.

## 4.1. Cálculo de la SIR

Se considera un sistema de 19 celdas, como se muestra en la Figura 4.1. Se supone que la potencia de transmisión es la misma para todas las celdas con una densidad de usuarios uniforme. La posición se puede caracterizar en términos de coordenadas polares  $(d, \varphi)$  con distribución conjunta  $f_{d,\varphi}(d, \varphi)$ . La distribución marginal de la distancia  $f_{den}(d)$  se puede obtener calculando la integral sobre  $f_{d,\varphi}(d, \varphi)$  como se muestra en la ecuación (4.1), donde  $d$  es la distancia hasta la estación base en servicio y la fase  $\varphi$  se puede suponer que se distribuye uniformemente en  $[-\pi, \pi]$ . En este esquema, todas las celdas pueden transmitir al mismo tiempo. Se genera interferencia entre celdas vecinas. La banda de frecuencia de los S recursos se divide en C bloques o chunks (bandas), cada uno de los cuales contiene  $L = S/C$  recursos. La ubicación de la estación móvil (MS, por sus siglas en inglés, Mobile Station) con respecto a cada estación base (BS, por sus siglas en inglés, Base Station) es conocida. Con esta localización se puede determinar la distancia entre la MS y la BS y la dirección de la Línea de Vista (LOS, por sus siglas en inglés, Line Of Sight) con respecto de la BS y MS [51].

$$f_{den}(d) = \int_0^{2\pi} f_{d,\varphi}(d, \varphi) d\varphi \quad (4.1)$$

Por simplicidad no consideramos el ruido de fondo. La ecuación para calcular la relación señal a interferencia (SIR, por sus siglas en inglés, Signal-to-Interference Ratio) de la  $i$ -ésima banda se muestra en la ecuación (4.2) [52].

$$SIR_i = \frac{P_t h_{z,i} z^{-\alpha}}{P_t \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha}} = \frac{h_{z,i} z^{-\alpha}}{\sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha}} \quad (4.2)$$

donde  $i = 1, \dots, C$ ,  $P_t$  es la potencia de transmisión y el desvanecimiento a pequeña escala (desvanecimientos rápidos) entre el móvil y una estación base interferente  $y \in \phi$

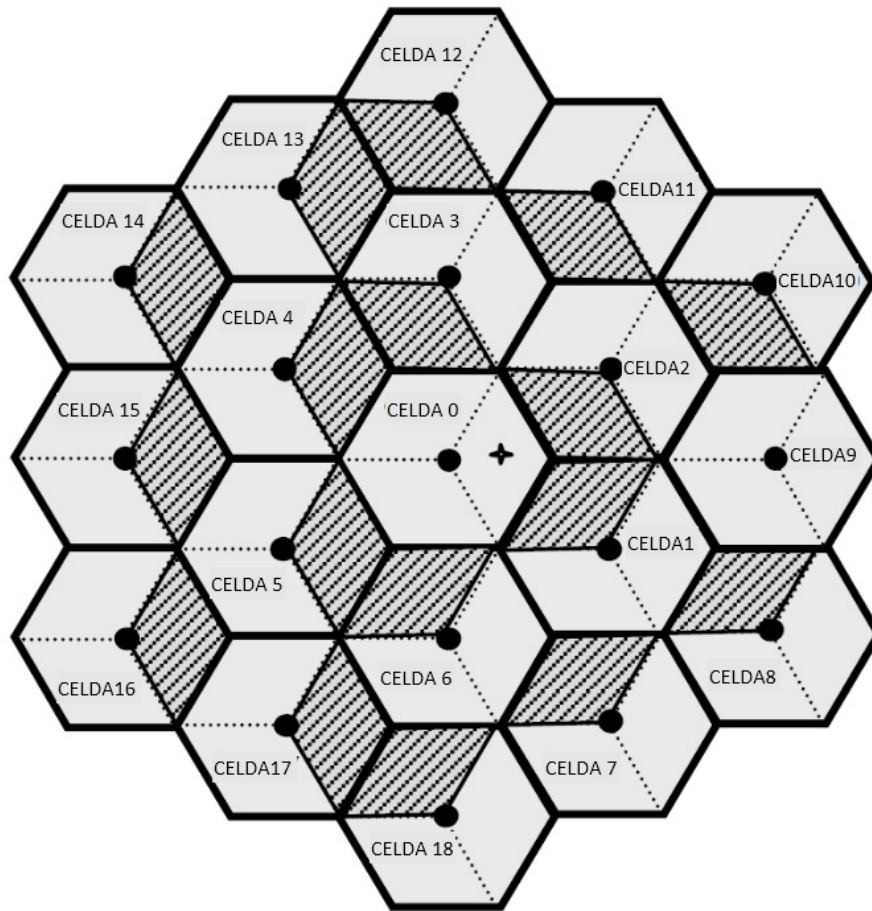


Figura 4.1: Sistema de 19 celdas.

en la  $i$  ésima banda esta dado por  $h_{y,i}$ . Es decir,  $\sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha}$ , representa la suma de las señales interferentes de las 19 celdas.  $h_{z,i}$  representa la señal de la celda de interés.

Las variables aleatorias interferentes se consideran como desvanecimiento con distribución Rayleigh, lo que implica que  $h_{y,i}$  se distribuyen exponencialmente. La pérdida por trayectoria esta dado por  $z^{-\alpha}$ , donde  $\alpha$  es el exponente de pérdida por trayectoria, y  $y^{-\alpha}$  es la pérdida de trayectoria entre el usuario de interés y la señal interferente.

Para calcular la tasa de transmisión se usa la ecuación de Shannon, la cual se

muestra en la ecuación (4.3), donde  $W$  es el ancho de banda.

$$S(\text{bit/s}) = W \log_2(1 + SIR_i) \quad (4.3)$$

## 4.2. Esquemas de asignación de recursos.

Cada usuario móvil podrá estimar la SIR en cada banda usando pilotos. Por lo tanto, un usuario móvil puede ordenar la calidad de sus bandas basadas en la SIR calculado y reportar la orden a la BS. Suponemos que cada usuario móvil reporta el orden de las  $C$  mejores bandas. La BS recibe este informe de ordenamiento de bandas de todos sus usuarios y asigna las bandas de la siguiente manera:

1. La BS primero considera las mejores bandas de todos los usuarios. Para cada banda  $B_i$ ,  $i = 1, \dots, C$ , la BS enlista todos los usuarios que reportaron  $B_i$  como su mejor banda. Si sólo un usuario reportó  $B_i$  como su mejor banda, la BS asigna la banda a ese usuario. Si varios usuarios reportaron  $B_i$  como su mejor banda, la BS asigna la banda siguiendo dos alternativas:
  - \* El usuario que presente la máxima SIR (Asignación de Máximo SIR).
  - \* La BS selecciona aleatoriamente uno de estos usuarios y asigna la banda a ese usuario (Asignación Aleatoria).
2. La BS considera entonces, solamente las bandas no asignadas y las segundas mejores bandas de los usuarios restantes (usuarios cuya mejor banda no fue asignada) y repite el procedimiento anterior con la segunda mejor banda.
3. Esta asignación continúa hasta que cada usuario tiene una banda asignada. Tenga en cuenta que en la  $i$ -ésima iteración de la BS se centra en la  $i$ -ésima mejor banda.



### 4.3. Análisis de estadísticas de orden.

Las estadísticas de orden de la SIR ( $SIR_{i:C}$ ) son necesarias para calcular la probabilidad de transmisión de un usuario. Las  $SIR_{i:C}$  no son independientes, por lo tanto, la estadísticas de orden dependientes se usan, entonces y se calcula la distribución condicional de  $SIR_{i:C}$ , dado  $SIR_{j:C}$  para  $j > i$ .

Sea  $SIR_{i:C}$  la variable aleatoria correspondiente a la  $i$ -ésima observación más grande de las  $C$  variables aleatorias originales, tal que  $SIR_{1:C} \leq SIR_{2:C} \leq \dots \leq SIR_{C:C}$ . La  $SIR_{i:C}$  también se denomina la  $i$ -ésima estadística de orden. Para calcular la función de distribución condicional utilizamos el teorema 1.

**Teorema 1.** Sea  $SIR_{1:C} \leq SIR_{2:C} \leq \dots \leq SIR_{C:C}$ , la función de distribución condicional de la  $i$ -ésima estadística de orden de  $SIR_{i:C}$  dado que la  $j$ -ésima estadística de orden  $SIR_{j:C}$  es igual a  $SIR_j$ , donde  $i < j$ , es igual a la probabilidad  $P = (SIR_{r:C} \leq \theta)$  el cual se denota por  $F_{r:C}(SIR)$ , está dado por la ecuación (4.4).

$$\begin{aligned}
 F_{r:C}(SIR) &= \sum_{j=r}^C (-1)^{(j-r)} \binom{j-1}{r-1} \cdot \sum_{1 \leq \dots \leq C} P_r \{ \max(SIR_{i:1}, \dots, SIR_{i:j}) \leq \theta \} = \\
 &= \sum_{j=r}^C (-1)^{(j-r)} \binom{j-1}{r-1} \sum_{1 \leq \dots \leq C} \int_0^{\infty} \frac{(j-1)!}{(i-1)!(j-i-1)!} \cdot \left[ \frac{1 - e^{-\theta z^\alpha \left( \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha} \right)}}{1 - e^{-\theta z^\alpha \left( \sum_{y \in \phi\{z\}} h_{y,j} y^{-\alpha} \right)}} \right]^{(i-1)} \cdot \\
 &= \left[ \frac{e^{-\theta z^\alpha \left( \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha} \right)} - e^{-\theta z^\alpha \left( \sum_{y \in \phi\{z\}} h_{y,j} y^{-\alpha} \right)}}{1 - e^{-\theta z^\alpha \left( \sum_{y \in \phi\{z\}} h_{y,j} y^{-\alpha} \right)}} \right]^{(j-i-1)} \cdot \\
 &= \frac{z^\alpha \left( \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha} \right) e^{-\theta z^\alpha \left( \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha} \right)}}{1 - e^{-\theta z^\alpha \left( \sum_{y \in \phi\{z\}} h_{y,j} y^{-\alpha} \right)}} d\theta \quad (4.4)
 \end{aligned}$$

Comprobación: Véase el Apéndice A.

La probabilidad de que al usuario se le asigne la  $i$ -ésima mejor banda se puede calcular viendo como ejemplo la distribución de los contenedores vacíos dados en las ecuaciones (4.5 y 4.6) [53].

$$P\{\mu_0(n, C) = k\} = C_C^k \left(1 - \frac{k}{C}\right)^n P\{\mu_0(n, C - k) = 0\} \quad (4.5)$$

$$P\{\mu_0(n, C) = 0\} = \sum_{l=0}^C C_C^l (-1)^l \left(1 - \frac{l}{C}\right)^n \quad (4.6)$$

donde  $C_C^k = \binom{C}{k}$ .

De la misma manera que las estadísticas de orden, se puede obtener la probabilidad de utilizar una tasa de servicio. Utilizando la expresión  $P_i * \lambda_i$  para una tasa de transmisión  $T_i$ , dependiendo de la  $SIR_i$ . Esto se muestra en la Figura 4.2, con los siete posibles MCS calculado en las secciones anteriores. Donde  $P_i$  es la probabilidad de que el usuario nazca en el área  $i$  ( $i = 1, \dots, 7$ ), y asignando el MCS correspondiente dependiendo del cálculo de la SIR. Para calcular esta probabilidad, también se asume una distribución de usuarios uniforme.

Usando la función de densidad de probabilidad de las estadísticas de orden en (4.4) y la Tabla (4.1), las nuevas probabilidades,  $P'_i$  se pueden obtener por la ecuación (4.7), estas probabilidades se muestran en la Tabla (4.2), que modifica las áreas de cobertura de la celda como se muestra en la Figura 4.3. Se observa que las áreas de mayor banda aumenta, ya que las estadísticas de orden modifican la probabilidad de transmitir en las primeras mejores bandas. Tenga en cuenta que se utiliza una densidad de usuarios uniforme, pero se puede utilizar cualquier otra distribución.

$$P'_i = P_i \cdot \int_{SIR_{min}}^{SIR_{max}} f_{r:C}(SIR_i) d\theta \quad (4.7)$$

Los usuarios pueden nacer en el estado activo, con tasa  $\lambda_i$  y en diferentes anillos

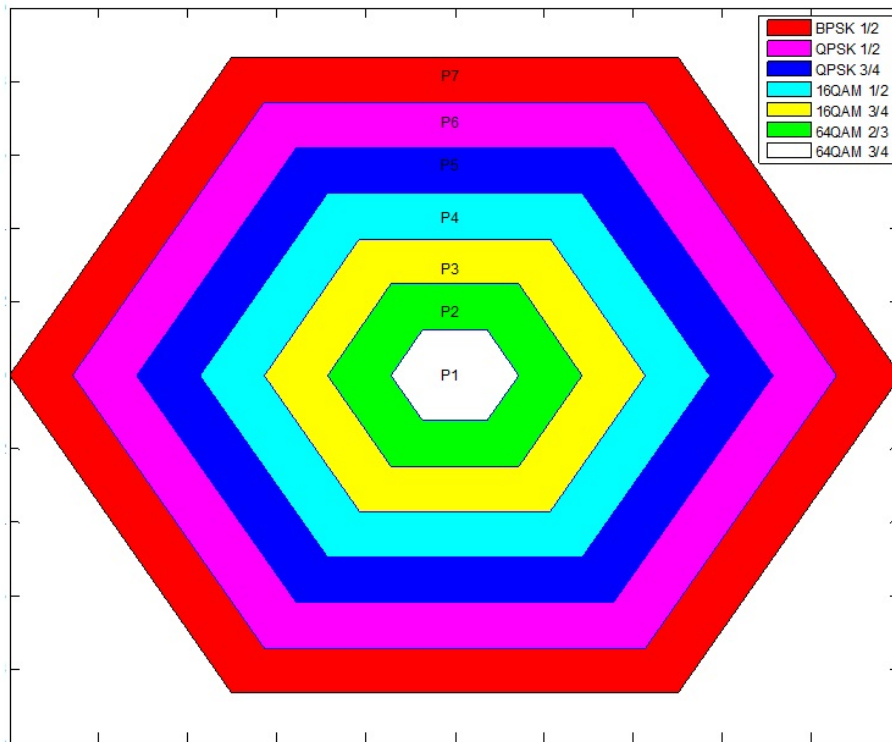


Figura 4.2: Probabilidad de áreas para cada MCS.

Tabla 4.1: Modulación y codificación vs SIR.

(Modulación, codificación)	$SIR_{min}$	$SIR_{max}$
(BPSK, 1/2)		4.3 dB
(QPSK, 1/2)	4.3 dB	8 dB
(QPSK, 3/4)	8 dB	11 dB
(16QAM, 1/2)	11 dB	14 dB
(16QAM, 3/4)	14 dB	19.4 dB
(64QAM, 2/3)	19.4 dB	23 dB
(64QAM, 3/4)	23 dB	

usando las probabilidades calculadas y mostrados en la Tabla (4.2), las tasas están dados por la ecuación (4.8).

Tabla 4.2: Probabilidades para cada MCS considerando estadísticas de orden.

