



Universidad Politécnica de Cartagena

Escuela Técnica Superior de Ingeniería de Telecomunicaciones

Proyecto Fin de Carrera

**Diseño De Sistemas De Subasta Dinámica
De Espectro En Radio Cognitiva**

Luis Pastor González

Director: Juan José Alcaraz Espín

Septiembre de 2011



Autor	Luis Pastor González
E-mail del autor	luispastor14@gmail.com
Director(es)	Dr. Juan José Alcaraz Espín
E-mail del director	juan.alcaraz@upct.es
Título del PFC	Diseño De Sistemas De Subasta Dinámica De Espectro En Radio Cognitiva
Descriptor	Software-defined Radio, Cognitive Radio, Spectrum Trading, Markov Decision Process, Dynamic Spectrum Access.
Resumen	
<p>En un sistema de radio cognitiva, el acceso dinámico al espectro (<i>Dynamic Spectrum Access, DSA</i>) hace referencia a un protocolo mediante el cual se permite a usuarios no licenciados o “secundarios” el uso temporal de una parte del espectro asignado a usuarios con licencia o “primarios”.</p> <p>En este trabajo se presenta una triple formulación: MDP, optimización mediante programación lineal y simulación, para modelar un escenario basado en un protocolo DSA de implementación centralizada con acceso basado en subasta, cuyo objetivo principal es la obtención del máximo rendimiento económico posible manteniendo la menor probabilidad de bloqueo para los usuarios primarios.</p>	
Titulación	Ingeniero de Telecomunicación
Departamento	Tecnologías de la Información y las Comunicaciones
Fecha de presentación	26 de septiembre de 2011

Índice

Abstracto.....	4
1. Introducción.....	5
1.1. Radio cognitiva. Conceptos básicos.....	7
1.2. DSA: Dynamic Spectrum Access.....	8
1.3. Perspectivas y posibles aplicaciones.....	11
1.4. Estandarización.....	11
1.5. Objetivos.....	12
2. Antecedentes y trabajos relacionados.....	15
2.1. Presentación de la Radio Cognitiva. J. Mitola y F. Capar.....	15
2.2. El proyecto de Timo Weiss y Friedrich Jondral.....	17
2.3. El proyecto XG de DARPA.....	18
2.4. El proyecto CORVUS.....	19
3. Modelo propuesto.....	22
3.1. Aproximación mediante MDP.....	26
3.1.1. Introducción a los Procesos de Decisión de Markov (MDP).....	29
3.1.2. Parámetros del escenario.....	33
3.1.3. Caracterización del modelo propuesto mediante un MDP.....	35
3.2. Optimización mediante programación lineal.....	50
3.2.1. Introducción a la programación lineal.....	50
3.2.2. Caracterización del modelo propuesto mediante programación lineal.....	52
3.3. Caracterización mediante simulación.....	60
3.3.1. Introducción a la Simulación por Eventos Discretos.....	61
3.3.2. Caracterización del modelo propuesto mediante simulación.....	63
3.3.3. Análisis estadístico de los resultados.....	70
4. Resultados.....	74
4.1. Escenarios propuestos.....	74
4.1.1. Escenario 1.....	75
4.1.2. Escenario 2.....	78
4.1.3. Escenario 3.....	80
4.1.4. Escenario 4.....	83
5. Conclusiones.....	86
Bibliografía.....	87

ABSTRACTO.

La **radio cognitiva (Cognitive Radio, CR)** está considerada como una solución prometedora a la ineficiencia existente en el uso del espectro radioeléctrico en las redes de acceso inalámbrico como consecuencia del esquema tradicional de asignación.

En un sistema de radio cognitiva, el **acceso dinámico al espectro (Dynamic Spectrum Access, DSA)** hace referencia a un protocolo mediante el cual se permite a usuarios no licenciados o “secundarios” el uso temporal de una parte del espectro asignado a usuarios con licencia o “primarios”. El objetivo de un protocolo DSA es permitir que los usuarios secundarios ocupen los huecos no utilizados de dicho espectro de una forma óptima, es decir, causando la menor disminución posible en la calidad de servicio (por ejemplo, probabilidad de bloqueo) de los usuarios primarios.

Los protocolos DSA tienen dos implementaciones: la **distribuida** y la **centralizada**. En la implementación centralizada existe un controlador principal que actúa como árbitro en la asignación, recolectando información sobre el uso del espectro para tomar en cada momento la decisión más apropiada a la situación, siempre con ciertas restricciones. En esta aproximación también puede existir un **proceso de subasta** para fijar, en tiempo real, un precio a las bandas del espectro disponibles para los usuarios secundarios. Por otro lado, en la implementación distribuida, son los usuarios secundarios quienes se encargan de “observar” el uso que se hace del espectro en cada momento y deciden si pueden o no entrar en el sistema.

Dado que estas decisiones tienen lugar en sistemas dinámicos (por ejemplo en redes con usuarios que entran y salen continuamente del sistema), la **programación dinámica** se presenta como una herramienta muy útil para llevar a cabo una formulación rigurosa de un problema de decisión DSA. Además, el proceso de ocupación del espectro, como otros muchos sistemas de comunicaciones, se puede caracterizar mediante un modelo Markoviano que, aplicado a la formulación de la programación dinámica, da lugar a un problema de **Procesos de Decisión de Markov (Markov Decision Process, MDP)**.

En este trabajo se presenta una triple formulación: MDP, programación lineal y simulación, para modelar un escenario basado en un protocolo DSA de implementación centralizada con acceso basado en subasta, cuyo objetivo principal es la obtención del máximo rendimiento económico posible manteniendo la menor probabilidad de bloqueo para los usuarios primarios.

1. INTRODUCCIÓN.

En la actualidad hay una creencia común de que existe una crisis en la disponibilidad de frecuencias que pueden ser económicamente explotadas. Esta idea surge de la sobrecarga existente en las frecuencias por debajo de los 3 GHz. A frecuencias superiores, como muestra la figura 1, existe un uso real muy bajo del espectro. La captura muestra una utilización del 0.5 % en la banda entre 3 y 4 GHz, que decrece hasta el 0.3 % en la banda entre 4 y 5 GHz. Por tanto, esta concepción de escasez de espectro parece contradictoria con los resultados obtenidos en la medición.

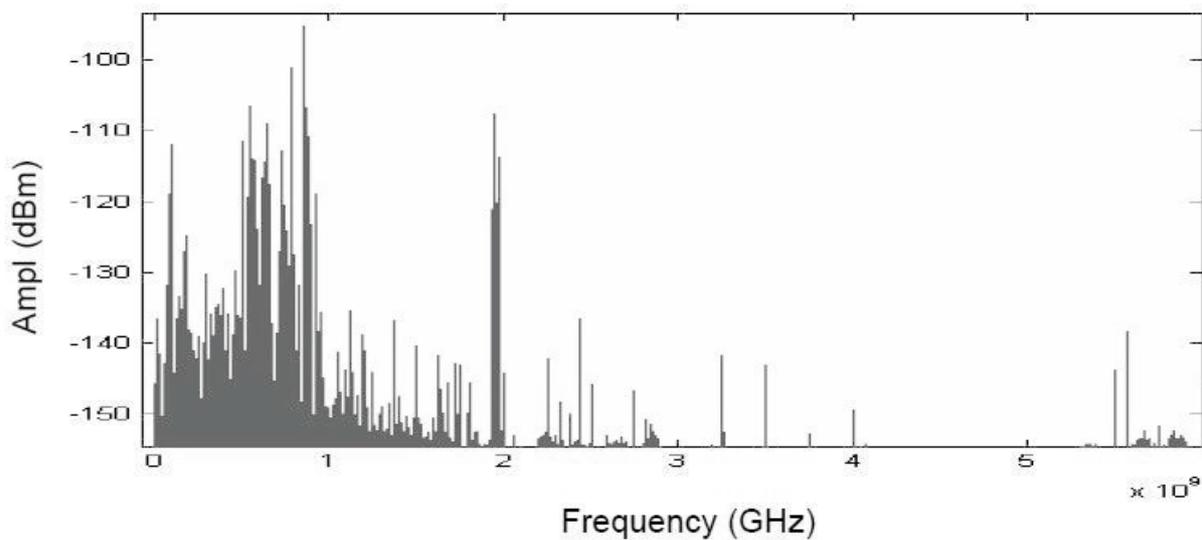


Figura 1: Captura del espectro hasta 6 GHz en un área urbana

La figura 1 corresponde a una medición tomada en Estados Unidos. Si se compara esta imagen con la asignación frecuencial establecida en dicho país (figura 2), se aprecia una discrepancia entre la asignación legal y el uso real del espectro [1].

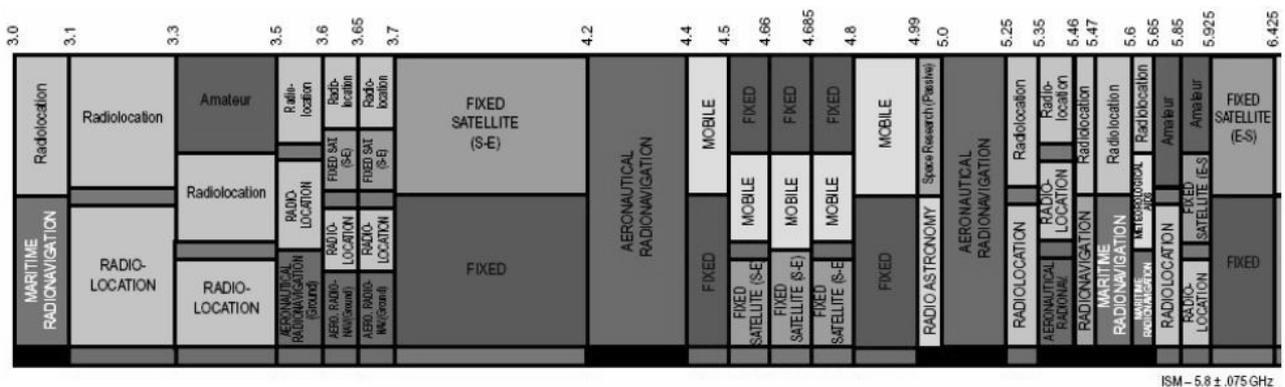


Figura 2: Asignación frecuencial en Estados Unidos

Esta diferencia entre la asignación y el uso real del espectro ocurre del mismo modo en el resto del mundo, y es la razón por la que es necesario un nuevo paradigma en la asignación del espectro. En España y, en general, en la Unión Europea, existe ya un conocimiento de este problema. El secretario de Estado español de Telecomunicaciones, Francisco Ros, ya adelantó en 2010 [2] la necesidad de un “uso más eficiente del espectro” radioeléctrico como “única forma de avanzar hacia un mercado único digital”.

En la actualidad, existe una demanda cada vez mayor de grandes anchos de banda por parte de los usuarios en sistemas de comunicaciones inalámbricas. Sin embargo, el espectro radioeléctrico es un recurso limitado al que los operadores obtienen acceso mediante un sistema de licencias basado en la asignación del ancho de banda por parte de una autoridad pública.

Hoy en día esta asignación es estática e inflexible, es decir, un determinado ancho de banda sólo puede ser utilizado por un operador y sus clientes (usuarios primarios -PU- o con licencia; *Licensed Users -LU-*). Se definen como **Usuarios Primarios**, aquellas entidades que poseen legalmente alguna banda frecuencial. No obstante, mientras ciertas bandas frecuenciales se usan prácticamente en su totalidad en ciertas áreas y durante ciertos intervalos de tiempo, también es cierto que otras regiones del espectro se ven a menudo infrutilizadas. Se da así una situación paradójica: mientras las limitaciones espectrales propician el desarrollo de nuevas aplicaciones *wireless*, existen partes del espectro no ocupadas (“huecos” u “oportunidades” en el espectro). Estos **huecos en el espectro** se definen como bandas de frecuencia asignadas a usuarios con licencia pero no utilizadas en determinados instantes y localizaciones, y que por tanto pueden ser aprovechadas por usuarios secundarios. Los **Usuarios Secundarios** se definen como aquellas entidades que pretenden hacer uso de los canales no utilizados en la banda de los LUs [3].