$P'_1$	0.25	64QAM, 3/4
$P'_2$	0.2065	64QAM, 2/3
$P'_3$	0.1766	16QAM, 3/4
$P'_4$	0.1413	16QAM, 1/2
$P'_5$	0.1168	QPSK, 3/4
$P'_6$	0.0870	QPSK, 1/2
$P'_7$	0.0217	BPSK, 1/2

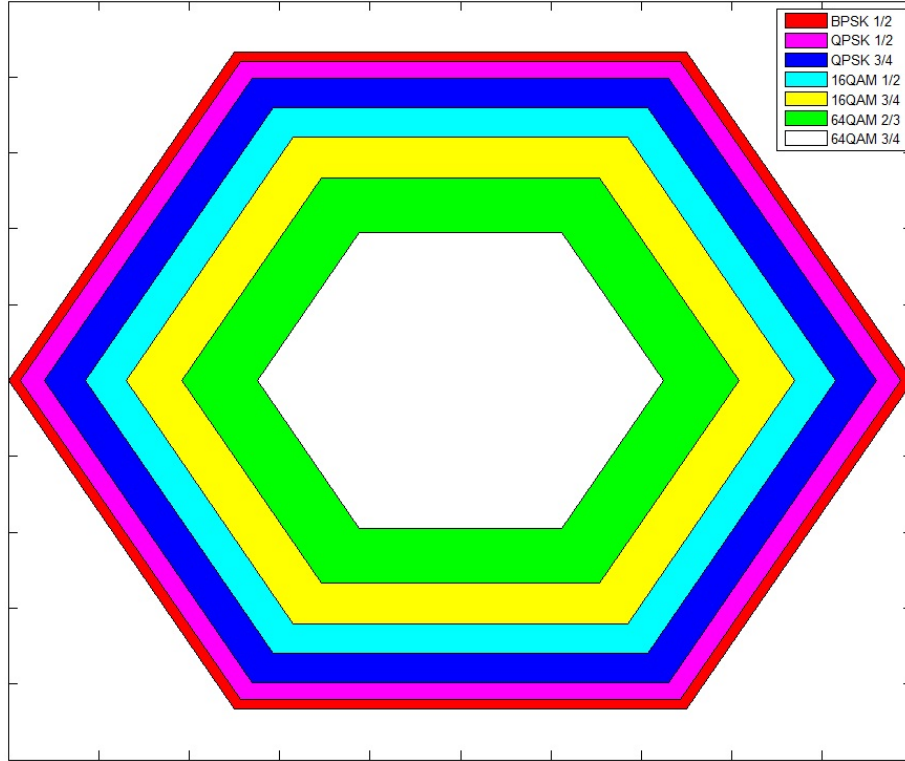


Figura 4.3: Probabilidad de áreas considerando estadísticas de orden.

$$\lambda_{nacer-estado-activo} = \begin{cases} P'_1 \lambda_1, & 64QAM3/4, & RB = 1 \\ P'_2 \lambda_1, & 64QAM2/3, & RB = 4 \\ P'_3 \lambda_1, & 16QAM3/4, & RB = 5 \\ P'_4 \lambda_1, & 16QAM1/2, & RB = 6 \\ P'_5 \lambda_1, & QPSK3/4, & RB = 8 \\ P'_6 \lambda_1, & QPSK1/2, & RB = 9 \\ P'_7 \lambda_1, & BPSK1/2, & RB = 12 \end{cases} \quad (4.8)$$

Tabla 4.3: Parámetro  $G_{Tun}$ .

Atracción alta	Atracción media	Atracción baja
17.08	2.87	1.15

Podemos utilizar esta misma metodología para una distribución de usuarios no uniforme, utilizando la ecuación (4.9) en vez de la ecuación (4.1) como se indica en el capítulo 2. Utilizamos un modelo de densidad de usuarios no uniforme exponencial. La definición matemática de la densidad exponencial de usuarios esta dado en (4.9).

$$f_{den}(d) = G_{Tun} \cdot G_{Exp} \cdot e^{-Ed} \quad 0 \leq d \leq \infty \quad (4.9)$$

donde  $G_{Exp}$  representa la densidad maxima en el modelo de densidad exponencial,  $E$  es el parámetro para controlar el nivel no uniforme,  $E$  refleja la pendiente de la densidad exponencial de usuarios. Si  $E$  es grande, presenta alta densidad y alta atracción cercana al centro de la celda. Si  $E$  es pequeño, se aproxima a una densidad de usuarios uniforme y baja atracción. Cuando  $E = 0$ , es exactamente igual a la densidad de usuarios uniforme.

$G_{Tun}$  representa la ganancia de ajuste del modelo de densidad de usuario no uniforme. La ganancia de ajuste  $G_{Tun}$  se puede obtener experimentalmente. La tabla 4.3 muestra un ejemplo de la ganancia  $G_{Tun}$  obtenida experimentalmente.

Usando un radio de celda  $R = 100$ ,  $G_{Exp} = 1$ ,  $E = 0,05$  y  $G_{Tun} = 17,08$  con una distribución de usuarios no uniforme, la densidad de la distancia dada por  $f_{den}(d)$  y la fase  $\varphi$  se puede suponer que se distribuye uniformemente en  $[-\pi, \pi]$  con la misma probabilidad  $p(\varphi) = 1/(2\pi)$ ,  $\varphi \in [-\pi, \pi]$ .

La Figura 4.4 muestra la distribución de usuarios en una celda con radio  $R = 100$ , tanto para una distribución uniforme como para una distribución no uniforme. Se observa que en la distribución los usuarios no uniforme se concentran cerca de la estación base.

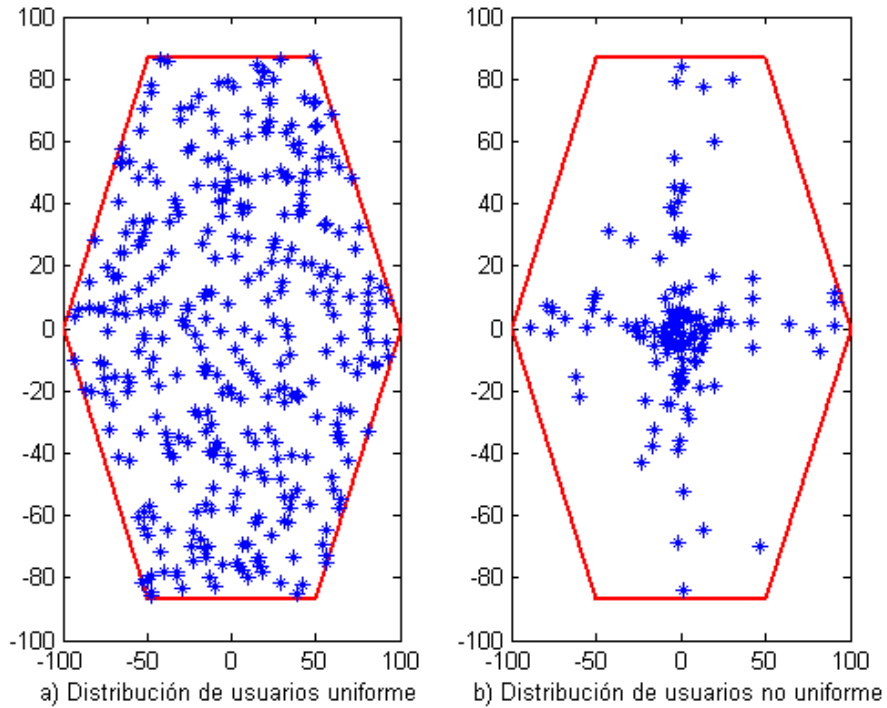


Figura 4.4: Modelo de distribución de usuarios.

Tabla 4.4: Probabilidades para cada MCS considerando estadísticas de orden.

$P'_i$	MCS	Uniforme	No uniforme
$P'_1$	64QAM, 3/4	0.25	0.274
$P'_2$	64QAM, 2/3	0.2065	0.217
$P'_3$	16QAM, 3/4	0.1766	0.1883
$P'_4$	16QAM, 1/2	0.1413	0.146
$P'_5$	QPSK, 3/4	0.1168	0.1278
$P'_6$	QPSK, 1/2	0.0870	0.0352
$P'_7$	BPSK, 1/2	0.0217	0.0116

Usando la misma metodología para una distribución de usuarios no uniforme,  $P'_i$  se puede obtener por (4.7), estas probabilidades se muestran en la tabla 4.4.

Tabla 4.5: Parámetros del sistema.

Parámetro	Valor
Diseño celular	Sistema hexagonal, sistema de 19 celdas
Ancho de banda del sistema	5 MHz
Desviación estandar de desvanecimiento	8 dB
Reuso de frecuencia	1
Radio de la celda	500 m
Exponente de pérdida de trayectoria	4 dB
Modelo de tráfico	AMR 12.2 Kbps 40 bytes por AMR 12.2 Kbps Paquete de voz VoIP 15 bytes por paquete SID

Tabla 4.6: Número de estados válidos.

Ancho de banda del canal (MHz)	Número máximo de bloques de recurso	Número de estados válidos
1.4	6	9
3	15	107
5	25	708
10	50	16887
15	75	144236
20	100	736887

## 4.4. Resultados numéricos

Esta sección presenta los resultados numéricos. Los parámetros del sistema se dan en la Tabla (4.5). Utilizando los datos de la Tabla (3.4) y la condición de estados válidos en la ecuación (3.4), se calcula el número de estados válidos para cada ancho de banda, y esto se muestra en la Tabla (4.6).

Calculamos los estados válidos cuando los bloques de recursos están completamente ocupados. Por ejemplo, para el ancho de banda de 5 MHz (25 bloques de

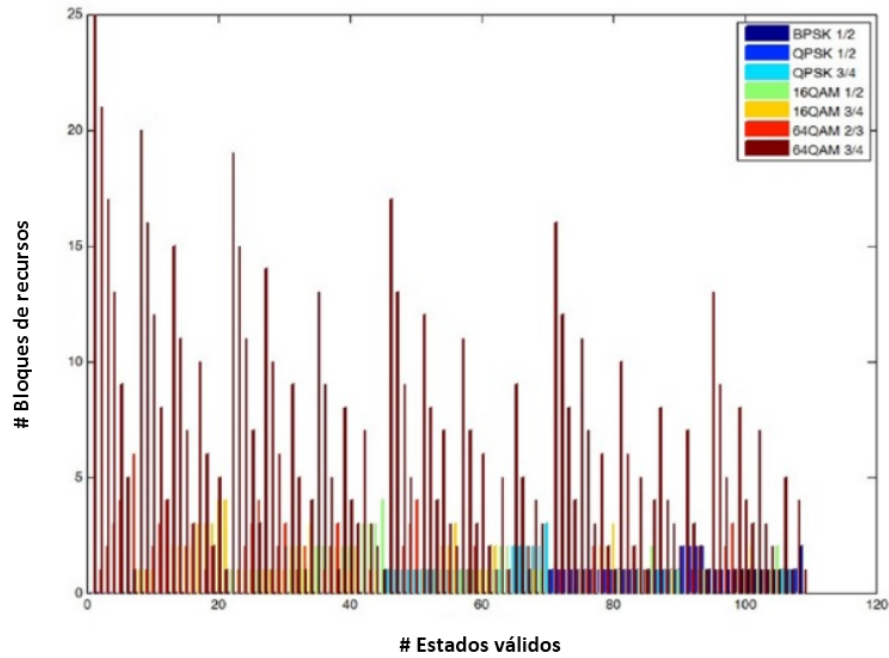


Figura 4.5: Estados válidos para 5 MHz.

recursos), hay 109 estados válidos cuando los bloques de recursos están ocupados por completo, esto se muestra en la Figura 4.5. En la Figura 4.6 podemos ver un caso específico de los 109 posibles estados válidos. Notamos que tenemos 6 usuarios con 64QAM, 3/4 (6 RBs), 1 usuario con 16QAM, 3/4 (5 RBs), 1 usuario con 16QAM, 1/2 (6 RBs) y 1 usuario con QPSK 3 / 4 (8 RBs), sumando los RBs tenemos 25 RBs. Esto es importante porque nos permite evaluar la capacidad del sistema.

Los usuarios arriban al sistema con una tasa de Poisson y permanecen en ella hasta que un paquete se transmite en  $t$  minutos (el tiempo medio de servicio, depende de la MCS mostrados en la Tabla 4.1). Usando las ecuaciones (3.6 - 3.12 y 4.8) y la condición en (3.4), podemos calcular la probabilidad de terminación forzada del sistema para diferentes valores de los recursos disponibles, con un ancho de banda de 5 MHz. La Figura 4.7 muestra la probabilidad de terminación forzada para el sistema con y sin estadísticas de orden descritas en las secciones anteriores.

La evaluación directa de la ecuación (4.5) no es factible, obtenemos  $P_1, \dots, P_C$



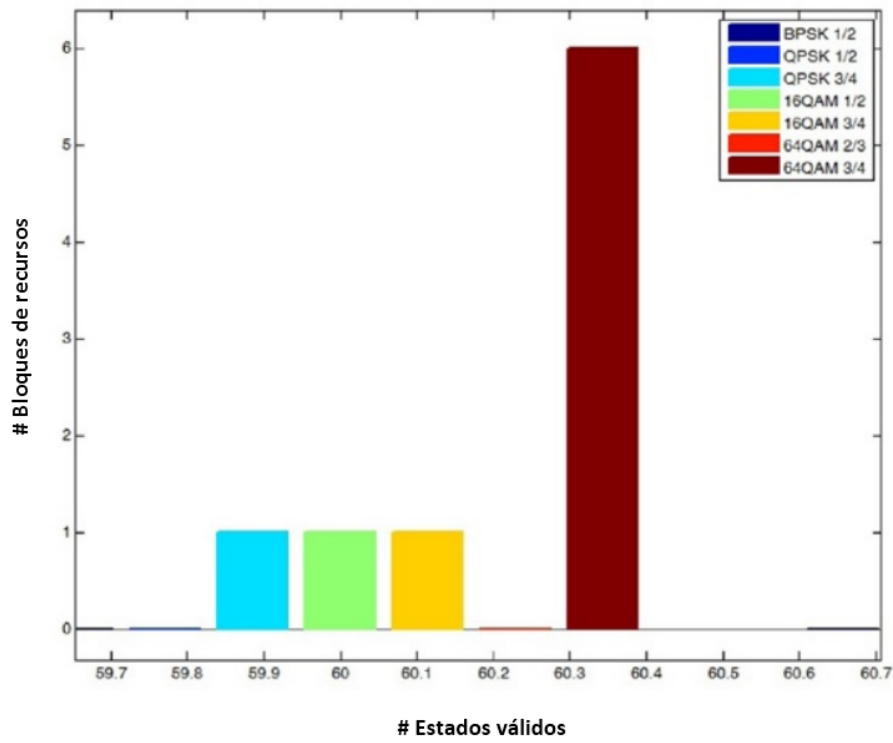


Figura 4.6: Un caso específico de estados válidos para 5 MHz.

usando simulaciones. La Figura 4.8 muestra la probabilidad de asignar la  $n$ -ésima mejor banda cuando  $C = 10$ .

Las Figuras 4.9 y 4.10 muestran los resultados de la SIR y la tasa de transmisión, respectivamente. Comparamos la asignación aleatoria y la asignación de máxima SIR y se observa que la asignación de máxima SIR tiene una ganancia significativa con respecto a la asignación aleatoria, ya que siempre busca transmitir cuando se tiene la máxima SIR.

Se observa en la Figura 4.9 que la probabilidad de terminación forzada se reduce significativamente cuando se utiliza el análisis de estadísticas de orden, porque la probabilidad de que el usuario transmita en su mejor banda es mayor. Este análisis puede servir para dimensionar redes reales.

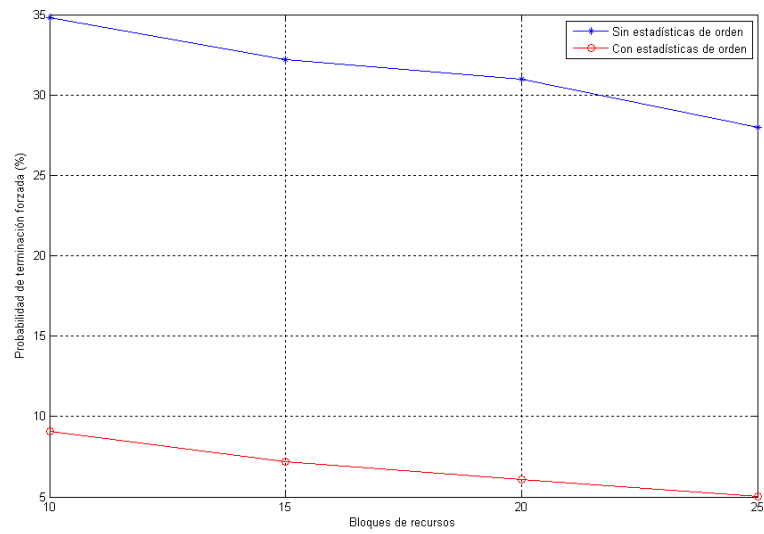


Figura 4.7: Probabilidad de terminación forzada.

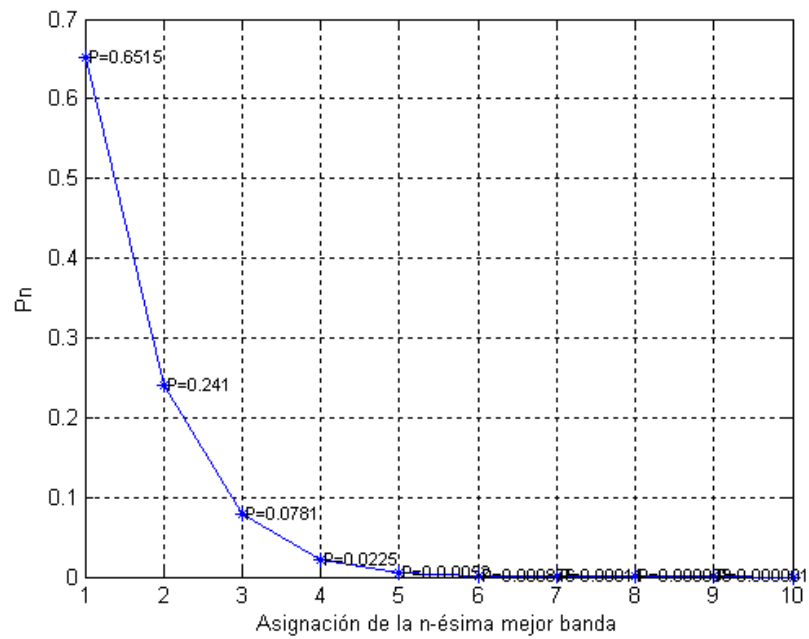


Figura 4.8: Probabilidad de asignación de la n-ésima mejor banda, C=10.

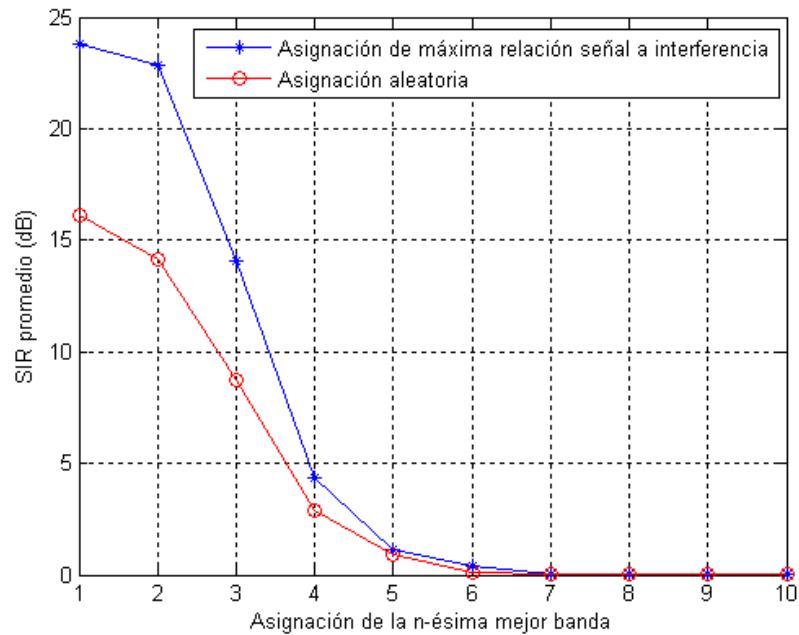


Figura 4.9: SIR vs asignación de la n-ésima mejor banda.

Ahora se desarrolla un análisis de desempeño para una distribución de usuarios no uniforme (modelo de densidad exponencial) y tráfico de Skype, utilizando la misma metodología. Skype es un servicio de VoIP proporcionado por Microsoft y actualmente es el más utilizado internacionalmente. El paquete de Skype llega cada 20 ms. El valor medio del tráfico es de 128 Kbps. Obtuvimos la probabilidad de bloqueo y la probabilidad de terminación forzada para este tipo de tráfico y distribución de usuarios. Estos se muestran en las Figuras 4.11 - 4.12.

Se observa en las Figuras 4.11 - 4.12, que la distribución no uniforme tiene un mejor rendimiento, porque los usuarios se concentran cerca de la estación base y con el análisis de las estadísticas de orden se observó que las áreas cercanas a la estación base aumentaron. En estas áreas, los usuarios utilizan menos recursos para la transmisión y transmiten a una mejor tasa.

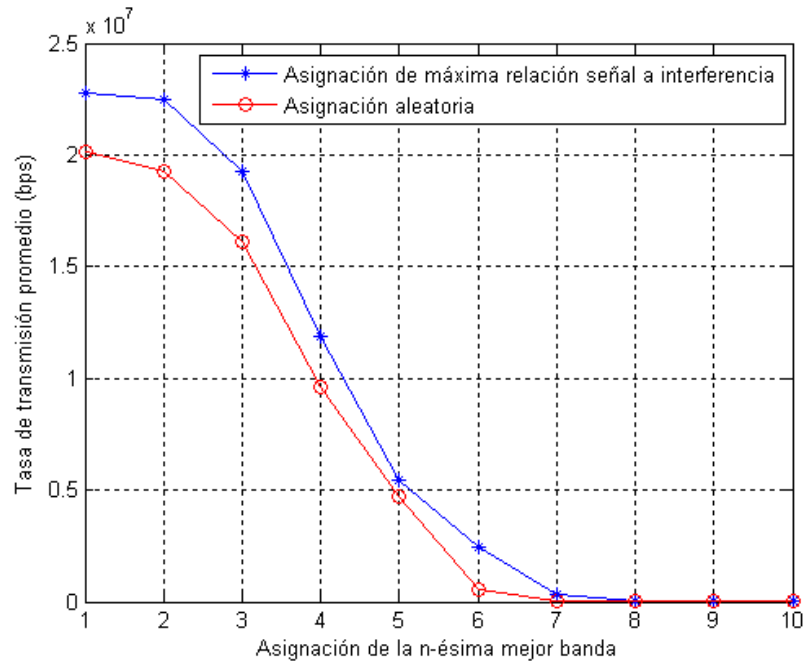


Figura 4.10: Tasa promedio de datos vs asignación de la n-ésima mejor banda.

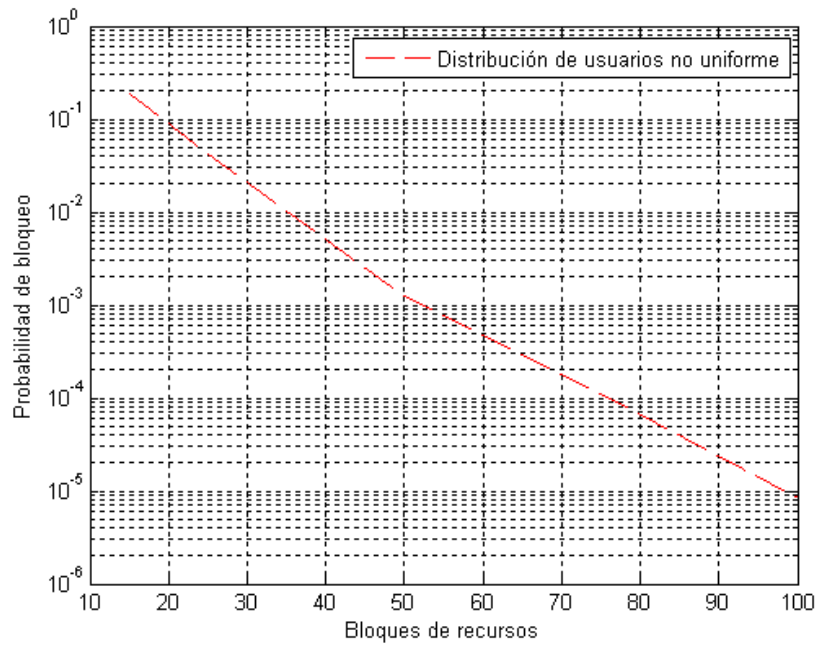


Figura 4.11: Probabilidad de bloqueo para tráfico skype.

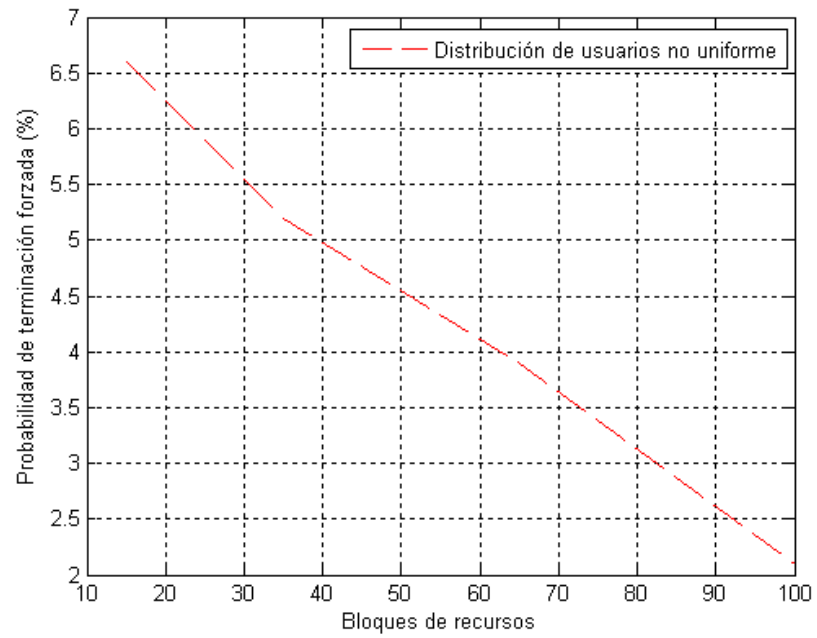


Figura 4.12: Probabilidad de terminación forzada para tráfico skype.

---

## Capítulo 5

# Esquema de acceso coordinado para sistemas celulares 5G

En la historia de las comunicaciones inalámbricas desde la primera generación (1G) a la cuarta generación (4G), el esquema de acceso múltiple ha sido la tecnología clave para distinguir los distintos sistemas celulares.

Es bien conocido que se utilizaron acceso múltiple por división de frecuencia (FDMA) para 1G, acceso múltiple por división de tiempo (TDMA) sobre todo para 2G, acceso múltiple por división de código (CDMA) para 3G y acceso múltiple por división de frecuencia ortogonal (OFDMA) para 4G, los cuales son principalmente esquemas de acceso múltiple ortogonales (OMA).

En estos esquemas de acceso múltiple convencionales, a los diferentes usuarios se les asignan los recursos ortogonales, ya sea en el dominio del tiempo, frecuencia o código con el fin de evitar o mitigar la interferencia interusuario. De esta manera, la ganancia de multiplexación se puede lograr con una complejidad razonable.

Sin embargo, el crecimiento rápido de Internet móvil impulsará a 1000 veces el aumento del tráfico de datos en 2020 para 5G. Por lo tanto, la eficiencia espectral se convierte en uno de los principales retos para manejar este tipo de tráfico de datos.

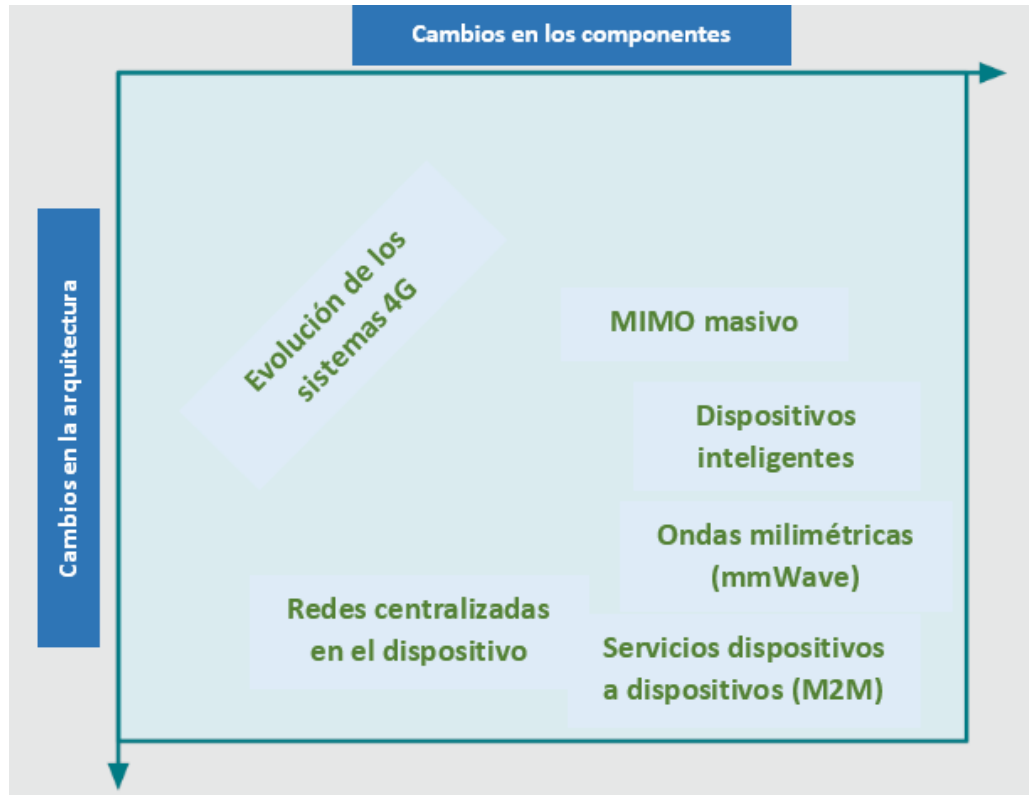


Figura 5.1: Propuestas de tendencias para 5G.