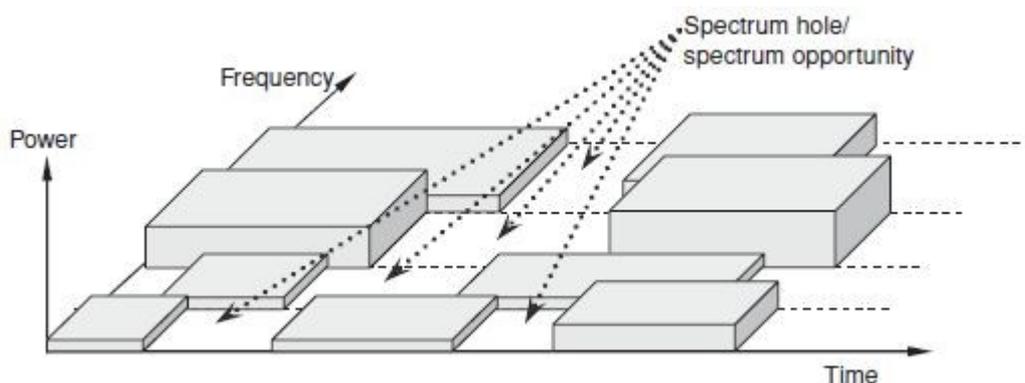


Figura 3: "Huecos" espectrales. Canales no utilizados por los usuarios primarios.

Ésta es una propiedad clásica de muchos sistemas de telecomunicaciones, en los que la intensidad de tráfico varía considerablemente a lo largo de un día, y constituye la base para los protocolos que se presentan en este trabajo.

Una parte de la solución se observa precisamente en la figura 1, donde existe un uso considerable en la banda por encima de los 5 GHz. Esto es debido a se trata de la banda UNII (*Unlicensed National Information Infrastructure*), una banda libre, al igual que la banda ISM (*Industrial, Scientific and Medical*) de 2.4 GHz, cuyas restricciones de uso son mínimas (hacen referencia a la potencia de transmisión, pero no se adjudican a un único propietario) desde el punto de vista regulatorio.

Surge así la necesidad de extender el uso no licenciado del espectro a otras bandas frecuenciales, aunque éstas sean licenciadas, de forma que se mantenga la calidad de servicio de los usuarios licenciados al mismo tiempo que se da respuesta a las necesidades de acceso de los usuarios sin licencia. Una aproximación para alcanzar este objetivo es el desarrollo de un sistema radio que sea capaz de detectar el uso espectral a lo largo de una banda concreta de frecuencias y hacer uso de los canales disponibles sólo si la comunicación no interfiere con un usuario con licencia. Este sistema se conoce con el nombre de Radio Cognitiva.

1.1. Radio cognitiva. Conceptos básicos.

El término **Radio Cognitiva** (CR, *Cognitive Radio*) hace referencia al conjunto de tecnologías cuyo objetivo es el incremento de la eficiencia en el uso del espectro radioeléctrico. La radio cognitiva proporciona mecanismos para que los usuarios sin licencia (secundarios, *Secondary Users -SU-*) puedan acceder a esas oportunidades o huecos que aparecen en las bandas licenciadas. Estos mecanismos se basan en la **radio definida por software (*software-defined radio*)**, que se trata de un sistema inalámbrico de comunicación cuyos parámetros (frecuencia, esquema de modulación, etc.) se pueden controlar de forma dinámica, en tiempo real, en función de las necesidades de la comunicación en un momento dado. Las principales características de la radio definida por software son la operación multibanda y el soporte multi-estándar, multi-servicio y multi-canal [4].

En cuanto a la nomenclatura, para los usuarios sin licencia se usarán indistintamente los términos “usuarios secundarios” o “SU's”; y de forma análoga “usuarios primarios” o “LU's” para los usuarios con licencia.

Evidentemente, esta cesión del espectro a usuarios secundarios se debe llevar a cabo teniendo presente en todo momento la minimización del impacto sobre la calidad de servicio de los usuarios primarios. Para ello, los usuarios de estos sistemas deben implementar tres funcionalidades básicas:

1. **Detección y análisis espectral:** un nodo tiene que tener la capacidad de ser consciente del entorno que le rodea, mediante un sondeo del espectro y/o la comunicación con otros nodos que posean información sobre la utilización espectral para poder detectar los patrones de uso de la banda de interés por parte de los usuarios primarios.
2. **Movilidad espectral:** el nodo debe ser capaz de reconfigurar su sistema de transmisión para adaptarse al entorno observado de forma imperceptible al usuario con licencia. Esta capacidad de reconfiguración es la que se describía antes como radio definida por software. La figura 4 representa esta característica.
3. **Compartición de recursos espectrales:** una red de radio cognitiva debe, además de respetar a los operadores con licencia, compartir los recursos disponibles con otros sistemas de radio cognitiva para conseguir una utilización espectral que sea equitativa y justa para todos los usuarios. Una vez que el usuario secundario pueda usar algún canal, llevará a cabo el acceso a la banda frecuencial en cuestión, siguiendo algún protocolo de acceso al medio, MAC (*Medium Access Control*) que permita la comunicación entre usuarios no licenciados con un impacto mínimo (pocas colisiones) sobre el resto de usuarios.

1.2. DSA: *Dynamic Spectrum Access*.

Dentro de los sistemas de radio cognitiva es frecuente el término “**Acceso Dinámico al Espectro**” o ***Dynamic Spectrum Access (DSA)***. Se trata del conjunto de mecanismos que gestionan el uso del espectro como respuesta a los cambios del sistema (canales disponibles, solicitudes de usuarios secundarios...) de acuerdo a ciertos objetivos (por ejemplo, maximizar el uso del espectro) y sujeto a ciertas restricciones (por ejemplo, minimizar la probabilidad de bloqueo de los usuarios primarios).

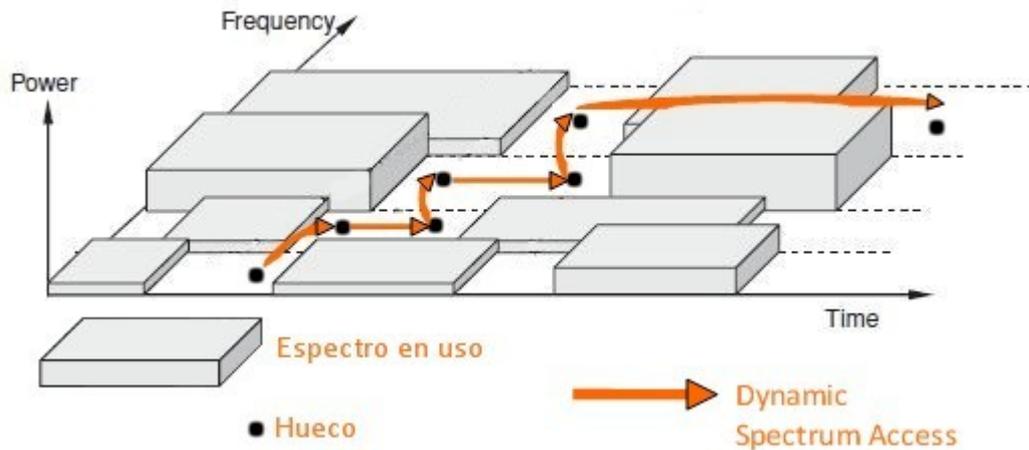


Figura 4: Acceso Dinámico al Espectro

Los modelos de DSA para radio cognitiva tienen dos clasificaciones. Por un lado, según el acceso que tengan los usuarios secundarios, existen tres tipos [5]:

1. **DSA de uso exclusivo.** En él, un usuario primario puede conceder a un usuario secundario el derecho al acceso exclusivo al espectro.
2. **DSA de uso compartido.** En él, los usuarios secundarios acceden al espectro cuando tienen la oportunidad, siempre evitando interrumpir a los usuarios licenciados. Dentro de este tipo de modelos existen dos aproximaciones, como muestra la figura 5:
 1. **Overlay.** En este caso el usuario accede a los huecos espectrales que dejan los usuarios primarios, por lo que deben llevar a cabo un proceso previo de monitorización para evitar la interferencia con otros usuarios.
 2. **Underlay.** En este caso, un usuario secundario puede transmitir al mismo tiempo que uno primario. Sin embargo, la potencia de transmisión del secundario estará limitada para que las interferencias estén por debajo de un umbral preestablecido. Para ello se suele hacer uso tecnología UWB.
3. **Modelos comunes.** En ellos, los usuarios secundarios pueden acceder libremente al espectro.

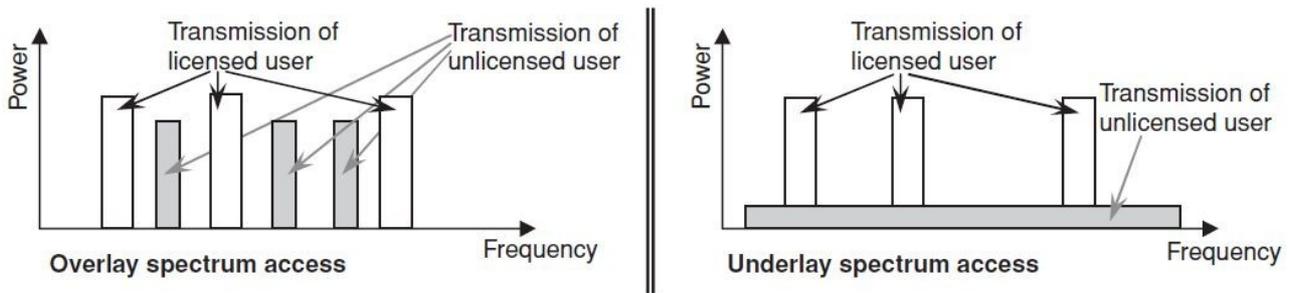


Figura 5: DSA de uso compartido. Izquierda: *overlay*. Derecha: *underlay*.

La segunda clasificación depende de la gestión que exista para el acceso:

1. En la **implementación centralizada**, existe una autoridad central que arbitra en la asignación de canales frecuenciales. Para ello, se encarga de recopilar la información de la red relativa a la ocupación espectral y a las necesidades de los usuarios secundarios para tomar la decisión más oportuna en cada momento. Generalmente esta decisión se basa en la solución de algún problema de optimización, como se verá en apartados posteriores.
2. Por otro lado, en la **implementación distribuida**, son los usuarios secundarios quienes toman las decisiones de acceso de forma autónoma según la información que ellos mismos se encargan de obtener del sistema. En comparación con la implementación centralizada, el esquema de decisión distribuido requiere una mayor cantidad de recursos computacionales en el terminal de usuario y, en general, la solución alcanzada no suele ser la óptima. No obstante, la implementación distribuida implica mayor cantidad de tráfico.

Los protocolos MAC para DSA pueden incluir también **protocolos de negociación por el espectro**. En situaciones donde exista una infrautilización del espectro, el operador licenciado para un determinado ancho de banda del mismo puede ofertar los canales disponibles a los usuarios secundarios. Para llevar a cabo esta tarea en tiempo real, es necesario que el protocolo utilizado implemente un sistema de negociación entre el propietario del espectro y los usuarios secundarios basado en parámetros como el coste del acceso (en términos económicos) y el tiempo de estancia en el sistema. Existen varios modelos de acceso dinámico basados en este tipo de sistemas.

1.3. Perspectivas y posibles aplicaciones.

En el marco regulatorio ya existe una concienciación de la necesidad de nuevas tecnologías para un uso eficiente de los recursos frecuenciales. Existen estudios de la FCC y la SPTF (*Spectrum Policy Task Force*) que evidencian importantes variaciones temporales y geográficas en el uso del espectro, con un rango de entre el 15% y el 85% [6]. Para poder hacer uso de estos “espacios”, la FCC norteamericana presentó en 2003 un documento oficial (*Notice of Proposed Rule Making* [7]) donde se presentaba a la radio cognitiva como un fuerte candidato para la implementación de la negociación y compartición del espacio espectral.

Las principales aplicaciones donde la FCC pretende hacer uso de la radio cognitiva son:

- Incrementar la potencia en transmisión en hasta 8 dB en áreas con baja densidad de población y reducido uso del espectro (áreas rurales). Los dispositivos de CR deberían ser capaces de percibir esta infrutilización e incrementar la potencia de transmisión en consecuencia.
- Cesión interrumpida del espectro a un usuario secundario por parte de un primario. El usuario secundario debe tener la capacidad de determinar cuál es el espectro no utilizado y reconocer las tramas de asignación y revocación del servicio de un usuario primario.
- Compartición coordinada y dinámica del espectro usando las capacidades de conocimiento temporal y espacial de los usuarios.
- Facilitar la interoperabilidad entre sistemas usando sistemas de CR para recibir y transmitir usando diferentes esquemas de modulación, codificación y corrección de errores.

1.4. Estandarización.

Dado que la gestión eficiente del espectro mediante sistemas de radio cognitiva implica muchos aspectos técnicos y económicos, la estandarización de estos sistemas es muy importante. Los procesos de estandarización son necesarios para el desarrollo e implementación de las redes de radio cognitiva [5].

No obstante, hasta ahora han surgido diferentes grupos que han trabajado individualmente, dando lugar a menudo a resultados incoherentes. Para resolver este problema surgió el **IEEE**

Standards Coordinating Committee (SCC) 41 dentro del programa *Next Generation Radio and Spectrum Management*, aglutinando una serie de componentes relacionados dentro del estándar **IEEE 1900** [8].

El objetivo del IEEE SCC 41 es solventar cualquier problema relacionado con el desarrollo, la implementación y el despliegue de la nueva generación de acceso radio y la gestión eficiente del espacio frecuencial. Este organismo está compuesto por cuatro grupos de trabajo y un grupo de estudio, siendo cada uno de ellos responsable de iniciar los procesos de estandarización de distintos aspectos de la radio cognitiva. Los principales componentes del estándar IEEE 1900 son los siguientes:

- **IEEE 1900.1.** Los objetivos principales de este estándar son los de **identificar y explicar los términos y conceptos** relacionados con la gestión del espectro, la radio definida por software, la radio adaptativa y otras tecnologías similares. Se usa como conexión entre el resto de grupos de trabajo del SCC 41, ya que todos hacen uso de las definiciones recogidas en éste.
- **IEEE 1900.2.** Este estándar se centra en las **recomendaciones sobre las interferencias y coexistencia de usuarios**. En radio cognitiva existen muchos dispositivos inalámbricos en el sistema, por lo que la optimización de sus parámetros de funcionamiento es crucial.
- **IEEE 1900.3.** Este grupo se centra en los procesos de **evaluación de los distintos componentes** de los sistemas de radio definida por software.
- **IEEE 1900.4.** Estudia las **propiedades de reconfiguración** de los sistemas para poder adaptarse a entornos *wireless* heterogéneos. Esta heterogeneidad será una característica clave en los sistemas inalámbricos de nueva generación, donde los dispositivos móviles deberán usar diferentes tecnologías al mismo tiempo.
- **IEEE 1900.A.** Es el grupo de estudio del SCC 41. Se encarga de la **certificación de los dispositivos** basados en protocolos de acceso dinámico al espectro (DSA).

1.5. Objetivos.

Ya se han expuesto las limitaciones existentes en los recursos radioeléctricos en los sistemas inalámbricos en general. Sin embargo existe una demanda de servicios *wireless* que

crece exponencialmente. No sólo aumenta el número de usuarios, sino también el ancho de banda que requieren las nuevas aplicaciones como video-telefonía o acceso inalámbrico a Internet. La forma de dar respuesta a esta demanda dada las limitaciones presentes en el espacio radioeléctrico constituye una de las principales áreas de investigación en el campo de sistemas de comunicaciones inalámbricas. La asignación de recursos y su optimización constituyen métodos generales para mejorar el funcionamiento de estas redes.

Como se detallará en apartados posteriores, este trabajo se centra en la formulación de problemas relacionados con la radio cognitiva, como la optimización desde el punto de vista de la asignación de recursos. Se presenta un modelo DSA centralizado de puja y subasta, en el que los usuarios secundarios ofrecen una cantidad máxima de dinero por el uso del espectro licenciado, que es propiedad de un único operador.

En el modelo descrito, **el objetivo principal será analizar el compromiso existente entre la calidad de servicio de los usuarios primarios y el beneficio económico derivado del arrendamiento del espectro a usuarios secundarios.**

Para un estudio riguroso del escenario descrito, se llevará a cabo una formulación del mismo basada en Procesos de Decisión de Markov (*Markov Decision Process, MDP*), en la optimización mediante programación lineal y la simulación por eventos discretos.

Cuando se enfrentan dos o más objetivos opuestos (en este caso calidad de servicio de los primarios frente a beneficio económico) en un problema de optimización, no existe una solución óptima, estrictamente hablando. El resultado tras el proceso de optimización será un frente de Pareto, es decir, un conjunto de puntos que relacionarán los dos objetivos buscados de tal forma que no será posible mejorar uno de ellos sin empeorar el otro.

La figura 6 muestra un frente de Pareto genérico. El área de color azul (T) representa el conjunto de puntos que relacionan ambos objetivos de forma no óptima, por ejemplo, P3 representa una relación entre f_1 y f_2 con un valor en el eje de ordenadas mayor que el que representa P1, con un valor menor del objetivo f_2 para un mismo valor de f_1 . La frontera inferior del área T (línea curva que une los puntos P1 y P2) es el frente de Pareto propiamente dicho, y une los puntos que representan la relación óptima entre ambos objetivos de forma que la mejora del primer objetivo (eje X) pasa por una disminución del segundo (eje Y).

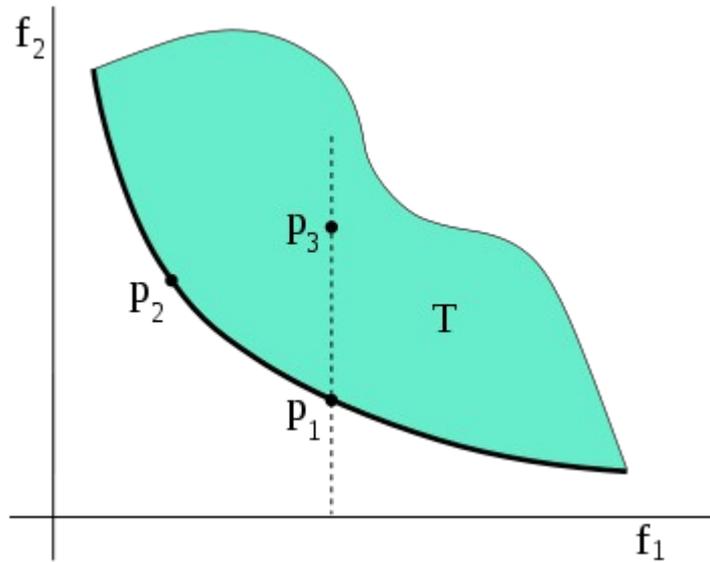


Figura 6: Frente de Pareto genérico.

La estructura del resto del documento es la siguiente:

- En el segundo apartado se presenta un resumen de la evolución de la radio cognitiva, desde su presentación hasta la implementación de algunos de los proyectos más ambiciosos en relación con la mejora de la eficiencia espectral.
- El tercer apartado presenta con todo detalle el modelo de radio cognitiva propuesto para el presente proyecto. Se subdivide en tres grandes apartados, en cada uno de los cuales se describe el proceso de caracterización de dicho modelo desde los tres puntos de vista ya comentados: Procesos de Decisión de Markov, optimización mediante programación lineal y simulación.
- En el cuarto apartado se presentan los escenarios propuestos para el modelo y los resultados obtenidos en cada uno de ellos mediante los tres procedimientos.
- En el quinto y último apartado se exponen las conclusiones obtenidas.

2. ANTECEDENTES Y TRABAJOS RELACIONADOS.

2.1. Presentación de la Radio Cognitiva. J. Mitola y F. Capar.

La idea fundamental que aparece tras el término “Radio Cognitiva” fue presentada por primera vez en 1999, en un artículo escrito por Joseph Mitola III y Gerald Q. Maguire [9], y definida en su totalidad un año después, en el 2000, también por J. Mitola III, en su tesis doctoral “*Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio*” [10]. La definición exacta que sugirió para los sistemas de radio cognitiva fue ésta:



Figura 7: J. Mitola III

“El término radio cognitiva identifica el punto donde los asistentes digitales personales (PDAs) inalámbricos y las redes relacionadas son lo suficientemente inteligentes desde el punto de vista computacional sobre los recursos radioeléctricos y las comunicaciones entre máquinas para:

- (a) detectar las necesidades de comunicación de los usuarios en función del entorno, y*
- (b) proporcionar los recursos radioeléctricos y servicios inalámbricos que más se ajusten a dichas necesidades.”*

Así, un usuario de CR es capaz de seleccionar de forma automática la mejor oportunidad para una transmisión, e incluso puede retardar o adelantar ciertas transmisiones en función de los recursos utilizados y previstos. La mayoría del trabajo de la tesis de Mitola [10] se centra en las capacidades de aprendizaje y razonamiento de los nodos de CR para alcanzar este objetivo.

No obstante, la idea de un sistema con la capacidad de seleccionar el mejor modo de funcionamiento de entre múltiples variantes técnicas, usando diferentes bandas de frecuencia, ya había sido introducida por P. Fettweis [11], en 1996, y M. Bronzel [12], en 1997, con el nombre de IBMS (*Integrated Broadband Mobile System*), donde también se presentaba el concepto de un canal de señalización universal para dar soporte a dicho sistema.

Aunque el término Radio Cognitiva fue definido originalmente como una extensión de la

radio definida por software, existe actualmente una redefinición por parte de la FCC (*Federal Communications Commission*) [7]:

“La radio cognitiva (CR) es un sistema radio capaz de cambiar sus parámetros de transmisión basándose en la interacción con el entorno en el que opera. La mayoría de sistemas de CR serán SDRs (*Software Defined Radios*), pero ni el funcionamiento software ni la programación en tiempo real son requisitos de un sistema CR”.

De forma implícita en la definición e implementación de un sistema de radio cognitiva, se asume un alto grado de flexibilidad, dado que el entorno radio es muy variable a causa de las interferencias y otras variaciones en los canales.

El concepto de “**Spectrum Pooling**”, que se puede traducir como “reparto de espectro” (entre usuarios primarios y secundarios) se introdujo por vez primera en el artículo de J. Mitola “*Cognitive radio for flexible mobile multimedia communications*” [13], en 1999. El espectro para redes móviles es limitado, pero en un momento dado, multitud de frecuencias están infrautilizadas. En el mismo artículo se introducía la posibilidad de que los usuarios con licencia pudieran ceder la porción no utilizada de espectro a los usuarios secundarios.

En 2002, F. Capar [14] reincidió en este concepto, definiéndolo como “una estrategia de compartición de recursos donde los usuarios con licencia tendrían la máxima prioridad”. Es decir, una vez que un usuario primario accediera a la banda de frecuencia, todos los usuarios secundarios deberían dejarla libre, dándole la prioridad al LU. Según Capar, existen dos posibilidades para que un usuario primario acceda al espectro:

1. Buscar frecuencias sin utilizar dentro del rango de frecuencias con licencia. En este caso el usuario primario tendría el derecho de reclamar los canales de frecuencias usados por los SU operando en dicha banda. Esta aproximación requiere que el usuario con licencia sea capaz de detectar a los usuarios secundarios e incluso comunicarse con ellos. Se asume además que el usuario primario no tiene por qué hacer uso de la totalidad del espectro disponible, por lo que estaría dispuesto a compartirlo bajo ciertas restricciones.
2. En la segunda aproximación, el usuario con licencia no tiene ningún conocimiento de los usuarios secundarios. En consecuencia, simplemente solicita algún canal de frecuencia dentro de su rango de operación, obligando a que cualquier usuario secundario que lo esté usando se cambie a otro canal disponible.

En ambas aproximaciones los usuarios secundarios tienen que monitorizar el espectro para detectar las frecuencias no utilizadas antes de obtener el acceso a los recursos. Además esta monitorización debe ser continua para poder detectar los usuarios primarios que van apareciendo en el sistema. En la primera aproximación el usuario primario podría informar al secundario de su intención de usar un determinado canal, eliminando así la necesidad de monitorización continua.

En función de las características físicas del espectro, Mitola [13] definió cuatro bandas espectrales denominadas “*pools*”.

1. *Very Low Band* (26.9 – 399.9 MHz)
2. *Low Band* (404 – 960 MHz)
3. *Mid Band* (1390 – 2483 MHz)
4. *High Band* (2483 – 5900 MHz)

Según los requisitos de ancho de banda, distancia de propagación y otras características de tráfico se usaría una u otra banda.

Capar [14], definió *pool* como “un espectro contiguo que puede ser utilizado por procesos de alquiler”. Fue Mitola quien introdujo la idea de alquilar el espectro no usado a los usuarios secundarios. En este caso sería necesaria la introducción de ciertas autoridades de gestión del espectro encargadas de administrar las bandas no utilizadas y su precio, en función del ancho de banda, la localización, el nivel de interferencias, etc. Con este esquema, los usuarios secundarios podrían programar sus transmisiones en función de la urgencia, los recursos disponibles y su precio. Por ejemplo, un email escrito durante un trayecto en taxi no se enviaría usando tecnología GPRS (que sería bastante cara). En su lugar, su envío se pospondría para cuando existiera una red LAN basada en radio cognitiva con un coste menor que el de una red móvil.