Por otra parte, debido al rápido desarrollo de Internet de las cosas (IOT), 5G tiene que soportar la conectividad masiva de los usuarios y/o dispositivos para satisfacer la demanda de baja latencia, dispositivos de bajo costo, y diversos tipos de servicios.

Para satisfacer estos requisitos, es necesario mejorar las tecnologías. Hasta ahora, se han propuesto algunos candidatos potenciales para hacer frente a los retos de 5G, tales como MIMO masivo, las comunicaciones de ondas milimétricas, la red ultra densa, y acceso múltiple no ortogonal (NOMA). Como se muestra en la Figura 5.1.

## 5.1. Tendencias para 5G

Se clasifican los impactos de las nuevas tecnologías como sigue:

- Cambios menores, tanto en el nodo y los niveles de arquitectura (por ejemplo, la introducción de libros de códigos y soporte de señalización para un mayor número de antenas). Nos referimos a estos como la evolución en el diseño.
- Cambios en el diseño de una clase de nodos de red (por ejemplo, la introducción de una nueva forma de onda). Nos referimos a estos cambios como componente.
- Cambios en la arquitectura del sistema (por ejemplo, la introducción de nuevos tipos de nodos o nuevas funciones de los ya existentes). Nos referimos a estos como los cambios de arquitectura.
- Cambios que tienen un impacto tanto a nivel de nodo como en la arquitectura. Nos referimos a estos como los cambios radicales.

Nos centramos en las tecnologías de punta (componentes, arquitectura o radicales). Creemos que las siguientes cinco tecnologías potencialmente podrían conducir a cambios tanto de arquitectura y diseño de componentes, según la clasificación de la figura 5.1.

- 1) Arquitectura centrada en el dispositivo: La arquitectura centrada en la estación base en los sistemas celulares puede cambiar en 5G. Puede ser el momento de reconsiderar los conceptos de enlace ascendente y descendente, así como canales de control y de datos, a una mejor información de las rutas de flujos con diferentes prioridades y propósitos hacia diferentes conjuntos de nodos dentro de la red.
- 2) Ondas milimétricas: Mientras el espectro en frecuencias de microondas escaso, en las ondas milimétricas es abundante. Tal espectro ha dado lugar a que los investigadores estén estudiando diferentes aspectos de la transmisión de ondas



milimétricas. Aunque lejos de ser completamente entendido, las tecnologías de ondas milimétricas ya han sido estandarizados para los servicios de corto alcance (IEEE 802.11ad) y desplegado para aplicaciones especializadas tales como backhaul de celdas pequeñas.

- 3) MIMO masivo: propone la utilización de un número elevado de antenas para multiplexar los mensajes de varios dispositivos en cada recurso de tiempo-frecuencia, centrándose en la energía radiada hacia los destinatarios al tiempo que minimiza la interferencia intra e intercelular. MIMO masivo puede requerir cambios importantes de arquitectura, sobre todo en el diseño de las estaciones base, y también puede conducir a nuevos tipos de implementaciones.
- 4) Los dispositivos inteligentes: las redes celulares 2G-3G-4G se construyeron bajo la premisa de tener un control completo en el lado de la infraestructura. Se argumenta que en los sistemas 5G se debe eliminar este supuesto de diseño y explotar la inteligencia en el lado del dispositivo dentro de las diferentes capas de la pila de protocolos, por ejemplo, al permitir la conectividad dispositivo a dispositivo (D2D) o la explotación de almacenamiento en el lado móvil. Si bien esta filosofía de diseño requiere principalmente un cambio en el nivel de nodo (el cambio de componentes), también tiene implicaciones a nivel arquitectura.
- 5) Comunicación máquina a máquina (M2M): Una inclusión natural de la comunicación M2M en 5G implica satisfacer tres requisitos fundamentalmente asociados con las diferentes clases de servicios de tasa de datos baja: soportar un número masivo de dispositivos de baja velocidad, mantener una tasa mínima de datos en prácticamente todas las circunstancias, y la transferencia de datos de muy baja latencia. Hacer frente a estos requisitos en 5G requiere nuevos métodos e ideas, tanto a nivel de componentes como de arquitectura.

En este capítulo, nos centramos en NOMA, que se espera que pueda aumentar el rendimiento del sistema y dar cabida a la conectividad masiva. Téngase en cuenta que 3GPP, LTE Rel-13, está realizando estudios hacia NOMA en la forma de transmisión de superposición de múltiples usuarios (MUST). NOMA permite que

varios usuarios compartan recursos en tiempo y frecuencia en la misma capa espacial a través de multiplexación en el dominio de la potencia y en el dominio de código.

Recientemente, varios esquemas de NOMA han atraído mucha atención, y por lo general se pueden dividir en dos categorías, multiplexación en el dominio de la potencia y en el dominio de código.

Aunque NOMA puede proporcionar ventajas atractivas, presenta algunos problemas, tales como el diseño avanzado del transmisor y el equilibrio entre el rendimiento y la complejidad del receptor.

## 5.2. Características de NOMA

En los esquemas convencionales de OMA, varios usuarios son asignados con recursos de radio que son ortogonales en el dominio tiempo, la frecuencia o del código. Idealmente, no hay interferencias entre usuarios múltiples, debido a la asignación de recursos ortogonal en OMA, la detección de un solo usuario, de manera sencilla se puede utilizar para separar las señales de los diferentes usuarios [54].

En teoría, se sabe que OMA no siempre puede alcanzar la capacidad de la suma de tasas de los sistemas inalámbricos multiusuario. En los sistemas de OMA convencionales, el número máximo de usuarios admitidos está limitada por la cantidad total de los recursos ortogonales y la granularidad de la calendarización.

Recientemente, NOMA se ha investigado para hacer frente a los problemas de OMA como se mencionó anteriormente. Básicamente, NOMA permite interferencias controlables para la asignación de recursos no ortogonales con el aumento tolerable en complejidad del receptor. En comparación con OMA, las principales ventajas de NOMA incluyen los siguientes [54].

- Mejora de la eficiencia espectral.

- Conectividad masiva: La asignación de recursos no ortogonales en NOMA indica que el número de usuarios o dispositivos compatibles no se limita estrictamente por la cantidad de recursos disponibles y su calendarización. Por lo tanto, NOMA tiene una capacidad de usuarios significativamente mayor que OMA mediante el uso de la asignación de recursos no ortogonal.
- Baja latencia de transmisión y el costo de señalización: En OMA convencional con transmisión basada en petición de recurso, en un primer momento un usuario tiene que enviar una petición de calendarización a la estación de base (BS). Luego, basándose en la petición recibida, la BS realiza la calendarización para la transmisión del enlace ascendente y envía una asignación de recurso sobre el canal de enlace descendente. Este procedimiento da como resultado una gran latencia y alto costo de señalización, que empeora o incluso es inaceptable en el escenario de la conectividad masiva prevista para 5G. Por el contrario, dicha calendarización dinámica no se requiere en algunos esquemas de enlace ascendente de NOMA, haciendo una transmisión de enlace ascendente libre de petición que puede reducir drásticamente la latencia de transmisión y la sobrecarga de señalización.

### 5.3. NOMA básico con receptor SIC

La Figura 5.2 ilustra el esquema NOMA básico a través de multiplexado de dominio de potencia con un receptor SIC en el enlace descendente. Téngase en cuenta que este esquema NOMA también se puede aplicar en el enlace ascendente. En la BS transmisora, se agregan en forma lineal señales para diferentes usuarios hasta en ciertas particiones de energía para equilibrar la suma de tasas de todos los usuarios multiplexados y la imparcialidad del rendimiento entre los usuarios individuales [54].

El receptor SIC se utiliza comúnmente para realizar la detección multiusuario (MUD). Debido al efecto cerca-lejos, las condiciones de canal pueden variar significativamente entre los usuarios. SIC se aplica a los usuarios con una relación de señal a ruido más interferencia (SINR) relativamente alta, y debe llevarse a cabo con el

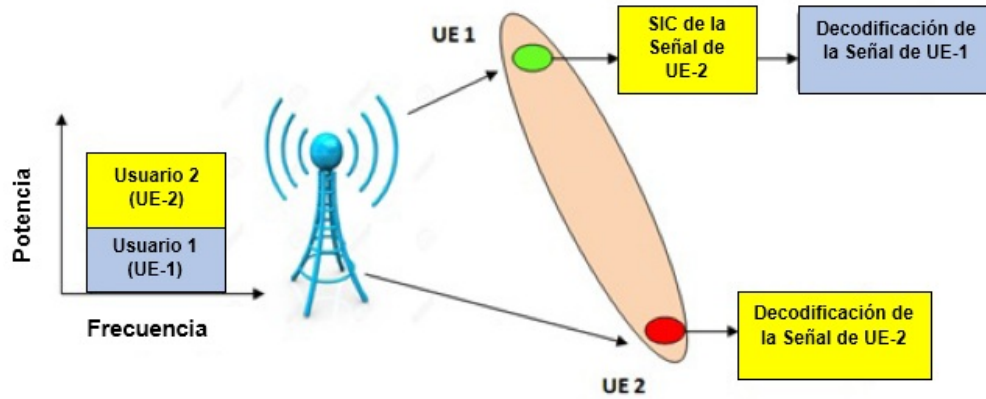


Figura 5.2: NOMA con receptor SIC.

fin de bajar la SINR.

Como podemos ver, la forma básica de NOMA con SIC explota la diferencia de SINR entre los usuarios, ya sea debido al efecto cerca-lejos o mediante la asignación de potencia no uniforme en el transmisor. Un esquema similar puede utilizarse para el enlace ascendente para aumentar la capacidad del sistema.

Por simplicidad se supone el caso de una sola antena transmisora y una receptora. La estación base transmite una señal ( $x_i$ ) al usuario  $UE - i$ , con potencia  $P_i$  (la suma de  $P_i$  es igual a  $P$ ) [54].

La señal recibida esta dada por:  $y_i = h_i x + w_i$ , donde  $x = \sqrt{P_1}x_1 + \sqrt{P_2}x_2$ ,  $h_i$  es el coeficiente del canal entre el usuario  $UE - i$  y la estación base,  $w_i$  denota el ruido AWGN incluyendo la interferencia inter celda [54].

En NOMA, el proceso SIC se implementa en el receptor del UE. El orden óptimo para la decodificación está en el orden decreciente de ganancia del canal normalizado por el ruido y la potencia de interferencia entre celdas (llamado simplemente ganancia del canal). Basado en este fin, se supone que cualquier usuario puede decodificar correctamente las señales de otros usuarios cuyo orden de decodificación viene antes del usuario correspondiente. Por lo tanto,  $UE - i$  puede eliminar la interferencia en-

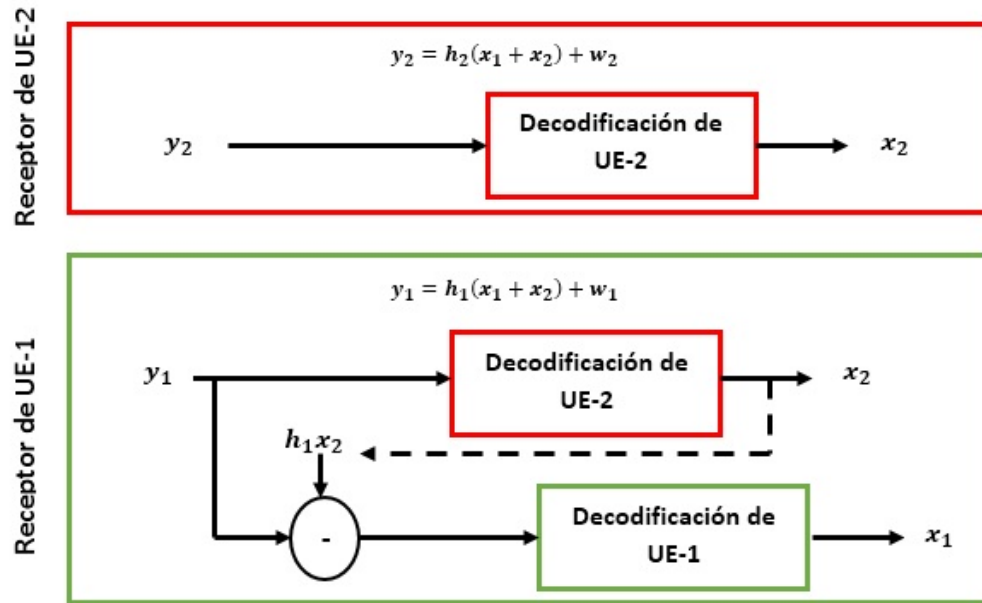


Figura 5.3: Proceso SIC.

tre usuarios de la  $j$ -ésima cuya ganancia de canal  $|h_j|^2$  es menor a la de los  $i$ -ésimos usuarios  $|h_i|^2$  ( $j < i$ ), es decir, las ganancias de los canales se ordenan de mayor a menor [54].

En la Figura 5.2,  $UE - 2$  no realiza la cancelación de interferencia ya que es el primero en el orden de decodificación.  $UE - 1$  primero decodifica  $x_2$  y resta de su componente de señal recibida  $y_1$ , y a continuación se decodifica  $x_1$  sin interferencia de  $x_2$ .

En el caso de UE-e, suponiendo que  $|h_1|^2 > |h_2|^2$ , UE-2 no realiza la cancelación de interferencia ya que viene primero en el orden de decodificación. El UE-1 decodifica primero  $x_2$  y resta su componente a la señal recibida  $y_1$ , luego decodifica  $x_1$  sin interferencia de  $x_2$ . Este proceso se muestra en la Figura 5.3.

## 5.4. NOMA Red

Básicamente, NOMA permite interferencias controlables por la asignación de recursos no ortogonales con el aumento tolerable en la complejidad del receptor, utilizando SIC. Un usuario puede eliminar la interferencia de otros usuarios, cuya ganancia de canal es menor que ese usuario. La precodificación multiusuario utilizada para NOMA de una sola celda tal vez no sea factible para el escenario NOMA red, ya que el precodificador para antenas de la estación base (BS) geográficamente separadas no forma realmente el haz físico que puede usarse fácilmente para NOMA intra-haz. Por esta razón proponemos un esquema para reducir la complejidad de la red y aumentar el rendimiento.

En la Figura 5.4 se muestra un sistema celular con dos celdas y cuatro usuarios. En cada celda se adopta un esquema NOMA de dos usuarios. Las estaciones base pueden cooperar entre sí para la transmisión y recepción conjunta de datos. El principal problema en la red NOMA es el diseño del precodificador, que se aplicará a los usuarios con condiciones de canal débiles para mitigar eficazmente la interferencia entre las celdas. En este caso son los usuarios 1 y 3. Si se asume que un precodificador perfecto es capaz de identificar las señales y decodificar adecuadamente cada señal, independientemente de la interferencia, las señales recibidas por los cuatro usuarios son dadas por la ecuación 5.1 [54].

$$y = HCx + w_i \quad (5.1)$$

donde  $H$  es el vector de ganancia de canal, es decir,  $H = [h_{0,j}, h_{1,j}, h_{2,j}, h_{3,j}]$ , donde  $j$  corresponde a la estación base ( $j = 0, 1$ ).  $C = [c_{0,j}, c_{1,j}, c_{2,j}, c_{3,j}]$  es el vector de precodificación,  $x = [x_0, x_1, x_2, x_3]$ , es el vector de las señales transmitidas de los cuatro usuarios,  $w_i$  representa el ruido AWGN incluyendo la interferencia inter-celda. La densidad espectral de potencia de  $w_i$  es  $N_{0,i}$ . La señal transmitida por cada estación base se puede expresar como sigue. BS-0:  $c_{0,0}x_0 + c_{1,0}x_1 + c_{2,0}x_2 + c_{3,0}x_3$ . Para BS-1:  $c_{0,1}x_0 + c_{1,1}x_1 + c_{2,1}x_2 + c_{3,1}x_3$ .

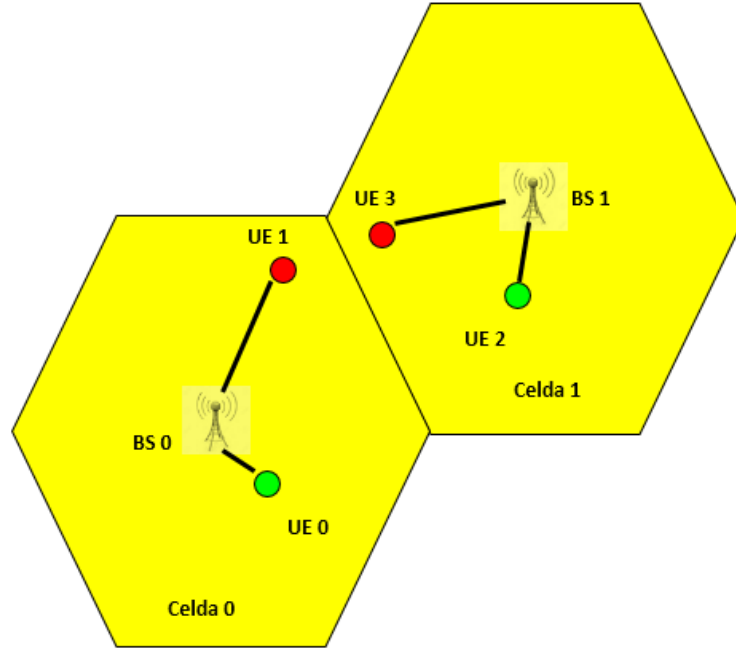


Figura 5.4: NOMA red.

Se observa que cada BS transmite información de cada usuario, esto es complicado cuando aumenta el número de usuarios y celdas, ya que aumenta el tiempo de procesamiento y la complejidad del predecodificador [55]. Por simplicidad suponemos que las interferencias de todas las celdas no afectan a los usuarios 0 y 2, ya que son usuarios cercanos a la estación base y presentan una ganancia de canal relativamente alta, sólo afectan a los usuarios 1 y 3. El vector de ganancia de canal se reduce a  $H = [h_{1,j}, h_{3,j}]$ . Entonces el predecodificador sólo se necesita para los usuarios 1 y 3,  $C = [c_{1,j}, c_{3,j}]$ . Las señales transmitidas por las estaciones base son, para BS-0:  $x_0 + c_{1,0}x_1 + c_{3,0}x_3$  y para BS-1:  $c_{1,1}x_1 + x_2 + c_{3,1}x_3$ . Por lo tanto, la interferencia de los usuarios 1 y 3 a los usuarios 0 y 2 puede ser cancelada correctamente. La SINR de cada usuario se puede expresar como sigue [54].

$$SINR_0 = \frac{|h_{0,0}|^2 P_0}{|h_{0,1}|^2 (P_2 + |c_{1,1}|^2 P_1 + |c_{3,1}|^2 P_3) + N_{0,0}} \quad (5.2)$$

Como se mencionó anteriormente, el usuario 0 puede cancelar las interferencias,

por lo tanto la ecuación 5.2 se reduce a la ecuación 5.3.

$$SINR_0 \approx \frac{|h_{0,0}|^2 P_0}{N_{0,0}} \quad (5.3)$$

De la misma manera se procede para el usuario 2, quedando la SINR como se muestra en la ecuación 5.4.

$$SINR_2 \approx \frac{|h_{2,1}|^2 P_2}{N_{0,2}} \quad (5.4)$$

Para el usuario 1, la SINR se calcula como sigue.

$$SINR_1 = \frac{CHP_1}{|h_{1,0}|^2 P_0 + |h_{1,1}|^2 P_2 + N_{0,1}} \quad (5.5)$$

donde  $C = [c_{1,j}, c_{3,j}]$ , y  $H = [h_{1,j}, h_{3,j}]$ , para  $j = 0, 1$ . De manera similar se sigue para el usuario 3.

$$SINR_3 = \frac{CHP_3}{|h_{3,0}|^2 P_0 + |h_{3,1}|^2 P_2 + N_{0,3}} \quad (5.6)$$

Podemos observar que la complejidad del problema aumenta a medida que aumenta el número de usuarios y el número de celdas y el predecodificador es necesario. Para simplificar el receptor se propone un esquema de operación coordinada, OMA para usuarios en el borde de la celda y NOMA para usuarios en el centro como se muestra en la Figura 5.5. Esto elimina el problema del predecodificador complejo, ya que los usuarios en el borde de la celda no necesitan dicho predecodificador, y la información de ganancia de canal de todos los usuarios del sistema ya no es necesaria. Este esquema simplemente requiere la información de los usuarios de la celda de servicio. Por lo tanto, se reduce el problema a un análisis de asignación de potencia. El análisis de la tasa de transmisión y la ganancia de canal para NOMA y OMA se desarrollan en la siguiente sección.



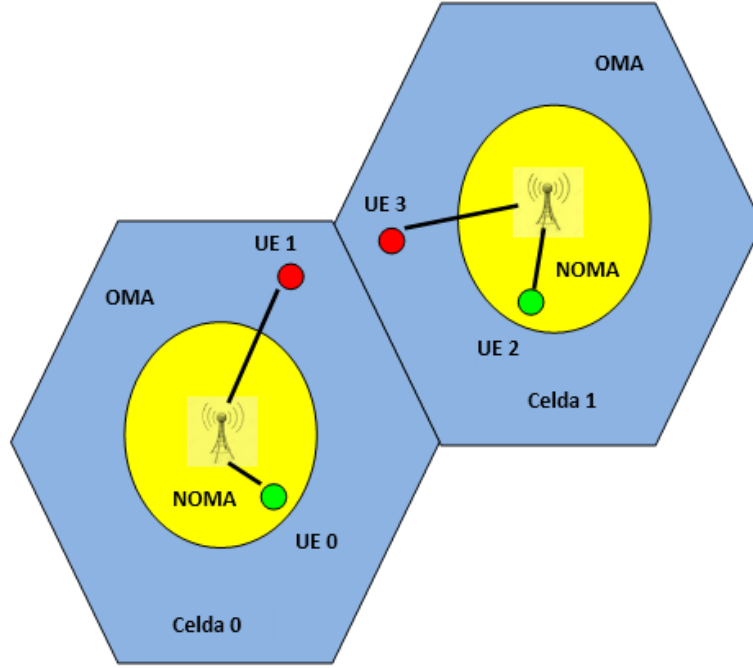


Figura 5.5: Esquema coordinado NOMA-OMA.

El receptor SIC se utiliza comúnmente para realizar la detección multiusuario. Debido al efecto cerca-lejos las condiciones de canal pueden variar significativamente entre los usuarios.

Por simplicidad se considera el caso de una sola antena transmisora y una receptora. Entonces, el problema es identificar los usuarios en la misma banda en la región NOMA y asignarles una potencia óptima, tal que la tasa de transmisión se maximice.

Se distribuyen uniformemente  $M$  usuarios en cada sub-banda. El ancho de banda se divide en múltiples sub-bandas. En cada sub-banda el servidor transmite a múltiples usuarios simultáneamente con diferentes potencias de transmisión para cada usuario.

La estación base transmite una señal  $x_i$  al usuario  $UE - i$ , con potencia  $P_i$ , en la misma banda. La señal recibida por el  $i$ -ésimo usuario está dada por la ecuación 5.7.

$$y_i = h_i x_i \sqrt{P_i} + h_i \sum_{l=i+1}^M x_l \sqrt{P_l} + h_i \sum_{l=1}^{i-1} x_l \sqrt{P_l} + I_i + v_i \quad (5.7)$$

donde:

- $M$  = Usuarios en la misma banda.
- $h_i$  = Ganancia de canal considerando las pérdidas por trayectoria, desvanecimiento (rápidos y lentos).
- $x_i$  = Señal transmitida para el usuario  $i$ .
- $P_i$  = Potencia transmitida del usuario  $i$ .
- $I_i$  = Interferencia intercelda.
- $v_i$  = AWGN experimentado por el usuario  $i$ .

Considerando las estadísticas de orden podemos considerar que las SINR están ordenados,  $\gamma_1 \geq \gamma_2 \geq \dots \geq \gamma_i \geq \gamma_{i+1} \geq \dots \geq \gamma_M$ . El usuario  $i$  puede decodificar y eliminar exitosamente la interferencia inter-usuario de los usuarios  $i + 1, \dots, M$  como se indica en el segundo termino de la ecuación 5.7. La potencia de la señal de los usuarios 1 hasta  $i - 1$ , que tienen mejores condiciones de canal que el usuario  $i$ , se pueden considerar como interferencia intra-celda por el usuario  $i$ , que es el tercer termino de la ecuación 5.7 [54].

El proceso SIC se implementa en el receptor del UE. El orden óptimo para decodificación está en el orden decreciente de la ganancia del canal (normalizado). Basado en esto, se considera que cualquier usuario puede decodificar correctamente las señales de otros usuarios cuyo orden de decodificación viene antes del usuario correspondiente. Las ganancias de los canales se ordenan de mayor a menor.

Por lo tanto, el usuario  $i$  puede decodificar y remover exitosamente la interferencia inter-usuario de los usuarios  $i + 1, \dots, M$  como se indica en el segundo termino de la ecuación 5.7 ( $h_i \sum_{l=i+1}^M x_l \sqrt{P_l}$ ).

Sin embargo, la potencia de la señal de los usuarios  $1, \dots, i - 1$ , que son los que tienen mejor condiciones de canal que el usuario  $i$ , se pueden considerar como interferencia intra-celda por el usuario  $i$ , que es el tercer término de la ecuación 5.7 ( $h_i \sum_{l=1}^{i-1} x_l \sqrt{P_l}$ ).

Definiendo la ganancia normalizada del usuario  $i$  como:  $\gamma_i = \frac{|h_i|^2}{N_{0,i}}$ . El throughput del usuario  $i$  (UE- $i$ ),  $R_i$ , se calcula con la ecuación 5.8, utilizando la ecuación de Shannon.

$$R_i = \log_2\left(1 + \frac{P_i \gamma_i}{\sum_{k=1}^{i-1} P_k \gamma_i + 1}\right) \quad (5.8)$$

Considerando el caso de 2 usuarios ( $i$  y  $j$ ), con  $\gamma_i > \gamma_j$ , obtenemos las tasas mostradas en la ecuación 5.9. Asumiendo que sólo 2 usuarios se multiplexan en una banda, el problema de optimización a resolver por el esquema de asignación de potencia cada TTI (intervalo de tiempo de transmisión) se puede formular como el problema de maximizar la suma de tasas.

$$R_i = \log_2(1 + P_i \gamma_i) \quad R_j = \log_2\left(1 + \frac{P_j \gamma_j}{1 + P_i \gamma_j}\right) \quad (5.9)$$

De esto se puede observar que la asignación de potencia de un usuario impacta en la tasa alcanzable de otros usuarios. Si  $M$  es grande el encabezado de señalización y las probabilidades de error de propagación debido a una decodificación incorrecta de SIC puede ser alto. Por lo tanto, el número de usuarios multiplexados sobre una misma sub-banda necesita ser limitado. En unos trabajos se sugiere que en una sub-banda el número de usuarios está limitado a 2 ó 3 ( $M=[2,3]$ ) [54].

De todos los algoritmos de detección multiusuario, SIC es una opción prometedora porque puede decodificar un número de señales concurrentes en lugar de tratarlos como ruido. El análisis teórico y experimental demuestra que SIC puede mejorar eficazmente el rendimiento de la red inalámbrica [54].

SIC funciona de la siguiente manera: sin pérdida de generalidad, suponga que hay  $M$  usuarios, cuyas potencias recibidas son  $P_1 \leq P_2 \leq \dots \leq P_M$ , siempre y cuando cumplan con las siguientes condiciones:

- Paso 1

$$\frac{P_M \gamma_M}{\sum_{k=1}^{M-1} P_k \gamma_M + 1} \geq \beta \quad (5.10)$$

- Paso 2

$$\frac{P_{M-1} \gamma_{M-1}}{\sum_{k=1}^{M-2} P_k \gamma_{M-1} + 1} \geq \beta \quad (5.11)$$

- ...

- Paso (M-i+1)

$$\frac{P_i \gamma_i}{\sum_{k=1}^{i-1} P_k \gamma_i + 1} \geq \beta \quad (5.12)$$

El receptor intenta decodificar las señales en el orden  $M, M - 1, \dots, 2, 1$ , para decodificar la potencia recibida  $P_i$ . SIC primero decodifica la potencia de la señal más fuerte, entonces la señal de la potencia más fuerte se elimina de la señal recibida, entonces SIC selecciona la potencia de la señal residual en el proceso iterativo y así sucesivamente hasta que el resto de la señal recibida no satisface la fórmula anterior.

## 5.5. Tasa de transmisión promedio

En OMA el ancho de banda de  $\alpha$  ( $0 < \alpha < 1$ ) Hz se asigna a UE-0 (UE-2) y el resto del ancho de banda,  $1 - \alpha$  Hz, se asigna a UE-1 (UE-3). La tasa de transmisión (normalizada) se calcula usando la ecuación de Shannon. El throughput de  $UE - i$ ,  $R_i$ , están dados en las ecuaciones 5.13 y 5.14, para la celda 0 (esto también aplica para la celda 1) [54].

Para OMA:

$$R_0 = \alpha \log_2 \left( 1 + \frac{P_0 |h_0|^2}{\alpha N_{0,0}} \right) \quad (5.13)$$

$$R_1 = (1 - \alpha) \log_2 \left( 1 + \frac{P_1 |h_1|^2}{(1 - \alpha) N_{0,1}} \right) \quad (5.14)$$

En NOMA, se asume una decodificación exitosa y una propagación sin error, el throughput de  $UE - i$ ,  $R_i$ , se representa en las ecuaciones 5.15 y 5.16.

$$R_0 = \log_2 \left( 1 + \frac{P_0 |h_0|^2}{N_{0,0}} \right) \quad (5.15)$$

$$R_1 = \log_2 \left( 1 + \frac{P_1 |h_1|^2}{P_0 |h_1|^2 + N_{0,1}} \right) \quad (5.16)$$

Puede observarse que la asignación de potencia para cada UE afecta en gran medida el rendimiento del usuario y por tanto el esquema de modulación y codificación (MCS) utilizado para la transmisión de datos de cada UE. Además, en NOMA todo el ancho de banda está disponible para cada usuario.