2.2. El proyecto de Timo Weiss y Friedrich Jondral.

La Universidad de Karlsruhe, en Alemania, ha desarrollado una arquitectura basada en radio cognitiva y reparto de espectro. Esta arquitectura, definida en 2004 por T. Weiss y F. Jondral [15], usa un esquema centralizado basado en OFDM. Weiss y Jondral partieron de un escenario basado en un punto de acceso parecido al que utiliza 802.11 (WiFi), consistente en una estación base y usuario móviles, ambos con capacidades de radio cognitiva, asumiendo que no hay cambios en los sistemas de usuarios primarios. El uso de OFDM tiene la ventaja de que es posible “llenar” de ceros las subportadoras utilizadas por los usuarios primarios, por lo que no se emite

ninguna potencia en éstas. Además, OFDM permite el análisis de la actividad espectral de los LUs sin ningún coste adicional, ya que se usa la FFT de todas formas.

En esta implementación, para detectar a los usuarios primarios, la estación base emite periódicamente, mediante *broadcast*, unas tramas llamadas “tramas de detección”, que son tramas que no contienen ningún dato. Durante ese periodo, todos los usuarios móviles llevan a cabo la monitorización del espectro, y toda la información obtenida debe recolectarse en la estación base, lo cual podría llevar un tiempo considerable usando técnicas tradicionales de acceso al medio.



Figura 8: Friedrich K. Jondral

En este aspecto se utiliza un protocolo llamado “*Boosting protocol*”, que omite la capa MAC y usa únicamente la capa física para la señalización. Todos los terminales móviles modulan un símbolo complejo a máxima potencia en aquellas subportadoras donde haya aparecido un usuario primario, es decir, sólo en aquéllas que no estuvieran ocupadas por un LU anteriormente. De esta forma, si todos los SUs hacen esto simultáneamente, la estación base recibirá una señal amplificada en las portadoras con nuevos LUs, con una alta fiabilidad en la detección. Aunque este mecanismo provoca interferencias a los usuarios primarios, se lleva a cabo durante un periodo lo suficientemente corto como para poder despreciarlas sin que afecte a la calidad de servicio.

Weiss y Jondral [15] también investigaron el problema de la interferencia mutua. Cada sistema interfiere en el otro, dado que la resolución espectral de la FFT realizada no es perfecta. La mejor solución para esta interferencia está en no utilizar las subportadoras adyacentes a aquéllas usadas por los usuarios primarios, dando lugar a cierta reducción en el ancho de banda disponible para el sistema de alquiler.

2.3. El proyecto XG de DARPA.

Para mejorar el uso del espectro radioeléctrico, DARPA (*Defense Advanced Research Projects Agency*) inició el programa *neXt Generation* (XG), en el que propuso una aproximación mucho más genérica que la de Weiss y Jondral con la intención de definir una arquitectura basada en las llamadas “radios inteligentes basadas en políticas” (*policy based intelligent radios*) [16]. En el proyecto XG existen sistemas radio definidos por software, que se pueden controlar mediante

ciertas políticas. Los dispositivos XG son capaces de (a) monitorizar el entorno para determinar cuál es el espectro libre, (b) usar ciertas restricciones basadas en la política usada para repartir los canales disponibles dentro de lo que se conoce como un dominio XG, formado por un grupo de terminales con las mismas características, y (c) determinar los mecanismos necesarios para acceder a dichos canales. Se deja abierta la posibilidad de implementar estos tres mecanismos de forma centralizada o distribuida.

2.4. El proyecto CORVUS.

CORVUS es un sistema de radio cognitiva desarrollado por profesores de las universidades de Berkeley (California, USA) y Berlín (Alemania) en 2004 [1]. Esta aproximación a la radio cognitiva parte de dos premisas básicas:

1. Abundancia de canales espectrales, disponibles para el uso de los usuarios secundarios, y
2. Los SUs usan técnicas de radio cognitiva para evitar las interferencias con los usuarios primarios presentes en el sistema.



Figura 9: Robert W. Broderick.
Desarrollador de CORVUS.

En este caso, los usuarios primarios no son conscientes de la existencia del sistema CR, por lo que no existe intercambio de información entre ambos tipos de usuarios, es decir, en CORVUS, los usuarios con licencia no señalizan de ninguna forma a los secundarios para el acceso al espectro. Los usuarios sin capacidades de CR se tratan como ruido en el sistema.

En esta red heterogénea, los SUs tienen un acceso restringido a la banda con licencia. Éstos pueden usarla siempre y cuando no hay un LU en ella. Un usuario primario, LU_x , puede tolerar una interferencia máxima de Δt_x unidades de tiempo. Pasado ese tiempo, el LU debe tener algún canal disponible para la comunicación, lo que significa que todos los SU interferentes deberán liberarlo en dicho periodo de tiempo.

Evidentemente es fundamental que un usuario secundario sea capaz de detectar la presencia de los usuarios primarios (al menos cada Δt_x). Esta detección se basa en las llamadas **Primary User Footprints (PUF)**, una “marca” o “huella” que indica la presencia de un LU. Las

PUF incluyen información sobre la asignación de frecuencias de la FCC [6], además de información como características locales de los LUs (tiempo medio de servicio, horas pico, etc.).

Inicialmente, antes de cualquier actividad, un SU debe monitorizar el espectro en busca de usuarios con licencia para asegurarse que no interfiere en su comunicación. Además deberá realizar esta tarea periódicamente para dejar libre el espectro tan pronto como detecte la presencia de un LU.

En el modelo propuesto en CORVUS, los usuarios secundarios forman los llamados “**Grupos de Usuarios Secundarios**” (**SUG, Secondary User Groups**) para coordinar su comunicación. Los miembros de un SUG pueden comunicarse entre sí mediante un esquema ad-hoc o mediante un punto de acceso dedicado. No se contempla la comunicación entre usuarios de diferentes SUG.

En cualquier caso, se asume que la comunicación entre SUs o entre SU y punto de acceso es siempre unicast, existiendo la posibilidad de que la gestión de las comunicaciones sea distribuida o centralizada (con una estación concreta actuando como autoridad).

Existen dos **patrones de tráfico** básicos en los usuarios secundarios:

1. Tráfico web, donde los usuarios secundarios buscan acceder a internet, lo cual implica la existencia de una estación base o punto de acceso que proporcione el acceso a la Red. Por ello, se trata de una situación apta para una aproximación basada en una arquitectura centralizada.
2. Redes ad-hoc, que abarca cualquier tipo de tráfico ad-hoc, sin asumir ninguna infraestructura predeterminada, generalmente con el propósito de establecer comunicación entre usuarios de un SUG. Se trata de un patrón óptimo para una aproximación basada en una arquitectura distribuida.

CORVUS opera en un amplio rango de frecuencias, desde decenas de MHz a varios GHz, creando una “banda libre virtual” con unas propiedades similares a las bandas ISM y UNII. En este rango de frecuencias se definen las ya mencionadas **Spectrum Pools (SP)** como un rango de frecuencias (no necesariamente contiguas) usadas por un SUG. Las SP de diferentes SUGs se pueden solapar, por lo que habrá usuarios de diferentes SUGs compitiendo por los mismos recursos. Cada SP se divide en n **subcanales**.

La figura 10 muestra la idea principal para el reparto de espectro en CORVUS. Los usuarios primarios poseen distintas partes del espectro, pero pueden estar inactivos en determinados momentos. Las zonas rojas indican que un LU está usando su espectro y, por tanto, un SU no podrá utilizarlo. La figura muestra tres comunicaciones activas de usuarios secundarios. Para cada comunicación, una pareja de usuarios secundarios elige una combinación de subcanales para formar un “**Enlace de Usuarios Secundarios**” (SUL, *Secondary User Link*).

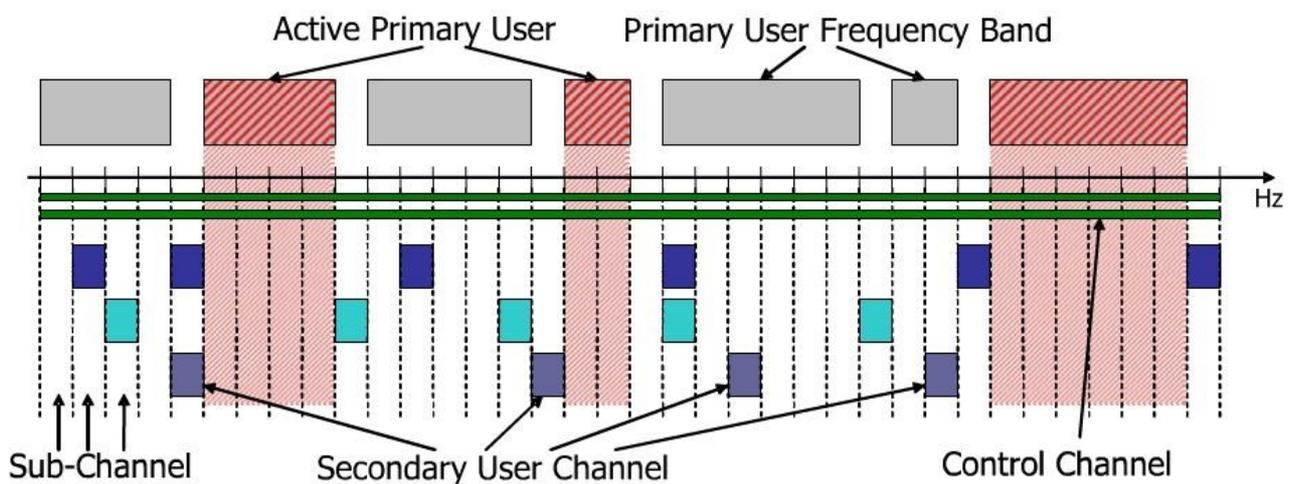


Figura 10: Reparto de espectro en CORVUS

Los usuarios secundarios, además utilizan canales lógicos dedicados para el intercambio de información de control y monitorización. En CORVUS se establecieron dos tipos de canales lógicos de control: un “**Canal de Control Universal**” (UCC, *Universal Control Channel*), y los “**Canales de Control de Grupos**” (GCCs, *Group Control Channels*). El UCC es único, y debe ser conocido por cualquier SU a priori, en otro caso, éste no podrá comunicarse con los demás. El propósito principal del UCC es el de anunciar la existencia de los grupos existentes y permitir la incorporación de nuevos usuarios a los mismos. Cada SUG cuenta, además, con un GCC para el intercambio de información entre sus miembros.

Estos canales de control transportan una cantidad limitada de información con una tasa de datos relativamente baja, y pueden ser:

- a) canales localizados dentro del espectro con licencia para este propósito,
- b) canales localizados en las bandas libres (ISM / UNII), o
- c) canales UWB (*Ultra Wide Band*).

3. MODELO PROPUESTO.

Como se indicaba en la introducción, en el sistema de asignación del espectro radioeléctrico que existe en la actualidad, son las administraciones públicas las encargadas de ceder dicho espectro a los distintos operadores. Generalmente, los operadores obtienen las bandas del espectro a través de un proceso de subasta pública. Este esquema de asignación se conoce con el nombre de “**mercado primario**”.

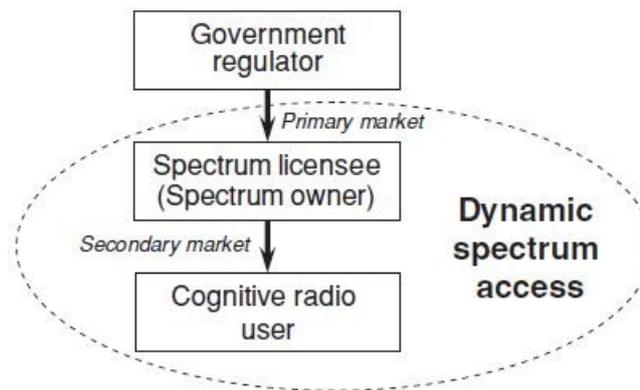


Figura 11: Mercados primario y secundario

Sin embargo, la creciente demanda de espectro y la existencia de franjas o canales no utilizados pone de manifiesto la ineficiencia de este método de asignación. Una forma práctica y viable de resolver esta ineficiencia es dotar a los propietarios del espectro (operadores) de la posibilidad de “vender” o “alquilar” las porciones no utilizadas del ancho de banda del que disponen, dando lugar a lo que se conoce como “**mercados secundarios**”. A diferencia de esquema de asignación primario, en los mercados secundarios la cesión del ancho de banda a usuarios sin licencia o secundarios se lleva a cabo en tiempo real.

En el modelo propuesto, la implementación DSA (*Dynamic Spectrum Access*) estará **centralizada**, es decir, existirá una entidad que permitirá o denegará el acceso de los usuarios secundarios al espectro libre de acuerdo a una determinada política, definida posteriormente. Esta entidad será, habitualmente, el operador propietario de la banda frecuencial.