Definimos a  $\beta$  como el umbral de la SINR y  $\sigma^2$  como la potencia de ruido, entonces la tasa de transmisión se puede calcular utilizando la ecuación de Shannon como se muestra en la ecuación 5.17.

$$R = W \log_2(1 + \beta) \quad (5.17)$$

donde  $W$  es el ancho de banda del canal.

La potencia recibida se puede condicionar como:  $P_1 \leq P_2 \leq \dots \leq P_M$ , entonces:

$$P_i \geq \beta(1 + \beta)^{i-1} \sigma^2 \quad (5.18)$$

- Comprobación:

$$\frac{P_{l+1}}{\sum_{i=1}^l P_i + \sigma^2} \geq \beta \quad (5.19)$$

$$P_{l+1} \geq \beta \left( \sum_{i=1}^l P_i + \sigma^2 \right) \geq \beta \left( \sum_{i=1}^l \beta(1 + \beta)^{i-1} \sigma^2 + \sigma^2 \right) \quad (5.20)$$

- Obtenemos:

$$\sum_{i=1}^l \beta(1 + \beta)^{i-1} \sigma^2 = \beta \sigma^2 \sum_{i=1}^l (1 + \beta)^{i-1} \quad (5.21)$$

- Definimos:

$$A = (1 + \beta) \quad (5.22)$$

$$\sum_{i=1}^l A^{i-1} = S \quad (5.23)$$

- Por lo tanto:

$$S = A^0 + A^1 + \dots + A^{l-1} = 1 + A + A^2 + \dots + A^{l-1} \quad (5.24)$$

$$S * A = A + A^2 + A^3 + \dots + A^l \quad (5.25)$$

- Por lo tanto:

$$S - S * A = S(1 - A) = 1 - A^l \quad (5.26)$$

- Despejando:

$$S = \frac{1 - A^l}{1 - A} \quad (5.27)$$

- Sustituyendo:

$$\sum_{i=1}^l (1 + \beta)^{i-1} = \frac{1 - (1 + \beta)^l}{1 - (1 + \beta)} = \frac{1 - (1 + \beta)^l}{-\beta} \quad (5.28)$$

- Por lo tanto:

$$\beta \sigma^2 \left( \frac{1 - (1 + \beta)^l}{-\beta} \right) = -\sigma^2 (1 - (1 + \beta)^l) \quad (5.29)$$

- Entonces:

$$P_{l+1} \geq \beta [-\sigma^2 + \sigma^2 (1 + \beta)^l + \sigma^2] = \beta (1 + \beta)^l \sigma^2 \quad (5.30)$$

- Por lo tanto:

$$P_l \geq \beta (1 + \beta)^{l-1} \sigma^2 \quad (5.31)$$

## 5.6. Ganancia de canal

En esta sección se presenta un análisis por estadística de orden para obtener la función de densidad de probabilidad de la ganancia del canal. En NOMA las ganancias de canal son variables aleatorias ordenadas, que pueden ser analizados por estadísticas de orden [56]. Si se considera un canal con desvanecimientos rápidos, entonces se puede modelar con variables aleatoria con función de distribución Rayleigh, es decir, el canal puede estar dado por  $h_{j,i}|z|^{-\alpha}$ , donde  $|z|^{-\alpha}$  es la pérdida por

trayectoria, con  $\alpha$  el exponente de pérdida por trayectoria ( $\alpha > 2$ ), por lo tanto  $h_{j,i}$  se distribuye exponencialmente. Hay  $N$  usuarios, de tal manera que el usuario 0 es el más cercano a la BS y el usuario  $N$  es el más lejano, es decir,  $h_{0,i} > h_{1,i} > \dots > h_{N,i}$ . Por lo tanto, las ganancias de canal son estadísticas de orden dependientes. Debemos encontrar la función de distribución de ganancias del canal, que se puede calcular como se muestra en la ecuación 5.32 [54].

$$H = \frac{|h_{j,i}|^2 |z|^{-\alpha}}{\sum x_i^2 |r|^{-\alpha} + N_{0,j}} \quad (5.32)$$

Para obtener las estadísticas de orden de  $H$ , tal que  $h_{0,i} > h_{1,i} > h_{2,i} > \dots$ , primero encontramos la probabilidad de tener la mayor ganancia de canal, la cual se muestra en la ecuación 5.33.

$$P(\max(h_{0,i}, h_{1,i}, h_{2,i}, \dots) \leq \theta) = P\left(\frac{|h_{0,i}|^2 |z|^{-\alpha}}{\sum x_i^2 |r|^{-\alpha} + N_{0,0}} \leq \theta, \frac{|h_{1,i}|^2 |z|^{-\alpha}}{\sum x_i^2 |r|^{-\alpha} + N_{0,1}} \leq \theta, \dots |h_{0,i} > h_{1,i} > h_{2,i} > \dots\right) \quad (5.33)$$

Considerando un sólo término (por ejemplo el primer término), se tiene:

$$P\left(\frac{|h_{0,i}|^2 |z|^{-\alpha}}{\sum x_i^2 |r|^{-\alpha} + N_{0,0}} \leq \theta\right) = P(|h_{0,i}|^2 \leq \theta |z|^\alpha (\sum x_i^2 |r|^{-\alpha} + N_{0,0})) = 1 - e^{-\theta |z|^\alpha (\sum x_i^2 |r|^{-\alpha} + N_{0,0})} \quad (5.34)$$

Suponemos que se tienen  $C$  usuarios, esto es,  $C$  valores de estadísticas de orden, entonces, por un resultado clásico la teoría de la probabilidad, la probabilidad  $P_{r,C}$  de realización de al menos  $r$  de eventos  $C$  ( $Y_1, \dots, Y_C$ ), está dado por la ecuación 5.35 [56].

$$P_{r,C} = \sum_{j=r}^C (-1)^{(j-r)} \binom{j-1}{r-1} \sum_{1 \leq \dots \leq C} Pr\{\max(Y_1, \dots, Y_j) \leq y\} \quad (5.35)$$

Dado que  $Y_i$  tiene una cdf conjunta arbitraria,  $F(y_1, \dots, y_C)$ , podemos obtener la expresión para la cdf  $F_{r,C}$ . Consideramos las variables  $Y_1, \dots, Y_C$  como estadísticas de orden con función de distribución  $F(y)$  y función de densidad  $f(y)$ . De la función de densidad marginal de  $Y_i$  y la función de densidad conjunta de  $Y_i$  y  $Y_j$  podemos obtener la función de densidad condicional de  $Y_i$ , dado  $Y_j$ , esto se muestra en la ecuación 5.36.

$$f_{Y_{(i)}}(y_i | Y_{(j)}) = \frac{(j-1)!}{(i-1)!(j-i-1)!} \left[ \frac{1 - e^{-\theta|z|^\alpha (\sum x_i^2 |r|^{-\alpha + N_0})}}{1 - e^{-\theta|z|^\alpha (\sum x_j^2 |r|^{-\alpha + N_0})}} \right]^{(i-1)} * \\ \left[ \frac{e^{-\theta|z|^\alpha (\sum x_i^2 |r|^{-\alpha + N_0})} - e^{-\theta|z|^\alpha (\sum x_j^2 |r|^{-\alpha + N_0})}}{1 - e^{-\theta|z|^\alpha (\sum x_j^2 |r|^{-\alpha + N_0})}} \right]^{(j-i-1)} * \\ \frac{|z|^\alpha (\sum x_i^2 |r|^{-\alpha + N_0}) (e^{-\theta|z|^\alpha (\sum x_i^2 |r|^{-\alpha + N_0})})}{1 - e^{-\theta|z|^\alpha (\sum x_j^2 |r|^{-c_1, 1x_1 + x_2 + c_3, 1x_3 + N_0})}} \quad (5.36)$$

Usando este resultado podemos obtener la expresión de la función de densidad condicional de la ganancia de canal, dado en la ecuación 5.37.

$$f_{H_{(i)}}(h_i | h_{(j)}) = \frac{(j-1)!}{(i-1)!(j-i-1)!} \left[ \frac{F(y_i)}{F(y_j)} \right]^{(i-1)} \left[ \frac{F(y_j) - F(y_i)}{F(y_j)} \right]^{(j-i-1)} \frac{f(y_i)}{F(y_j)} \quad (5.37)$$

Para obtener la función de distribución usamos  $P_{r,C}$ , y obtenemos la ecuación 5.38.

$$F_{r,C}(H) = \sum_{j=r}^C (-1)^{(j-r)} \binom{j-1}{r-1} \sum_{1 \leq \dots \leq C} \int_0^\infty f_{H_{(i)}}(h_i | h_{(j)}) \cdot d\theta \quad (5.38)$$

Utilizando el esquema mostrado en la Figura 5.5 y los siguientes parámetros del sistema para OMA con igual ancho de banda e igual potencia de transmisión se asignan a cada UE ( $\alpha = 0,5, P_0 = P_1 = P_2 = P_3 = P/2$ ). En NOMA



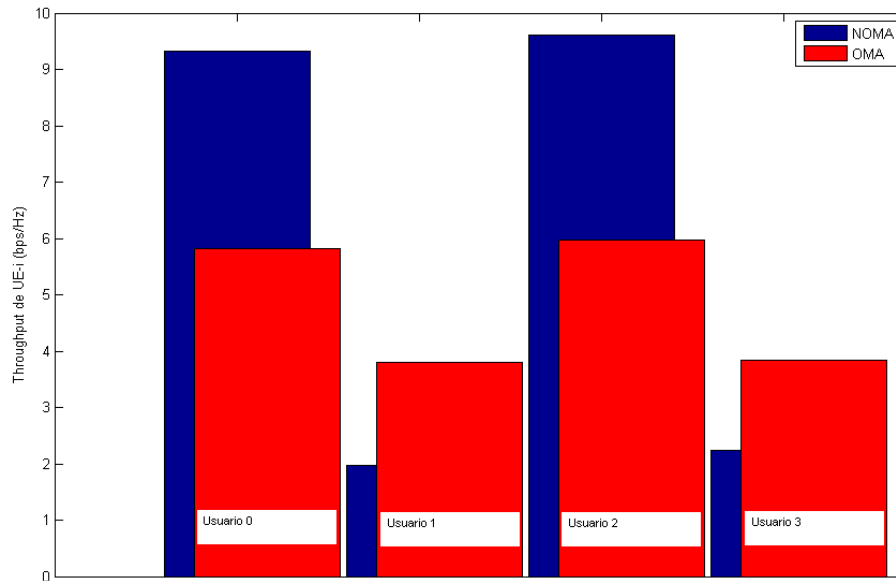


Figura 5.6: *Throughput* para diferentes usuarios con NOMA y OMA.

$$P_0 = P/5, P_1 = 4P/5, P_2 = P/5, P_3 = 4P/5.$$

Se asume que la ganancia de canal del usuario 0 es mayor que la ganancia de canal del usuario 1. De manera similar, la ganancia de canal del usuario 2 es mayor que la ganancia de canal del usuario 3. Usando el análisis de ganancia de canal con estadísticas de orden y un precodificador ZF (zero-forcing), tenemos el resultado en la Figura 5.6. Se observa que para los usuarios cercanos a la BS el esquema NOMA presenta un mejor rendimiento, mientras que para los usuarios al borde de la celda el esquema OMA presenta un mejor rendimiento, las tasas de usuario se calculan para OMA  $R_0 = 5,8231bps/Hz$ ,  $R_1 = 3,8064bps/Hz$ ,  $R_2 = 5,9638bps/Hz$ ,  $R_3 = 3,8303bps/Hz$  y para NOMA:  $R_0 = 9,3262bps/Hz$ ,  $R_1 = 1,9682bps/Hz$ ,  $R_2 = 9,6071bps/Hz$ ,  $R_3 = 2,2317bps/Hz$  [54].

NOMA tiene una ganancia en los usuarios cerca de la estación base y OMA en los usuarios en el borde de la celda. Las ganancias son: para OMA  $R_1 = 48,4\%$ ,  $R_3 = 41,7\%$  y para NOMA  $R_0 = 37,71\%$ ,  $R_2 = 37,9\%$ . Con estos resultados podemos observar que el esquema coordinado tiene una ganancia mayor del 40% que la red NOMA en el borde de la celda (anillo 2), como se muestra en la Figura 5.7 [54].

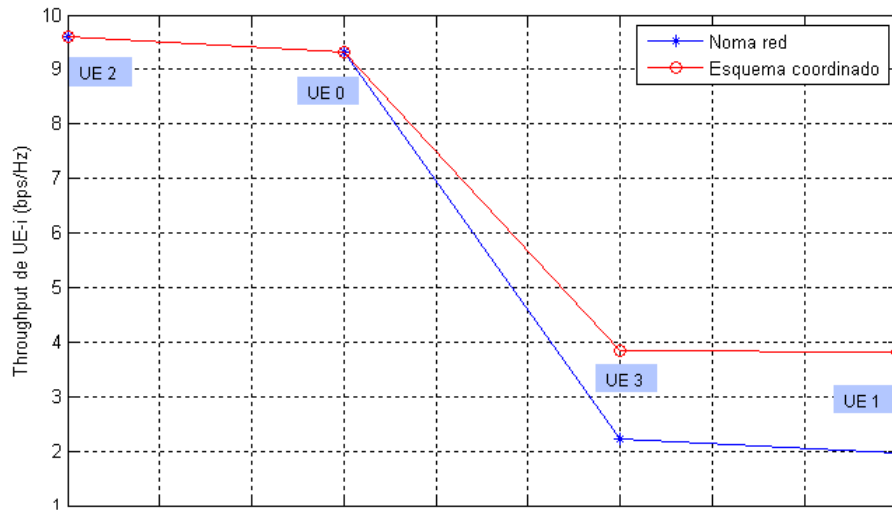


Figura 5.7: *Throughput* NOMA red vs esquema coordinado.

En este capítulo, propusimos un esquema de operación coordinado entre OMA y NOMA para reducir la complejidad del receptor en NOMA con receptor SIC. Red NOMA se espera que aumente el rendimiento del sistema y dar cabida a la conectividad masiva. Las ganancias alcanzables se muestran prometedoras, del orden del 37% en comparación con OMA en el centro de la celda y OMA es superior en más del 42% en comparación con NOMA en el borde de la celda. Por lo tanto, OMA puede coexistir con NOMA para 5G, siempre buscando aumentar el rendimiento del sistema, y mejorar la calidad de los usuarios.

---

# Conclusiones generales y trabajo a futuro

El presente trabajo de tesis doctoral se centró en el estudio de la calendarización persistente para sistemas celulares y en particular en la evaluación de su desempeño a través de la tasa de transmisión, la probabilidad de terminación forzada y la probabilidad de bloqueo. La importancia de este esquema de calendarización radica en que es un esquema que elimina el problema de cuello de botella que presenta la calendarización dinámica y se prevé que seguirá vigente durante varios años más debido a que es considerado para los sistemas celulares 5G.

El trabajo comienza con una revisión de aspectos generales como son los esquemas de acceso al medio para los sistemas celulares y el estudio detallado de los esquemas de calendarización.