Además, y como se comentó en la introducción, se trata de un **modelo de puja y subasta**, es decir, existirán diferentes tipos de usuarios secundarios, que estarán definidos por la máxima cantidad de dinero dispuestos a pagar por el acceso.

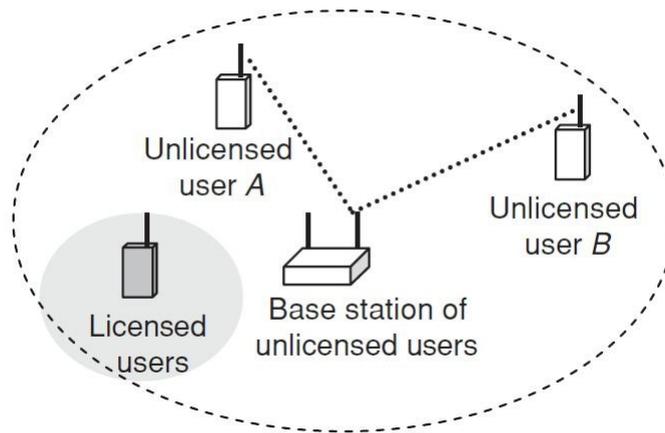


Figura 12: DSA centralizado

Así, será el operador con licencia quien, en función de la situación del sistema (canales o huecos disponibles, usuarios secundarios ya en el sistema...) establecerá la cantidad mínima dispuesto a aceptar, es decir, actuará como subastador del espectro. Esta cantidad, evidentemente será mayor (más restrictiva) a medida que el número de usuarios en el sistema aumente y, por tanto, exista un riesgo mayor de perjudicar a los usuarios primarios.

En relación con los sistemas de tarificación en sistemas celulares y de radio cognitiva existe mucha literatura. Por ejemplo, en el trabajo de J. Hou [17], el precio impuesto para el uso del espectro se utiliza para controlar la tasa de llegada. Si el precio elegido por el operador es mayor que el habitual, un usuario secundario podría decidir no realizar la llamada. Como resultado, la solicitud de llamada (o el servicio del que se trate) se pospondrá con la esperanza de que el precio exigido disminuya. Se trata de una aproximación muy similar a la utilizada en el modelo propuesto, descrita en el párrafo anterior, y en la que las llegadas de usuarios secundarios decrecerán a medida que aumente el coste del acceso.

En los modelos basados en subasta se enfatiza el aspecto económico del modelo DSA estudiado. Por otro lado, existen modelos basados únicamente en la optimización del servicio (maximizar las transferencias de usuarios secundarios minimizando la interferencia con los primarios). En estos modelos se antepone el aspecto técnico al económico. Evidentemente el aspecto económico en los sistemas de radio cognitiva es muy importante, ya que supone el incentivo necesario para que los usuarios primarios estén dispuestos a ceder el espectro que poseen.

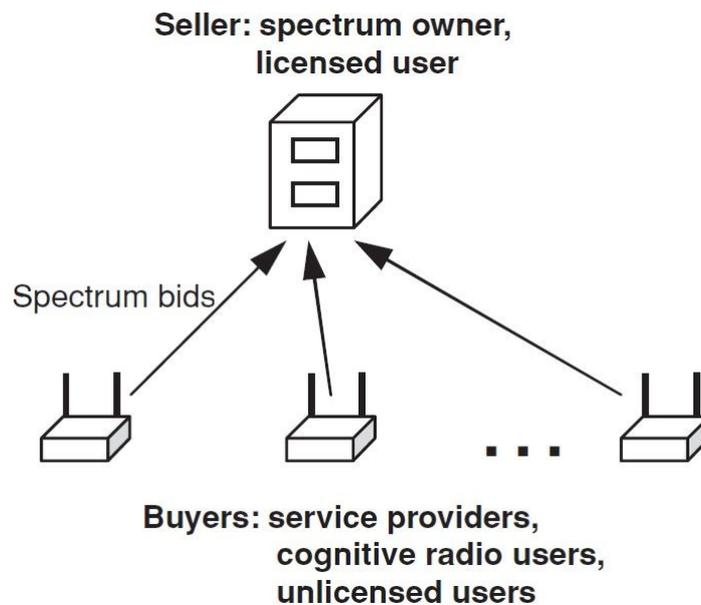


Figura 13: Modelo basado en subasta

Esta cantidad exigida estará determinada por el **control**. El término control determina, en un momento dado (definido por la cantidad de usuarios en el sistema), si es posible que un usuario no licenciado acceda al sistema en función de la ocupación del mismo y de la “puja” del solicitante. Por tanto habrá distintos controles, siendo unos más restrictivos que otros, lo que se traduce en que la puja necesaria para acceder al espectro será mayor en los controles más restrictivos. Es decir, con cada usuario que acceda al sistema, se determinará un nuevo control, que normalmente será más restrictivo que el anterior para poder mantener un equilibrio óptimo entre el beneficio que se obtiene del acceso y su impacto en la calidad de servicio de los usuarios primarios.

Además, en el sistema modelado **siempre tendrán prioridad los usuarios primarios**, lo que significa que su acceso al espectro no se verá directamente afectado por el control existente en ese momento. Si un usuario primario llega al sistema y existe algún canal disponible, tendrá acceso al mismo. La única forma de que la llegada de un usuario licenciado sea rechazada es que, en el momento de la solicitud, todos los canales estén ocupados. Por ello, **no se expulsa a los usuarios secundarios** que hayan accedido al espectro. Un usuario sin licencia sólo abandonará el sistema cuando su servicio finalice. Además, se asume que cada usuario entrante (de cualquier tipo) únicamente puede ocupar un canal.

Se define la **política** del sistema como el conjunto de controles que se aplican para todas las posibles combinaciones de ocupación de usuarios primarios y secundarios. Cada una de estas combinaciones se define como un **estado** del sistema y, por tanto, habrá un control para cada

estado. Además existen dos grandes grupos de políticas: las estacionarias y las no estacionarias.

En una **política estacionaria**, existe una asociación única entre un estado y un control, es decir, en un estado genérico i -ésimo, siempre se utilizará el mismo control. Por otro lado, en una **política no estacionaria**, esta asociación no es tan estricta y, en un estado cualquiera, cada control tendrá asociada una probabilidad, de forma que en un mismo estado se podrán tomar distintos controles.

En este Proyecto la política que se buscará será de tipo estática, siendo el objetivo fundamental del mismo encontrar la **política óptima**, es decir, un mapeo control-estado tal que se obtenga el mejor compromiso entre los objetivos buscados (maximizar el beneficio económico minimizando el bloqueo de los usuarios primarios), o lo que es lo mismo, se busca obtener el **Frente de Pareto**, ya descrito en la introducción.

La mayor dificultad en la búsqueda de esta política óptima consiste en que se debe optimizar no sólo el rendimiento esperado en el estado en el que se tome la decisión, sino también el rendimiento esperado en los estados futuros a los que se llegue como consecuencia de la decisión tomada.

Se entiende por **escenario** el conjunto de parámetros que definen el sistema propuesto: tasas de llegada de usuarios primarios y secundarios, tiempo medio de estancia en el sistema de ambos, número de canales del sistema y ponderación entre ambos objetivos (el peso de cada uno de ellos en el proceso de optimización).

La búsqueda y validación de la política óptima se llevará a cabo mediante **tres aproximaciones diferentes**, cuyos resultados se compararán y verificarán en diferentes escenarios:

1. Aproximación mediante Procesos de Decisión de Markov (MDP).
2. Optimización mediante programación lineal.
3. Validación mediante simulación.

Las tres aproximaciones se codificarán en MATLAB, un software matemático que ofrece un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio (lenguaje "M").

Cada una de estas aproximaciones se detalla en los siguientes subapartados.

3.1. Aproximación mediante MDP.

Una **cadena de Markov** es esencialmente un proceso estocástico en el que, en cada instante de tiempo, t , una variable aleatoria, X_t , tiene un valor entero aleatorio. La característica fundamental de las cadenas de Markov es que la distribución de probabilidad de la variable aleatoria en el instante siguiente X_{t+1} depende del valor de la variable en el instante actual, X_t . El valor de esta variable en cada momento se define como **estado**, existiendo una correspondencia entre el estado de la cadena de Markov y el estado del sistema que se modele (usuarios en el sistema).

Dado que el número de canales en el espectro es limitado, el conjunto de valores que podrá tomar X_t es finito, y se asumirá además que la distribución de probabilidades del estado siguiente sólo depende del valor de X_t y es totalmente independiente de t .

El rango de aplicación de este tipo de modelos es muy amplio. Básicamente incluye cualquier sistema dinámico cuya evolución en el tiempo implique aleatoriedad y sea posible una enumeración clara de los estados que puede presentar el sistema. En el caso particular del presente Proyecto, el sistema de acceso al espectro mediante radio cognitiva contará con diferentes estados claramente definidos por la ocupación que exista en cada momento y la aleatoriedad vendrá dada por las llegadas y tiempos de estancia en el sistema de los usuarios. Por tanto, su caracterización mediante un modelo de Markov, con probabilidades de transición entre estados, será relativamente sencilla.

Según cómo sea la evolución entre estados, las cadenas de Markov se clasifican en dos grandes grupos: las cadenas de Markov en tiempo continuo y en tiempo discreto.

En las **cadenas de Markov en tiempo discreto (DTMC)** el estado cambia en instantes discretos de tiempo que se identifican por una variable entera, n . En cada periodo o **etapa**, el estado del sistema se denota por X_n , que pertenece a un conjunto finito, S , de posibles estados, denominado **espacio de estados**. En general se asume que este espacio de estados es finito, o sea, $S = \{1, 2, \dots, m\}$, para cierto entero positivo m . Este tipo de cadenas de Markov se describen mediante sus probabilidades de transición p_{ij} : cuando el estado es i , hay una probabilidad p_{ij} de que el siguiente estado sea j . Matemáticamente se expresa así:

$$p_{ij} = P\{X_{n+1} = j | X_n = i\}, \quad i, j \in S \quad (1)$$

Una suposición clave en la definición de las DTMC es que las probabilidades de transición

p_{ij} se aplican siempre que se visita el estado i , sin importar qué sucedió en el pasado ni cómo se alcanzó el estado i . A esta característica se le denomina **propiedad de Markov**, cuya expresión matemática es la siguiente:

$$P\{X_{n+1}=j|X_n=i, X_{n-1}=i_{n-1}, \dots, X_0=i_0\}=P\{X_{n+1}=j|X_n=i\}=p_{ij} \quad (2)$$

Para todos los instantes n , y todos los estados $i, j \in S$, y todas las posibles secuencias i_0, \dots, i_{n-1} , de estados anteriores. Es decir, la distribución de probabilidad de X_{n+1} depende del pasado únicamente en lo que éste está representado por el estado actual X_n .

Además, las probabilidades de transición p_{ij} deben ser no negativas y sumar uno:

$$\sum_{j=1}^m p_{ij}=1, \forall i \quad (3)$$

En general, se permite que la probabilidad p_{ii} sea positiva, en cuyo caso el siguiente estado es el mismo que el actual.

Todos los elementos de una cadena de Markov quedan “codificados” en la **matriz de probabilidades de transición**, que puede verse como una matriz que, en cada una de sus filas, contiene una función de distribución de probabilidad para el estado siguiente. La primera fila es la distribución del estado siguiente al 1, la segunda es la del estado siguiente a 2, etc.

$$P = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1m} \\ p_{21} & p_{22} & \dots & p_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ p_{m1} & p_{m2} & \dots & p_{mm} \end{bmatrix}$$

En estos sistemas también resultan muy útiles los **grafos de transición de probabilidad**, cuyos nodos son los estados y cuyos arcos son las posibles transiciones. Los arcos pueden asociarse a sus correspondientes probabilidades, p_{ij} , permitiendo visualizar el modelo entero.

Por otro lado, en las **cadena de Markov en tiempo continuo (CTMC)**, las transiciones entre estados no se dan en intervalos fijos de tiempo, sino que los estados evolucionan en tiempo continuo.