Posteriormente se realizó una revisión de las distintas funciones de distribución de usuarios, para realizar una evaluación que se acerque más a la realidad, así como los distintos tipos de servicios para los nuevos sistemas celulares, esto con el fin de mostrar que la propuesta desarrollada seguirá vigente durante varios años.

Con el estudio del rendimiento del esquema de calendarización persistente, obtuvimos el número de bloques de recursos para diferentes esquemas de modulación y codificación basados en la SIR, teniendo en cuenta las estadísticas de orden. Esto se complementó con un análisis de teletráfico para usuarios persistentes que consideran las condiciones del canal y las interferencias en las celdas vecinas, lo que se refleja

en las tasas de transmisión de acuerdo con el esquemas de modulación y codificación asignado. El modelo para el análisis de teletráfico se muestra en la Figura 3.3. Este modelo es bidimensional y está limitado por los estados válidos del sistema. Las estadísticas de orden de la SIR son necesarias para calcular la probabilidad de asignación de banda para el usuario. Teniendo en cuenta las estadísticas de orden, las áreas con mejores condiciones de canal aumentan como se puede ver en la Figura 4.3 y con esto aumentan las probabilidades de transmisión a una mejor velocidad de transmisión que se refleja en un mejor esquema de modulación y codificación, esto se muestra en la Tabla 4.4. Se analizaron dos enfoques diferentes para la asignación de recursos: asignación aleatoria y asignación de máxima SIR, en los resultados se observa que la asignación de máxima SIR tiene una ganancia significativa de hasta 20 % con respecto a la asignación aleatoria, como se muestra en la Figura 4.9. Se observa en la Figura 4.7 que la probabilidad de terminación forzada se reduce significativamente cuando se usa el análisis con estadísticas de orden, debido a que la probabilidad de que el usuario transmita en su mejor banda es mayor. Se observa que las estadísticas de orden modelan los sistemas celulares con mayor precisión. Como ejemplo, los resultados numéricos muestran que la probabilidad de terminación forzada se corrige en más del 70 % en comparación con el análisis sin estadísticas de orden para la distribución de usuario uniforme y el tráfico VoIP. Siguiendo la misma metodología, se obtuvieron la probabilidad de bloqueo y la probabilidad de terminación forzada para el tráfico Skype y la distribución exponencial de usuario no uniforme, que son servicios considerados para 5G. Este análisis puede servir para dimensionar redes reales.

También se desarrolló una propuesta de operación coordinada entre OMA y NOMA para reducir la complejidad del receptor en el dominio de la potencia, con SIC en el receptor. Se espera que NOMA Red aumente el rendimiento del sistema y permita una conectividad masiva. Las ganancias alcanzables se muestran prometedoras y para nuestro esquema son mayores en más del 42 % en comparación con NOMA Red en el borde de la celda. Por lo tanto, OMA puede coexistir con NOMA para 5G, siempre buscando aumentar el rendimiento del sistema y mejorar la calidad de los usuarios.

Como trabajo futuro podemos mencionar lo siguiente:

- Evaluar el desempeño del sistema con una cola de espera, y tomar como parámetro de calidad de servicio el retardo en los usuarios que esperan en la cola.
- Evaluar el desempeño de nuestra metodología desarrollada para otros tipos de servicios como son tráfico Gtalk y tráfico Asterisk, con otras distribuciones de usuarios no uniforme.
- Realizar un análisis para una calendarización semipersistente y calcular el costo de la persistencia.
- Desarrollar un modelo para la asignación óptima de potencia para la selección de usuarios NOMA.
- Evaluar el esquema coordinado propuesto para NOMA en el dominio de código.

---

# Apéndice A

Supongamos que tenemos  $C$  bandas, es decir,  $C$  valores de las estadísticas de orden, luego, por un resultado clásico en la teoría de probabilidad, la probabilidad  $P_{r,C}$  de la realización de al menos  $r$  de los  $C$  eventos  $A_1, \dots, A_C$  esta dado por:

$$P_{r,C} = \sum_{j=r}^C (-1)^{(j-r)} \binom{j-1}{r-1} S_j, \quad (39)$$

donde:

$$S_j = \sum_{1 \leq \dots \leq C} Pr\{A_{i:1}, \dots, A_{i:j}\} = \sum_{1 \leq \dots \leq C} Pr\{\max(Y_{i:1}, \dots, Y_{i:j}) \leq y\} \quad (40)$$

$$S_j = \sum_{1 \leq \dots \leq C} Pr\{A_{i:1}, \dots, A_{i:j}\} = H_{j:j}(y) \quad (41)$$

Sustituyendo en la ecuación (8), obtenemos:

$$P_{r,C}(y) = \sum_{j=r}^C (-1)^{(j-r)} \binom{j-1}{r-1} H_{j:j}(y) \quad (42)$$

Sea  $Y_1, Y_2, \dots, Y_C$  una muestra aleatoria de una población absolutamente continua con cdf  $F(y)$  y función de densidad  $f(y)$ , y sea  $Y_{(1:C)} \leq Y_{(2:C)} \leq \dots \leq Y_{(C:C)}$  denotan las estadísticas de orden obtenidas de esta muestra. Entonces, la distribución condicional de  $Y_{(i:C)}$ , dado que  $Y_{(j:C)} = y_j$  para  $j > i$ , es la misma que la distribución de la  $i$ -ésima estadística de orden en una muestra de tamaño  $j-1$  de una población cuya distribución  $F(y)$  simplemente truncada a la derecha en  $y_j$ .

De la función de densidad marginal de  $Y_{(i:C)}$  y la función de densidad conjunta de  $Y_{(i:C)}$  y  $Y_{(j:C)}$  obtenemos la función de densidad condicional de  $Y_{(i:C)}$ , dado que  $Y_{(j:C)} = y_j$ , como:

$$f_{Y_{(i:C)}}(y_i | Y_{(j:C)} = y_j) = f_{Y_{(i:C)}, Y_{(j:C)}}(y_i, y_j) / f_{Y_{(j:C)}}(y_j) = \frac{(j-1)!}{(i-1)!(j-i-1)!} \left[ \frac{F(y_i)}{F(y_j)} \right]^{(i-1)} \cdot \left[ \frac{F(y_j) - F(y_i)}{F(y_j)} \right]^{(j-i-1)} \cdot \frac{f(y_i)}{F(y_j)} \quad (43)$$

Aquí  $i < j$  y  $-\infty < y_i, y_j$ . Tenga en cuenta que  $F(y_i)/F(y_j)$  y  $f(y_i)/F(y_j)$  son la cdf y función de densidad de la población cuya distribución se obtiene truncando la distribución  $F(y)$  en la derecha en  $y_j$ .

Por otro lado, la SINR de la  $i$ -ésima banda esta dado por:

$$SINR_i = \frac{h_{z,i} z^{-\alpha}}{\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha}} \quad (44)$$

donde  $\sigma^2$  es la varianza del ruido,  $i = 1, \dots, C$  y el desvanecimiento a pequeña escala entre la estación móvil y la estación base interferente  $y \in \phi$  en la  $i$ -ésima banda esta dado por  $h_{y,i}$ . Las variables aleatorias interferentes son considerados como desvanecimientos Rayleigh, lo cual implica que  $h_{y,i}$  se distribuye exponencialmente. La pérdida por trayectoria esta dado por  $z^{-\alpha}$  donde  $\alpha$  es el exponente de pérdida por trayectoria. Un nodo está configurado para estar en cobertura si su SINR recibido es mayor que un umbral  $\theta$ .

Las estadísticas de orden de las  $SINRs$  ( $SINR_{i:C}$ ) ordenadas son necesarias para calcular la probabilidad de transmisión del usuario. Obtenemos la función de distribución de las estadísticas de orden de las bandas, es decir,  $P(SINR_{i:C} \leq \theta)$ , tal que  $SINR_{C:C} \geq SINR_{C-1:C} \geq \dots \geq SINR_{1:C}$ . Comenzamos primero con la distribución

de la mejor banda  $SINR_C$ .

$$P(C, \theta) = P(\text{máx}\{SINR_{i:1}, \dots, SINR_{i:C}\} \leq \theta) = P\left(\frac{h_{z,1}z^{-\alpha}}{\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,1}y^{-\alpha}} \leq \theta, \dots,\right.$$

$$\left. \frac{h_{z,C}z^{-\alpha}}{\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,C}y^{-\alpha}} \leq \theta \mid SINR_{C:C} \geq SINR_{C-1:C} \geq \dots \geq SINR_{1:C}\right) \quad (45)$$

Tomando un sólo termino como ejemplo, dado que las variables aleatorias inter-ferentes son consideradas como desvanecimientos Rayleigh, lo cual implica que  $h_{y,i}$  se distribuye exponencialmente, obtenemos:

$$\begin{aligned} P\left(\frac{h_{z,C}z^{-\alpha}}{\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,C}y^{-\alpha}} \leq \theta\right) &= P\left(h_{z,C}z^{-\alpha} \leq \theta(\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,C}y^{-\alpha})\right) \\ &= P\left(h_{z,C} \leq \theta z^\alpha(\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,C}y^{-\alpha})\right) \end{aligned} \quad (46)$$

Debido a que  $h_{y,i}$  esta distribuido exponencialmente se obtiene:

$$P\left(h_{z,C} \leq \theta z^\alpha(\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,C}y^{-\alpha})\right) = 1 - e^{-\theta z^\alpha(\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,C}y^{-\alpha})} \quad (47)$$

Utilizando la ecuación (12), obtenemos la función de densidad de cada estadística de orden, condicionado a estadísticas de orden superiores ( $j > i$ ).

$$\begin{aligned} f_{SINR_{(i:C)}}(SINR_i | SINR_{(j:C)} = SINR_j) &= f_{SINR_{(i:C)}, SINR_{(j:C)}}(SINR_i, SINR_j) / f_{SINR_{(j:C)}}(SINR_j) \\ &= \frac{(j-1)!}{(i-1)!(j-i-1)!} \left[ \frac{F(SINR_i)}{F(SINR_j)} \right]^{(i-1)} \left[ \frac{F(SINR_j) - F(SINR_i)}{F(SINR_j)} \right]^{(j-i-1)} \cdot \frac{f(SINR_i)}{F(SINR_j)} \end{aligned}$$



$$\begin{aligned}
 &= \frac{(j-1)!}{(i-1)!(j-i-1)!} \cdot \left[ \frac{1 - e^{-\theta z^\alpha (\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha})}}{1 - e^{-\theta z^\alpha (\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,j} y^{-\alpha})}} \right]^{(i-1)} \\
 &\quad \left[ \frac{e^{-\theta z^\alpha (\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha})} - e^{-\theta z^\alpha (\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,j} y^{-\alpha})}}{1 - e^{-\theta z^\alpha (\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,j} y^{-\alpha})}} \right]^{(j-i-1)} \cdot \frac{z^\alpha (\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha}) e^{-\theta z^\alpha (\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha})}}{1 - e^{-\theta z^\alpha (\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,j} y^{-\alpha})}} \quad (48)
 \end{aligned}$$

Utilizando la ecuación (8), obtenemos la probabilidad que  $P = (SINR_{r:C} \leq \theta)$  la cual denotamos por  $F_{r:C}(SINR)$  y está dado por:

$$\begin{aligned}
 F_{r:C}(SINR) &= \sum_{j=r}^C (-1)^{(j-r)} \binom{j-1}{r-1} \cdot \sum_{1 \leq \dots \leq C} P_r \{ \max(SINR_{i:1}, \dots, SINR_{i:j}) \leq \theta \} = \\
 &\quad \sum_{j=r}^C (-1)^{(j-r)} \binom{j-1}{r-1} \sum_{1 \leq \dots \leq C} \int_0^\infty \frac{(j-1)!}{(i-1)!(j-i-1)!} \cdot \left[ \frac{1 - e^{-\theta z^\alpha (\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha})}}{1 - e^{-\theta z^\alpha (\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,j} y^{-\alpha})}} \right]^{(i-1)} \\
 &\quad \left[ \frac{e^{-\theta z^\alpha (\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha})} - e^{-\theta z^\alpha (\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,j} y^{-\alpha})}}{1 - e^{-\theta z^\alpha (\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,j} y^{-\alpha})}} \right]^{(j-i-1)} \cdot \frac{z^\alpha (\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha}) e^{-\theta z^\alpha (\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha})}}{1 - e^{-\theta z^\alpha (\sigma^2 + \sum_{y \in \phi\{z\}} h_{y,j} y^{-\alpha})}} d\theta \quad (49)
 \end{aligned}$$

Por facilidad no se considera el ruido de fondo, entonces la ecuación (18) se reduce a:

$$\begin{aligned}
 F_{r:C}(SIR) &= \sum_{j=r}^C (-1)^{(j-r)} \binom{j-1}{r-1} \cdot \sum_{1 \leq \dots \leq C} P_r \{ \max(SIR_{i:1}, \dots, SIR_{i:j}) \leq \theta \} = \\
 &\quad \sum_{j=r}^C (-1)^{(j-r)} \binom{j-1}{r-1} \sum_{1 \leq \dots \leq C} \int_0^\infty \frac{(j-1)!}{(i-1)!(j-i-1)!} \cdot \left[ \frac{1 - e^{-\theta z^\alpha \left( \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha} \right)}}{1 - e^{-\theta z^\alpha \left( \sum_{y \in \phi\{z\}} h_{y,j} y^{-\alpha} \right)}} \right]^{(i-1)} \cdot
 \end{aligned}$$

$$\left[ \frac{e^{-\theta z^\alpha \left( \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha} \right)} - e^{-\theta z^\alpha \left( \sum_{y \in \phi\{z\}} h_{y,j} y^{-\alpha} \right)}}{1 - e^{-\theta z^\alpha \left( \sum_{y \in \phi\{z\}} h_{y,j} y^{-\alpha} \right)}} \right]^{(j-i-1)} \cdot \frac{z^\alpha \left( \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha} \right) e^{-\theta z^\alpha \left( \sum_{y \in \phi\{z\}} h_{y,i} y^{-\alpha} \right)}}{1 - e^{-\theta z^\alpha \left( \sum_{y \in \phi\{z\}} h_{y,j} y^{-\alpha} \right)}} d\theta \quad (50)$$

---

## Referencias

- [1] R. Rom and M. Sidi, *Multiple Access Protocols Performance and Analysis*, Springer-Verlag.
- [2] L. Dai; B. Wang; Y. Yuan; S. Han; C. Lin I. and Z. Wang, *Non-orthogonal Multiple Access for 5G: Solutions, Challenges, Opportunities, and Future Research Trends*, IEEE Communications Magazine, vol. 53, no. 9, pp. 74 - 81, 2015.
- [3] 3GPP TR 36.804, *Base Station Radio Transmission and Reception*, 2008.
- [4] Farooq Khan, *LTE for 4G Mobile Broadband Air Interface Technologies and Performance*, Cambridge, U.K., Cambridge Univ. Press, 2009.
- [5] M. Chen; V. Leung and R. Xiaorong, *IEEE Access Special Section Editorial: 5G Wireless Technologies: Perspectives on the Next Generation of Mobile Communications and Networking*, IEEE Access, vol. 2, pp. 1686 - 1688, 2014.
- [6] Nokia Siemens Networks. (2011). *2020: Beyond 4G: Radio Evolution for the Gigabit Experience*. [Online]. Available: <http://nbn.com/sites/default/files/document/nokia-siemens-networks-beyond-4g-white-paper-online-20082011-0.pdf>.
- [7] Qualcomm Inc. (Oct. 2013). *The 1000x Data Challenge*. [Online]. Available: <http://www.qualcomm.com/solutions/wireless-networks/technologies/1000x-data>.
- [8] (Nov. 2012). *Ericsson Mobility Report On the Pulse of the Networked Society*. [Online]. Available: <http://www.ericsson.com/res/docs/2012/ericsson-mobility-report-november-2012.pdf>.