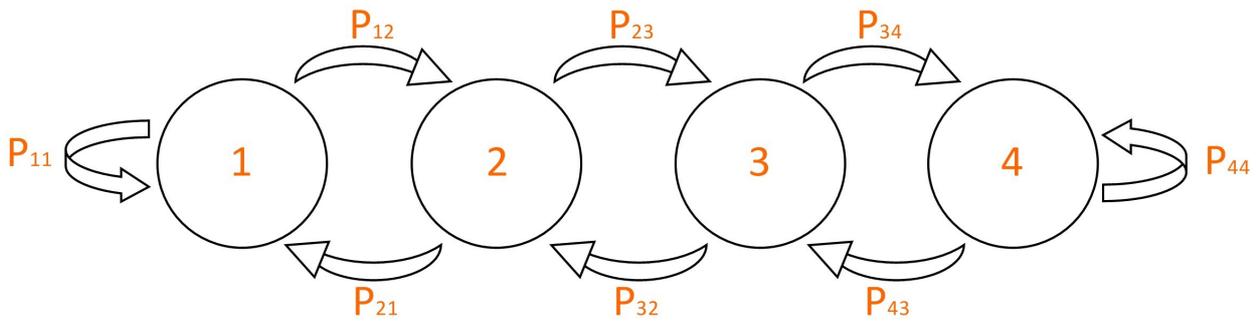


Figura 14: Grafo de transición de probabilidad genérico

En estos procesos, las transiciones se producen de acuerdo a ciertas probabilidades de transición, pero los tiempos de permanencia en cada estado se modelan como **variables aleatorias continuas**.

Para describir el proceso en tiempo continuo, se introducen dos variables aleatorias de interés:

- X_n : El estado justo después de la transición n-ésima.
- T_n : El tiempo transcurrido desde la transición (n-1) y la n-ésima.

También se adoptan las siguientes suposiciones:

- Si el estado actual es i , el tiempo hasta la siguiente transición está distribuido exponencialmente con parámetro v_i y es independiente de la historia anterior del proceso y del estado siguiente.
- Si el estado actual es i , el estado siguiente será j con probabilidad p_{ij} , independientemente de la historia anterior y del tiempo de permanencia en i .

Con estas suposiciones, tenemos que:

$$P\{X_{n+1}=j, T_{n+1} \geq t | X_n=i\} = P\{X_{n+1}=j | X_n=i\} P\{T_{n+1} \geq t | X_n=i\} = p_{ij} e^{-v_i t} \quad (4)$$

Y el tiempo esperado para la siguiente transición es:

$$E\{T_{n+1} | X_n=i\} = \int_0^{\infty} t v_i e^{-v_i t} dt = \frac{1}{v_i} \quad (5)$$

Se puede interpretar v_i como el número medio de transiciones desde el estado i hacia otros estados. Por tanto, v_i se denomina **tasa de salida** del estado i . Puesto que sólo una fracción p_{ij} de dichas transiciones se producen hacia el estado j , se define la **tasa de transición entre i y j** como:

$$q_{ij} = p_{ij} v_i \quad (6)$$

Representando el número medio de transiciones entre i y j por unidad de tiempo de permanencia en i . De las definiciones anteriores se deduce que, dadas las tasas de transición q_{ij} , se pueden obtener las tasas v_i usando la fórmula:

$$v_i = \sum_{j=1}^m q_{ij} \quad (7)$$

Y las probabilidades de transición usando la fórmula:

$$p_{ij} = \frac{q_{ij}}{v_i} \quad (8)$$

Aunque el modelo permite transiciones de i hacia i (auto-transiciones), el efecto no es observable por la ausencia de memoria de la distribución exponencial. Se asume entonces que:

$$p_{ii} = 0 \quad (9)$$

3.1.1. Introducción a los Procesos de Decisión de Markov (MDP).

Aunque con un modelo simple basado en cadenas de Markov se puede analizar el funcionamiento de sistemas basados en eventos discretos, se trata de técnicas esencialmente descriptivas, es decir, que permiten evaluar el rendimiento de un sistema que opera con unas normas de funcionamiento preestablecidas. No obstante, en este Proyecto el objetivo es precisamente diseñar esas normas de funcionamiento del sistema para que su rendimiento sea el mejor posible, o sea, obtener una **política óptima**.

La base de las técnicas que permiten hallar la política óptima es la **programación dinámica**. Se trata de una extensa teoría que data de los años 60 y cuya importancia ha ido creciendo progresivamente en paralelo con la capacidad de cómputo de los ordenadores. Se

basa, principalmente, en la resolución de un problema muy complejo mediante la combinación de las soluciones de una serie de subproblemas de menor complejidad [5]. Su aplicación en el ámbito de las comunicaciones también ha ido en aumento desde que, a mediados de los 70, pioneros de la telemática como Dimitri Bertsekas [18] o Leonard Kleinrock [19] aticipasen su potencial para la ingeniería de protocolos.

Además, la programación dinámica hace uso de unos algoritmos muy simples y de fácil aplicación, no sólo a problemas de redes, sino a todo tipo de casos prácticos. Evidentemente, bajo la simplicidad de los mismos subyace un sofisticado análisis matemático que dota de rigor a estas técnicas. La programación dinámica aplicada al presente Proyecto se describirá con mayor detalle en apartados posteriores.

El citado modelo Markoviano, aplicado a la formulación basada en la programación dinámica, da lugar a lo que se conoce como **Procesos de Decisión de Markov (MDP)**. Se trata de sistemas de eventos discretos en los que, en cada etapa del sistema, se puede actuar sobre alguna variable capaz de influir en la evolución futura del sistema, permitiendo cierto grado de control sobre su rendimiento. Si el sistema se modela mediante una cadena de Markov, la variable de control, denominada genéricamente como u , determinará las probabilidades de la matriz de transición. Habrá, por tanto, una función $P(u)$, que devolverá una matriz de transición por cada valor de u .

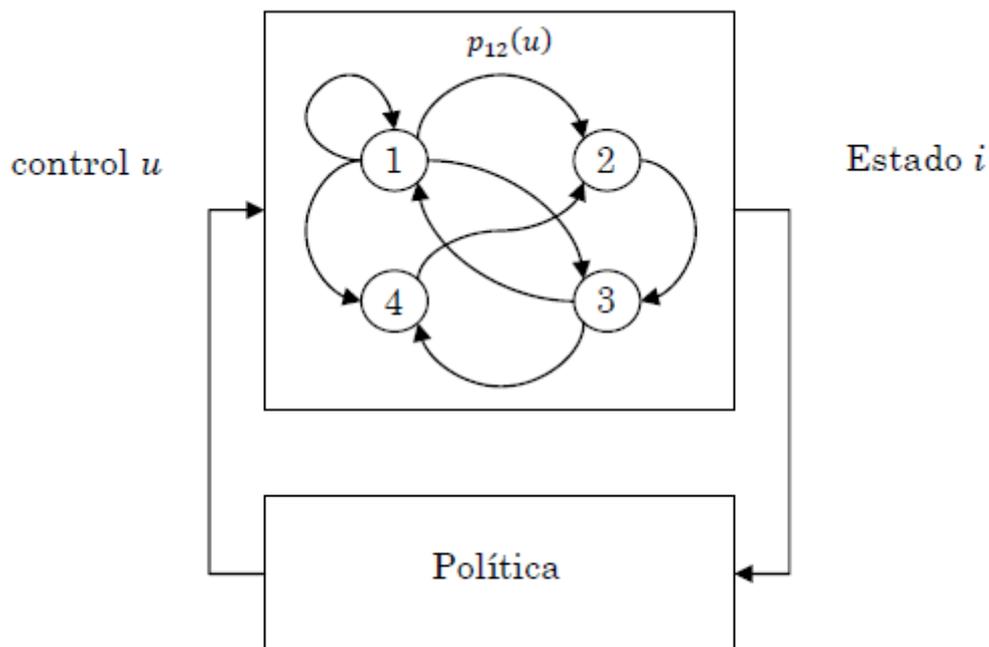


Figura 15: MDP

En los MDPs se deben asociar los estados (o las transiciones) a un determinado coste o beneficio, siendo el objetivo último, su minimización o maximización, respectivamente, es decir, hallar la política óptima.

Para ilustrar la gran utilidad de los métodos basados en MDP, a continuación se citan algunos casos de uso en el campo de las comunicaciones y en algunos otros:

- Asignación de recursos en transmisión inalámbrica, Calidad de Servicio, etc [20].
- **Detección del espectro radio disponible en radio cognitiva** [21].
- Control de admisión y enrutamiento en redes 802.16 (WiMAX) [22].
- Control de admisión y asignación de recursos en servidores web [23].
- Ajuste dinámico de los precios de los servicios en redes en función de la congestión [24].
- Asignación de precios para *stock options* [25].
- Desarrollo de estrategias de inteligencia artificial para videojuegos [26].
- Desarrollo de algoritmos de soporte para el diagnóstico médico [27].

Elementos de los MDPs.

A continuación se describen los elementos básicos de los Procesos de Decisión de Markov y se hace una breve introducción a la notación utilizada:

En los MDP, se parte de un sistema de estados discretos estocástico, del que se asume que existe un conocimiento en todo momento de su estado, es decir, su estado es “observable”, y cuyo espacio de estados S , es “contable”.

Un MDP es un sistema de este tipo en el que se definen tres elementos:

1. Las **acciones de control**, que se toman cada vez que se produce una transición,
2. los **costes** asociados a las acciones y a las transiciones entre estados o a los estados y
3. las **probabilidades de transición**, cuyo valor puede depender de las acciones de control.

Dado un estado $i \in S$ del sistema, la acción de control u sobre ese estado se debe seleccionar de un conjunto finito de estados U , es decir, $u \in U$.

De acuerdo a estos conceptos, se puede reescribir la definición de política como la regla que indica qué acción de control debe ser seleccionada en cada etapa. Las políticas consideradas

para este trabajo tienen dos propiedades: (1) no se seleccionan aleatoriamente y (2) son políticas estacionarias.

A partir de esta nomenclatura, una política estacionaria es un mapeo del espacio de estados, S , al conjunto de acciones de control posibles, U , es decir, es una función de la forma $u=\mu(i)$, con $i \in S$. En muchas ocasiones, sólo un subconjunto de las acciones de control de U están permitidas en el estado i . Dicho subconjunto, representado como U_i , contiene las acciones “admisibles” en i . Es decir, una política será “admisibles” si $\mu(i) \in U_i$ para todo $i \in S$. El símbolo π representa generalmente cualquier tipo de política (estacionaria o no) y el símbolo μ suele emplearse para políticas estacionarias.

Una vez seleccionada la acción de control $\mu(i)$, la función de distribución de probabilidad del próximo estado j queda configurada como $p_{ij}(\mu(i))$, para todo $j \in S$. Por tanto, una política μ determina toda la matriz de transición de la DTMC.

Otro elemento fundamental de las DTMC es que cada estado i se asocia con un coste (o beneficio) que depende, en general, de i y de la acción de control u . Se suele emplear la función $g(i,u)$ para representar dicho coste. En muchas ocasiones los costes (beneficios) se asocian a las transiciones entre estados. Este tipo de coste se representará mediante la función $\tilde{g}(i,u,j)$, que proporciona el coste de usar u en i y pasar al estado j . Dada esta función, se empleará el **coste definido por unidad de tiempo**, obtenido de la siguiente forma:

$$g(i, u) = \sum_{j \in S} p_{ij}(u) \tilde{g}(i, u, j) \quad (10)$$

El horizonte del problema es el número de etapas que se debe considerar en la resolución del mismo. Por tanto, si el número de etapas es N , para un determinado estado inicial $x_0 \in S$ se define el coste de una política π de la siguiente forma:

$$J_{\pi}(x_0) = E \left\{ \sum_{k=0}^{N-1} a^k g(x_k, u_k) \right\} \quad (11)$$

Donde x_k representa el estado en la etapa $k = 0, 1, \dots, N-1$; u_k es la decisión de control tomada a partir de la política π admisible, aunque no necesariamente estacionaria, y a es un escalar positivo tal que $0 < a \leq 1$ denominado factor de descuento, que se emplea en muchas ocasiones para reducir la importancia relativa de los costes/beneficios de futuras etapas frente a las etapas más inmediatas. La operación $E\{\cdot\}$ corresponde al cálculo del valor esperado, o

esperanza matemática.

El coste representado anteriormente corresponde a un problema de **horizonte finito**. En los problemas de **horizonte infinito**, el coste total se define así:

$$J_{\pi}(x_0) = \lim_{N \rightarrow \infty} E \left\{ \sum_{k=0}^{N-1} a^k g(x_k, u_k) \right\} \quad (12)$$

El interés de este Proyecto se centra en estudiar el comportamiento asintótico del modelo Markoviano, es decir, las probabilidades de transición cuando $n \rightarrow \infty$, siendo n el intervalo de tiempo durante el que se observa el sistema. En este caso se habla de **estado estacionario** del sistema.

La minimización del coste total tiene sentido sólo si $J_{\pi}(x_0)$ es finito pero, en muchos problemas de horizonte infinito, como el que se describe, es posible que $J_{\pi}(x_0) = \infty$ para todas políticas y todos los posibles estados iniciales. Por ello se define el **coste medio por estado** de la siguiente forma:

$$CTMC: \quad \lim_{N \rightarrow \infty} \frac{1}{E\{t_N\}} E \left\{ \int_0^{t_N} g(x(t), u(t)) dt \right\} \quad (13)$$

$$DTMC: \quad \lim_{N \rightarrow \infty} \frac{1}{N} E \left\{ \sum_{k=0}^{N-1} g(x_k, u_k) \right\} \quad (14)$$

3.1.2. Parámetros del escenario.

Cada escenario formulado estará definido por los siguientes parámetros:

- λ_{LU} : Tasa de llegada de los usuarios primarios al sistema. Representa el número medio de llegadas al sistema por unidad de tiempo. Sigue una distribución de Poisson.
- λ_{SU} : Tasa de llegada de los usuarios secundarios al sistema. Representa el número medio de llegadas al sistema por unidad de tiempo. Sigue una distribución de Poisson.
- μ_{LU} : Tasa de salida de los usuarios primarios del sistema. Representa la frecuencia con la

que los usuarios primarios abandonan el sistema, o de forma equivalente, da una idea de su tiempo medio de estancia en el sistema.

- μ_{su} : Tasa de salida de los usuarios secundarios del sistema. Representa la frecuencia con la que los usuarios secundarios abandonan el sistema, o de forma equivalente, da una idea de su tiempo medio de estancia en el sistema.

En cuanto a la relación entre estas tasas, como se indicaba en la introducción al modelo propuesto, se trata de un sistema con rechazo, particularmente con una cola de longitud nula. Esto quiere decir que cuando un usuario (de cualquier tipo) llega al sistema, bien accede a un canal disponible o bien se rechaza. Por tanto con independencia de cómo sean las tasas de llegada con respecto a las de salida, **el sistema siempre será estable**. Simplemente variará la cantidad de tráfico rechazado.

- **N** : Número de canales disponibles en el espectro. Número máximo de usuarios admisibles en el sistema.
- **Tipos de usuarios secundarios**. Los usuarios secundarios del escenario a modelar se clasificarán en distintas clases en función de la cantidad de dinero que estén dispuestos a ofrecer por el acceso. Su ordenación irá de mayor a menor cantidad ofrecida. Así, si se parte de un escenario con tres tipos de SUs, los usuarios de tipo 1 ofrecerán una cantidad mayor que los usuarios de tipo 2, quienes a su vez ofrecerán una cantidad mayor que los usuarios de tipo 3.
- **Precios ofrecidos por cada tipo de usuario**. Por simplicidad, se asumirá que los precios ofrecidos estarán dentro de un conjunto de valores: $B = \{b_1, b_2, \dots, b_m\}$, expresados en coste por unidad de tiempo.
- **Probabilidad de cada tipo de usuario secundario, P_j** . En un sistema de este tipo, no tiene por qué existir la misma cantidad de usuarios secundarios de cada tipo. Esto se modela con el parámetro P_j , que representa la probabilidad de que, cuando exista una llegada de un SU, éste sea de tipo j . Evidentemente, se debe cumplir que

$$\sum_{j=1}^m P_j = 1 \quad (15)$$

- **Controles.** Este parámetro dependerá directamente de los tipos de usuarios secundarios que existan. Concretamente habrá tantos controles como clases de SUs, también identificados por enteros positivos, más el control “0”, que denegará el acceso a cualquier usuario no licenciado. Los controles también estarán ordenados de mayor a menor nivel de restricción, como muestra la siguiente tabla:

Control	Nivel de restricción
0	• No se permite el acceso de ningún usuario secundario
1	• Sólo se permite el acceso de usuarios secundarios de tipo 1.
2	• Sólo se permite el acceso de usuarios secundarios de tipos 1 y 2.
...	
n	• Sólo se permite el acceso de usuarios secundarios entre 1 y n .

Tabla 1: Tipos de controles del sistema

3.1.3. Caracterización del modelo propuesto mediante un MDP.

El primer paso para la formulación del problema será establecer una asociación o **mapeo entre el estado del sistema**, que se identificará con un entero positivo, **y la ocupación del espectro**. Sabiendo que sólo existen dos tipos de usuarios (primarios y secundarios), cada estado será una combinación de ambos dentro del número de canales disponibles en el espectro. La tabla 2 muestra esta asociación para un ejemplo concreto con dos canales disponibles en el espectro.

Número de usuarios		Estado del sistema
Primarios	Secundarios	
0	0	1
1	0	2
2	0	3
0	1	4
1	1	5
0	2	6

Tabla 2: Mapeo usuarios-estado

El mapeo representado en la tabla 2 se almacena en una matriz, denominada Y , con la misma distribución que la tabla: la fila representa el estado del sistema, por tanto tendrá N_{st} filas, siendo N_{st} el número de estados; y tendrá dos columnas: la primera representará el número de usuarios primarios en el sistema y la segunda, el número de usuarios secundarios:

$$Y|_{N_{st} \times 2}; \quad j = \begin{cases} 1 \rightarrow Y_{ij} \equiv \text{Número de LUs en el estado } i - \text{ésimo} \\ 2 \rightarrow Y_{ij} \equiv \text{Número de SUs en el estado } i - \text{ésimo} \end{cases} \quad (16)$$

Se observa que existe una relación directa entre el número de canales del sistema, N , y el número de estados del mismo, N_{st} . Esta relación viene dada por la expresión siguiente:

$$N_{st} = (N + 1)^2 - \frac{N(N + 1)}{2} \quad (17)$$

Cadena de Markov en tiempo continuo (CTMC) asociada al modelo propuesto

La naturaleza del modelo propuesto da lugar a que su **caracterización inicial** deba llevarse a cabo mediante una **cadena de Markov en tiempo continuo (CTMC)**, ya que se parte de unas tasas de llegada y salida de los usuarios (ver apartado “Parámetros del escenario”, página 33).

Se empezará generando la **matriz de tasas (rates)**, R , asociada a la CTMC. Para el proceso de optimización que se llevará a cabo más adelante, será necesario obtener **tantas matrices de tasas como controles existan**. Es decir, para cada uno de los controles disponibles en el escenario, se calculará su matriz de tasas asociada que, como se describirá posteriormente, se utilizará en la búsqueda de la política óptima.

El proceso para el cálculo de cada elemento (tasa de transición) de la matriz es el siguiente:

1. Sea r_{ij} **el elemento a calcular**. Éste representa la tasa (o probabilidad) asociada a la transición del estado i al estado j .
2. Se recurre a la **matriz Y** para observar el número de usuarios de cada tipo, tanto en el estado origen, i , como en el estado destino, j . Se denotará por L_i el número de LUs en el estado i y S_i , el número de SUs.
3. Se comprueba la **diferencia entre ambas ocupaciones**, restando el número de usuarios primarios del estado j al número de usuarios primarios en el estado i anterior, y haciendo lo propio con los usuarios secundarios. En este momento, si el valor absoluto de la suma de las dos diferencias es mayor que 1, la transición estudiada representará la llegada o salida

de dos o más usuarios, lo cual no se contempla. Es decir, **la tasa o probabilidad de transición será nula si entre el estado i y el j hay más de un usuario de diferencia**. Si no es así, se sigue con el paso 4.

4. En función de la llegada o salida que se produzca:
 1. Si llega un usuario primario: $R_{ij} = \lambda_{LU}$.
 2. Si abandona un usuario primario: $R_{ij} = L_i \cdot \lambda_{LU}$.
 3. Si llega un usuario secundario, la tasa dependerá del control, u , al que se asocie la matriz generada:
 1. Si $u = 0 \rightarrow R_{ij} = 0$.
 2. Si $u = 1 \rightarrow R_{ij} = \lambda_{SU} \cdot P_1$.
 3. Si $u = 2 \rightarrow R_{ij} = \lambda_{SU} \cdot (P_1 + P_2)$.
 4. Si $u = n \rightarrow R_{ij} = \lambda_{SU} \cdot (P_1 + P_2 + \dots + P_n) = \lambda_{SU}$. Siendo n el máximo control posible.
 4. Si abandona un usuario secundario: $R_{ij} = S_i \cdot \lambda_{SU}$.

La figura 16 muestra de forma más gráfica el proceso descrito:

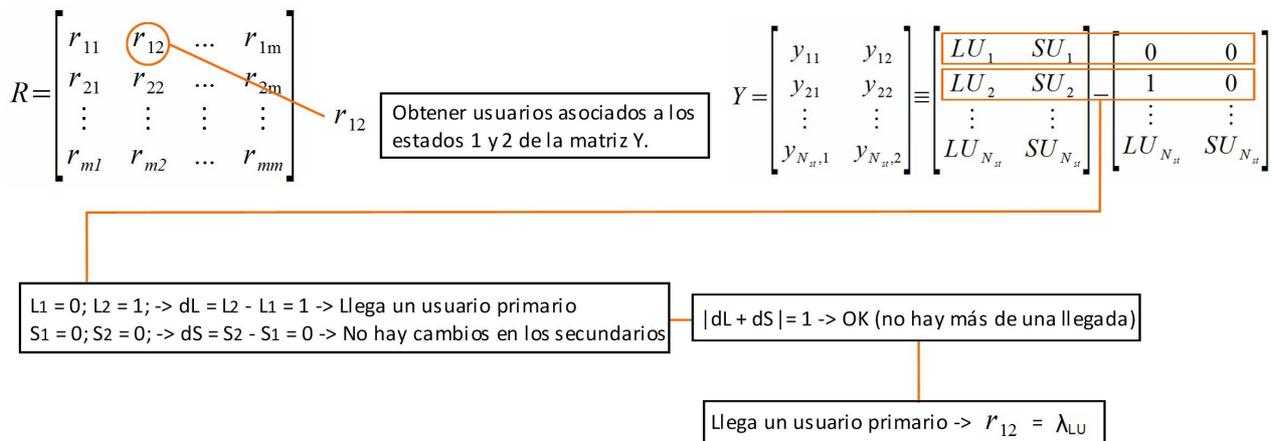


Figura 16: Ejemplo de cálculo de una tasa de transición

Uniformización de la CTMC. Transformación a cadena en tiempo discreto.

Para hacer un análisis del estado estacionario, se aprovecha la propiedad de ausencia de memoria de la distribución exponencial para definir una cadena de Markov en tiempo discreto a partir de la cadena en tiempo continuo. Este proceso se denomina **uniformización de la CTMC**.

La idea consiste en dividir el tiempo en pequeños intervalos de duración fija δ . La cadena discreta es, en principio, una aproximación de la cadena continua, pero tiende a ser exacta

cuando δ se aproxima a 0.

En este punto, se hace necesaria una breve introducción a la programación dinámica y la ecuación de Bellman.

La programación dinámica (DP), se basa en una idea muy sencilla denominada **principio de optimalidad**. Considérese un sistema con $\alpha = 1$ que pasa por N etapas, y un eje temporal en el que se indexan dichas etapas del sistema en orden decreciente, es decir, la primera etapa será la etapa N y la última, la etapa 1. Se pueden interpretar estos índices como el número de etapas que quedan hasta el final.