- 
- [9] Ericsson. (Jun. 2013). *5G Radio Access Research and Vision*. [Online]. Available: <http://www.ericsson.com/res/docs/whitepapers/wp-5g.pdf>.
- [10] T. L. Marzetta, *Noncooperative cellular wireless with unlimited numbers of base station antennas*, IEEE Trans. Wireless Commun., vol. 9, no. 11, pp. 3590 - 3600, Nov. 2010.
- [11] T.A. Levanen; J. Pirskanen; T. Koskela; J. Talvitie and M. Valkama, *Radio Interface Evolution Towards 5G and Enhanced Local Area Communications*, IEEE Access, vol. 2, pp. 1005 - 1029, September, 2014.
- [12] M. Majid; P. Popovski; M. Zeeshan; D. Lopez; T. Quek and M. Guizani, *IEEE Access Special Section Editorial: Physical and Medium Access Control Layer Advances in 5G Wireless Networks*, IEEE Access, vol. 5, pp. 27845 - 27849, 2017.
- [13] C.X. Wang et al., *Cellular architecture and key technologies for 5G wireless communication networks*, IEEE Commun. Mag., vol. 52, no. 2, pp. 122 - 130, Feb. 2014.
- [14] L. Atzori; I. Antonio, and G. Morabito, *The Internet of Things: A survey*, Comput. Netw., vol. 54, no. 15, pp. 2787 - 2805, Oct. 2010.
- [15] Y. Saito; Y. Kishiyama; A. Benjebbour, and T. Nakamura, *Nonorthogonal multiple access (NOMA) for cellular future radio access*, IEEE VTC Spring, pp. 1 - 5, Jun. 2013.
- [16] A. Ahmad; M. Husain; H. Tembine; O. Mahammed and A. Jamalipour, *IEEE Access Special Section Editorial: Optimization for Emerging Wireless Networks: IoT, 5G, and Smart Grid Communication Networks*, IEEE Access, vol. 5, pp. 2096 - 2100, March 2017.
- [17] A. Benjebbour; Y. Saito; Y. Kishiyama; A. Li; A. Harada and T. Nakamura, *Concept and Practical Considerations of Non-Orthogonal Multiple Access (NOMA) for Future Radio Access*, International Symposium on Intelligent Signal Processing and Communication Systems, pp. 770 - 774, Nov. 2013.

- 
- [18] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge Univ. Press, 2005.
- [19] Z. Yuan, G. Yu, and W. Li, *Multi-User Shared Access for 5G*, Telecommunications and Network Technology, vol. 5, no. 5, pp. 28 - 30, May 2015.
- [20] ITU-R M.2083-0, *IMT Vision - Framework and overall objectives of the future development of IMT for 2020 and beyond*, Sept. 2015
- [21] 3GPP TR 38.913, *Study on Scenarios and Requirements for Next Generation Access Technologies*, 3GPP Tech. Rep., V14.0.0, Oct. 2016.
- [22] M. Simsek, A. Aijaz, M. Dohler, J. Sachs, G. Fettweis, *5G-Enabled Tactile Internet*, IEEE Journal on Selected Areas in Communications, vol. 34, no. 3, pp. 460 - 473, Mar. 2016.
- [23] *The 5G Infrastructure Public Private Partnership. 5G PPP White Papers on Energy, Automotive, Factories, and eHealth Vertical Sectors*, Aug. 20, 2016.
- [24] 3GPP TR 36.881, *Study on latency reduction techniques for LTE*, 3GPP Tech. Rep., V14.0.0, Jun. 2016.
- [25] R1-167309, *Semi-persistent scheduling for 5G new radio URLLC*, 3GPP TSG-RAN WG1 86, Aug. 2016.
- [26] H. Shariatmadari, S. Iraji, R. Jntti, *Analysis of Transmission Methods for Ultra-Reliable Communications*, IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), pp. 2303 - 2308, Sept. 2015.
- [27] D. Jiang, H. Wang, E. Malkamaki, E. Tuomaala, *Principle and Performance of Semi-persistent Scheduling for VoIP in LTE System*, IEEE WiCom, pp. 2861 - 2864, Sept. 2007.
- [28] F. Capozzi; G. Piro; L.A. Grieco; G. Boggia and P. Camarda, *Downlink Packet Scheduling in LTE Cellular Networks: Key Design Issues and a Survey*, IEEE Communications Surveys & Tutorials, vol. 15, no. 2, pp. 678 - 700, Second Quarter 2013.

- 
- [29] J. Dajie; W. Haiming; E. Malkamaki and E. Tuomaala, *Principle and Performance of Semi-persistent Scheduling for VoIP in LTE System*, International Conference on Wireless Communications, Networking and Mobile Computing, WiCom 2007, pp. 2861 - 2864, 2007.
- [30] W. Haiming; J. Dajie and T. Esa, *Uplink Capacity of VoIP on LTE System*, Asia-Pacific Conference on Communications, APCC 2007, pp. 397-400, 2007.
- [31] P. Jani; K. Niko; H. Tero and M. Martti, *Persistent Packet Scheduling Performance for Voice-over-IP in Evolved UTRAN DOWnlink*, IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 1-6, 2008.
- [32] S. Jun-Bae and L. Victor, *Performance Modeling and Stability of Semi-Persistent Scheduling with Initial Random Access in LTE*, IEEE Transactions on Wireless Communications, vol. 11, no. 12, Dec. 2012.
- [33] S-E. Elayoubi; O.B. Haddada and B. Fourestie, *Performance Evaluation of Frequency Planning Schemes in OFDMA-based Networks*, IEEE Transactions on Wireless Communications, vol. 7, no. 5, May. 2008.
- [34] B. Jason; A. Nusratn; and Y.K. Jamil, *Delay Models for Static and Adaptive Persistent Resource Allocations in Wireless Systems*, IEEE Transactions on Mobile Computing, vol. 15, no. 9, September 2016.
- [35] P. Schulz; M. Matthe'; H. Klessig; M. Simsek; G. Fettweis; J. Ansari; S.A. Ashraf; B. Almeroth; J. Voigt; I. Riedel; A. Puschmann; A. Mitschele-Thiel; M. Mller; T. Elste and M. Windisch, *Latency Critical IoT Applications in 5G: Perspective on the Design of Radio Interface and Network Architecture*, IEEE Communications Magazine, vol. 55, no. 2, pp. 70 - 78, February 2017.
- [36] L. Xin, G. Xinyu, D. Xin Deng, Z. Lin and L. Wenyu, *Resource Allocation Algorithm for VoLTE with Semi-Persistent Scheduling*, 81st IEEE Vehicular Technology Conference (VTC Spring), pp. 1 - 5, May 2015.

- 
- [37] Y. Zhong; T.Q.S. Quek; and X. Ge, *Heterogeneous Cellular Networks With Spatio-Temporal Traffic: Delay Analysis and Scheduling*, IEEE Journal on Selected Areas in Communications, vol. 35, no. 6, pp. 1373 - 1386, June 2017.
- [38] R. Abreu; P. Mogensen; and K.I. Pedersen, *Pre-scheduled Resources for Retransmissions in Ultra-Reliable and Low Latency Communications*, IEEE Wireless Communications and Networking Conference (WCNC), pp. 1 - 5, 2017.
- [39] C. Li; A. Yongacoglu; and C. D'Ámours, *Downlink Coverage Probability with Spatially Non-uniform User Distribution Around Social Attractors*, 24th International Conference on Telecommunications (ICT), pp. 1 - 5, 2017.
- [40] C. Li; A. Yongacoglu; and C. D'Ámours, *Heterogeneous cellular network user distribution model*, 8th IEEE Latin-American Conference on Communications (LATINCOM), pp. 1 - 6, 2016.
- [41] R. Kirichek; M. Golubeva; V. Kulik; and A. Koucheryavy, *The home network traffic models investigation*, 18th International Conference on Advanced Communication Technology (ICACT), pp. 97 - 100, 2016.
- [42] J. Sahin, L. Liu and E. Perrins, *On the Queueing Performance of HARQ Systems with Coding over Finite Transport Blocks*, IEEE Globecom Workshops (GC Wkshps), pp. 1-7, Dec. 2015.
- [43] T. Kwon, H. Lee, S. Choi, J. Kim and D. Cho, *Design and Implementation of a Simulator Based on a Cross-Layer Protocol Between MAC and PHY Layers in a WiBro Compatible IEEE 802.16e OFDMA System*, IEEE Communications Magazine, vol. 43, no. 12, pp. 136 - 146, Dec. 2005.
- [44] *Air Interface for Broadband Wireless Access Systems*, IEEE 802.16e Rev2/D9 Std., Aug. 2009.
- [45] *Air Interface for Broadband Wireless Access Systems - Advanced Air Interface*, IEEE 802.16m/D1 Std., July 2009.

- 
- [46] 3GPP TS 36.300 v8.4.0, *Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (EUTRAN); Over All Description; Stage 2*, Mar. 2008.
- [47] 3GPP2, *Medium Access Control Layer for Ultra Mobile Broadband (UMB) Air Interface Specification*, C.S0084-002-0 v3.0, Nov. 2008.
- [48] 3GPP TSG-RAN WG2 Meeting 57, R2-070476, *Uplink Scheduling for VoIP*, Feb. 2007.
- [49] 3GPP TS 36.213 V9.0.1, *3GPP Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures*, Dec. 2009.
- [50] AA Florea, L Martinod, P Mege, and H Nguyen, *Multi-layer Realistic Voice Capacity Evaluation in LTE Rel. 9 and Performance Comparison with PMR and GSM*, In World Telecommunications Congress (WTC), pp. 1 - 6, 2012.
- [51] 3GPP TR 25.996 V6.1.0, *Spatial Channel Model for Multiple Input Multiple Output (MIMO) Simulations*, 2003-2009.
- [52] R.K. Ganti and K. Kuchi, *SINR Order Statistics in OFDMA Systems*, Global Communications Conference (GLOBECOM), pp. 4159 - 4164, 2012.
- [53] V. Kolchin, B. Sevastianov, and V. Chistiakov, *Random Allocations*, Vh Winston, 1978.
- [54] L.A. Vasquez and D. Lara *A Coordinated Operation Access Scheme for 5G Cellular Systems*, Ninth International Conference on Ubiquitous and Future Networks, pp. 604 - 606, 2017.
- [55] S. Han; C.L. I; Z. Xu and Q. Sun, *Energy Efficiency and Spectrum Efficiency Co-Design: From NOMA to Network NOMA*, IEEE COMSOC MMTTC E-Letter, vol. 9, no. 5, 2014.
- [56] H. David and H. Nagaraja, *Order Statistics*, Wiley Online Library, 1970.



- 
- [57] P. Zhang; J. Lu; Y. Wang and Q. Wang, *Cooperative Localization in 5G Networks: A Survey*, ICT Express, Available online, ISSN 2405 - 9595, pp. 27 - 32, March 2017.
- [58] J.F. Kurose, M. Schwartz, and Y. Yemini, *Controlling Window Protocols for Time-Constrained Communication in Multiple Access Networks*, IEEE Transactions on Communications, vol.: 36, no.: 1, pp. 41 - 49, Jan 1988.
- [59] S.R. Sachs, *Alternative Local Area Network Access Protocols*, IEEE Communications Magazine, vol. 26, no. 3, pp. 25 - 45, March 1988.
- [60] F.F. Kuo, *Protocols and Techniques for Data Communication Networks*, Prentice Hall, New Jersey, 1981.
- [61] W. Stallings, *Data and Computer Communications*, Macmillan Inc., New York, 1985.
- [62] M.B. Pursley, *The Role of Spread Spectrum in Packet Radio Networks*, Proceedings of IEEE, vol. 75, no. 1, pp. 116 - 34, January 1987.
- [63] D.M. Rodriguez, *Sistemas Inalámbricos de Comunicacin Personal*, S.A. Marcombo, 2002.
- [64] Report ITU-R M.2370-0, *IMT traffic estimates for the years 2020 to 2030*, 2015.
- [65] Report ITU-R M.2320-0, *Future technology trends of terrestrial IMT systems*, 2014.
- [66] Recommendation ITU-R M.1457-13, *Detailed specifications of the terrestrial radio interfaces of International Mobile Telecommunications-2000 (IMT-2000)*, 2017.
- [67] Recommendation ITU-R M.1645, *Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000*, 2003.
- [68] H. Jin, C. Cho, N. Song, and D. Sung, *Optical Rate Selection for Persistent Scheduling with HARQ in Time-Correlated Nakagami-m Fading Channels*, IEEE

- Transactions on Wireless Communications, vol. 10, no. 2, pp. 637 - 647, Feb. 2011.
- [69] C.E Uc-Rios and D. Lara-Rodriguez, *An ON-OFF Synchronous Sectored Operation for Packet Data Transmission in Cellular Systems*, Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th, pp. 1142 - 1146, 2007.
- [70] J. Castaneda-Camacho; C.E. Uc-Rios and D. Lara-Rodriguez, *Reverse Link Erlang Capacity of Multiclass CDMA Cellular System Considering Nonideal Antenna Sectorization*, IEEE Trans. Vehicular Technology, vol. 52, no. 6, pp. 1476 - 1488, Nov. 2003.
- [71] F. Yong; P. Lunden; M. Kuusela and M. Valkama, *Efficient Semi-Persistent Scheduling for VoIP on EUTRA Downlink*, IEEE 68th Vehicular Technology Conference, VTC 2008-Fall , pp. 1 - 5, 2008.
- [72] 3GPP TR 25.913 V7.3.0, March 2006.
- [73] 3GPP TR 25.814 V7.1.0, Sep 2009.
- [74] 3GPP TSG-RAN WG2 Meeting 57bis, R2-071227, *Number of Control Symbols*, St. Julians, Malta, March, 2007.
- [75] 3GPP TSG RAN WG1 Meeting 47bis, R1-070098, *Persistent Scheduling in E-UTRA*, Sorrento, Italy, January 2007.
- [76] 3GPP TSG RAN WG2 54, R2-062164, *Uplink Resource Allocation Scheme*, Tallinn, Estonia, August - Sept 2006.
- [77] 3GPP TSG RAN WG1 Meeting 47, R1-063275, *Discussion on Control Signaling for Persistent Scheduling of VoIP*, Riga, Latvia, November 2006.
- [78] S. N. Datta and S. Kalyanasundaram, *Optimal Power Allocation and User Selection in Non-Orthogonal Multiple Access Systems*, IEEE Wireless Communications and Networking Conference, pp. 1 - 6, 2016.
- [79] J. Gozalvez, *Fifth-Generation Technologies Trials*, IEEE Vehicular Technology Magazine, vol. 11, no. 2, pp. 5 - 13, 2016.