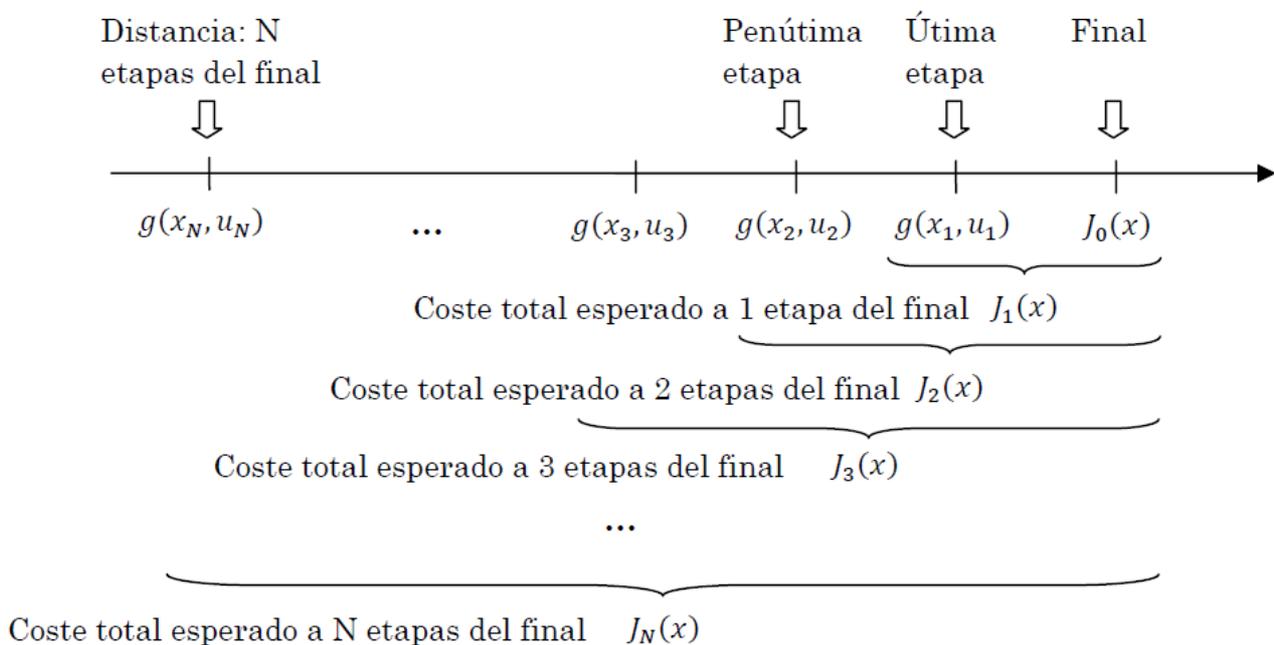


Figura 17: Funciones de coste por estado y coste hasta el final en un problema DP de N etapas

Se considera cualquier política admisible π . La función $J_0(x)$ denota el coste final o terminal. El coste total de la política π es:

$$C_{\pi}(x_N) = E \left\{ J_0(x_0) + \sum_{k=1}^N g(x_k, u_k) \right\} \quad (18)$$

Donde los valores de x_k y u_k vienen determinados por la política y la matriz de transición del sistema. Se puede dividir este coste en costes parciales. El coste de π a falta de una etapa para acabar es:

$$C_1(x_1) = E\{g(x_1, u_1) + J_0(f(x_1, u_1))\} \quad (19)$$

Donde se emplea la función $f(x, u)$, que devuelve el estado al que pasa el sistema cuando se encuentra en el estado x y aplica un control de valor u . Esta función se denomina **ecuación del sistema**. Dado que el sistema modelado es estocástico, el resultado de esta función es una variable aleatoria, y por eso se debe calcular la esperanza matemática. En general, el coste de π a k etapas del final viene dado por:

$$C_k(x_k) = E\{g(x_k, u_k) + C_{k-1}(f(x_k, u_k))\} \quad (20)$$

Se observa que se trata de una ecuación recursiva, en la que en cada recursión se añade un término más del sumatorio que aparece en la expresión original del coste total.

El principio de optimalidad consiste esencialmente en que el mínimo coste $J_N(x)$ para N etapas se puede calcular empleando la misma recursión, es decir, empezando por el coste final $J_0(x)$, calculando el mínimo coste a falta de una etapa $J_1(x)$,

$$J_1(x) = \min_{u \in U(x)} E\{g(x, u) + J_0(f(x, u))\} \quad (21)$$

y así sucesivamente, dando lugar a la siguiente fórmula recursiva:

$$J_k(x) = \min_{u \in U(x)} E\{g(x, u) + J_{k-1}(f(x, u))\} \quad k = 1, 2, \dots, N \quad (22)$$

De acuerdo a esta expresión, resulta lógico pensar que el coste óptimo en un problema de horizonte infinito es

$$J^*(x) = \lim_{N \rightarrow \infty} J_N(x) \quad (23)$$

para todos los estados $x \in S$. Esto se traduce en que aplicando sucesivamente la recursión DP, **el coste converge**, y lo hace **hacia el coste óptimo**.

Asumiendo que dicho límite existe y es $J^*(x)$, si se aplica la iteración DP sobre dicho coste, se puede suponer que no va a mejorar, ya que es el óptimo, pero tampoco va a empeorar, ya que la iteración DP devuelve el valor mínimo a la derecha de la ecuación. Por tanto, se puede plantear la siguiente ecuación, conocida como **ecuación de Bellman**:

$$J^*(x) = \min_{u \in U(x)} E \{ g(x, u) + \alpha J^*(f(x, u)) \} \quad x \in S \quad (24)$$

Se trata de una ecuación funcional, donde la incógnita es la función J . Puesto que J está definida para un conjunto contable de valores (S), puede verse también como un vector y, por tanto, la ecuación de Bellman es equivalente a un sistema de ecuaciones, cuya solución es el coste total del problema de coste infinito para cada uno de los posibles estados iniciales.

Sustituyendo u por $\mu(x)$, siendo ésta la política estacionaria que proporciona el control correspondiente al estado x , es razonable suponer que aquella μ que proporcione el valor mínimo a la derecha de la ecuación de Bellman debe ser la **política óptima** buscada.

En un MDP se definía $g(i, u)$ como el coste esperado por unidad de tiempo cuando se aplica el control u en el estado i . Por tanto, la esperanza matemática a la derecha de la ecuación de Bellman se debe aplicar al término donde aparece la función J . La ecuación de Bellman para MDPs queda de la siguiente forma:

$$J^*(i) = \min_{u \in U_i} \left[g(i, u) + \alpha \sum_{j \in S} p_{ij}(u) J^*(j) \right] \quad \forall i \in S \quad (25)$$

Recordando la ecuación 13, la función de coste para el caso del problema CTMC se definía de la siguiente forma:

$$\lim_{N \rightarrow \infty} \frac{1}{E\{t_N\}} E \left\{ \int_0^{t_N} g(x(t), u(t)) dt \right\} \quad (13)$$

Donde t_N es el instante de tiempo en el que se produce la transición N -ésima y $g(i, u)$ es el coste definido por unidad de tiempo. En la CTMC, se denotará la tasa de transición desde el estado i al estado siguiente cuando la acción de control es u como $\tau_i(u)$ y sus probabilidades de transición se denotan por $p_{ij}(u)$, $i \neq j$.

Se define el coste $G(i, u)$ como el **coste esperado en cada estado**:

$$G(i, u) = \frac{g(i, u)}{v_i(u)} = g(i, u) \tau_i(u) \quad (26)$$

Donde $\tau_i(u)$ representa el tiempo medio de permanencia en el estado i -ésimo cuando el control es u . Si los costes están definidos a partir de la transición entre estados, $\tilde{g}(i, u, j)$, los costes $G(i, u)$ se definen de la misma forma:

$$G(i, u) = \frac{g(i, u)}{v_i(u)} = \frac{\sum_{j \in S} p_{i,j}(u) \tilde{g}(i, u, j)}{v_i(u)} \quad (27)$$

Con este planteamiento, la ecuación de Bellman para el problema de coste medio por etapa de tipo CTMC tomará la siguiente forma:

$$h(i) = \min_{u \in U_i} \left[G(i, u) - \lambda \tau_i(u) + \sum_{j=1}^n p_{ij}(u) h(j) \right], \text{ para } i=1, 2, \dots, n \quad (28)$$

Una vez presentada la ecuación de Bellman, y volviendo a la transformación CTMC - DTMC, se plantea el siguiente proceso de uniformización:

1. Tasas asociadas a la DTMC:

1. Se escoge una tasa de transición uniforme γ tal que $\gamma > v_i(u)$, para todo i, u . Para ello, en primer lugar se obtiene la tasa de transición total para cada estado, sumando todos los elementos de cada fila. A continuación, se establece γ como el mayor de los elementos obtenidos, garantizando así la condición inicial.
2. Se definen las probabilidades de transición del DTMC asociado de la siguiente forma:

$$\tilde{p}_{i,j} = \begin{cases} \frac{v_i(u)}{\gamma} p_{ij}(u) & \text{si } i \neq j \\ 1 - \frac{v_i(u)}{\gamma} & \text{si } i = j \end{cases} \quad (29)$$

donde $p_{ij}(u)$ se calcula como

$$p_{ij}(u) = \frac{r_{ij}(u)}{\tau_i} \quad (30)$$

siendo $r_{ij}(u)$ la tasa de transición asociada a la CTMC (elemento de la matriz R).

2. Costes asociados a la DTMC.

Los costes asociados a la CTMC también tendrán su transformación para poder resolver el problema mediante una DTMC:

$$\tilde{G}(i, u) = \frac{G(i, u)}{\tau_i(u)} = G(i, u) v_i(u) \quad (31)$$

La ecuación de Bellman del problema discreto asociado queda así [18]:

$$\tilde{h}(i) = \min_{u \in U_i} \left[\tilde{G}(i, u) - \lambda + \sum_{j=1}^n p_{ij}(u) \tilde{h}(j) \right], \forall i = 1, 2, \dots, n \quad (32)$$

Particularmente, en el problema DSA planteado, se definirá el coste $g(i, u)$ por unidad de tiempo como una combinación lineal de dos vectores:

$$g(i, u) = \alpha \cdot g_L(i, u) - (1 - \alpha) \cdot g_U(i, u) \quad (33)$$

siendo $\alpha > 0$ y $g_L(i, u)$ el coste asociado al coste (bloqueo) de los usuarios primarios,

$$g_L(i, u) = \begin{cases} 1, & \text{si } i = N \\ 0, & \text{en caso contrario} \end{cases} \quad (34)$$

es decir, sólo habrá coste para los usuarios primarios cuando todos los *slots* o canales del espectro estén ocupados. Se observa que no depende directamente del control, u , ya que éste no afecta a los usuarios con licencia: $g_L(i, u) \approx g_L(i)$.

Por otro lado, el vector $g_U(i, u)$ representará el beneficio asociado a los usuarios secundarios, y éste sí dependerá del control que haya en cada estado. Nótese que su peso en la combinación del coste global es negativo, ya que la minimización buscada en el coste global pasa por una maximización del beneficio económico.

Sea $P(r|u)$ la probabilidad condicionada de que, dado un control u , el precio aceptado para el siguiente usuario secundario aceptado sea b :

$$P(r|u) = \begin{cases} \frac{p_r}{\sum_{j=u}^m p_j}, & \text{si } r \geq u \\ 0, & \text{en otro caso} \end{cases} \quad (35)$$

y sea $\tilde{g}_u(i, u, j)$ el beneficio medio asociado a la transición del estado i hacia el j , cuya expresión es:

$$\tilde{g}_U(i, u, j) = \begin{cases} p_U \sum_{r=1}^m b_r P(r|u), & \text{si } j = i + 1 \\ 0, & \text{en otro caso} \end{cases} \quad (36)$$

donde

$$p_U = \frac{\lambda_U}{\lambda_U + \lambda_L} \quad (37)$$

denota la probabilidad de que el siguiente usuario entrante sea secundario. A partir de estas expresiones, la expresión del beneficio $g_U(i, u)$ queda:

$$g_U(i, u) = \sum_{j=1}^N \tilde{g}_U(i, u, j) p_{i,j}(u) = p_{i,i+1}(u) p_U \sum_{j=1}^N b_j P(j|u). \quad (38)$$

Por tanto, se puede reescribir la ecuación 32 de la siguiente forma:

$$\tilde{h}(i) = \min_{u \in U_i} \left[\alpha \cdot g_L(i, u) + (1 - \alpha) \cdot g_U(i, u) v_i(u) - \lambda + \sum_{j=1}^n p_{ij}(u) \tilde{h}(j) \right], \forall i = 1, 2, \dots, n \quad (39)$$

Con esta definiciones de coste es posible, mediante un ajuste adecuado del parámetro α , construir el frente de Pareto descrito en la introducción, que determinará el máximo beneficio posible para una determinada probabilidad de bloqueo de los usuarios primarios.

Cálculo de la política óptima. El algoritmo "Policy Iteration".

Como se comentaba en apartados anteriores, existirá una política óptima para cada valor del parámetro α escogido. Por ello, el frente de Pareto buscado se construirá a partir de un barrido de dicho parámetro, es decir, **cada punto del frente de Pareto representará la política óptima asociada a un valor de α .**

Por conveniencia en la notación, la iteración DP, presentada en la ecuación 22 (página 39) se representará de forma compacta mediante un operador (T), de la siguiente forma:

$$(TJ)(x) = \min_{u \in U(x)} E\{g(x, u) + \alpha J(f(x, u))\} \quad x \in S \quad (40)$$

Es decir, que el operador T aplicado sobre un vector de costes J obtiene el resultado de hacer sobre dicho vector una iteración DP. La notación $(TJ)(x)$ indica que se está accediendo a la posición x del vector TJ . La iteración DP en un problema de horizonte finito sería, por tanto:

$$J_k(x) = (TJ_{k-1})(x) \quad k = 1, 2, \dots, N \quad (41)$$

De forma muy similar al operador T , se define el operador T_μ para una determinada política estacionaria μ :

$$(T_\mu J)(x) = E\{g(x, \mu(x)) + \alpha J(f(x, \mu(x)))\} \quad x \in S \quad (42)$$

donde $\mu(x)$ proporciona el control correspondiente al estado x . Tanto T como T_μ permiten expresar ciertas operaciones de forma muy compacta. Por ejemplo, aplicar la iteración DP k veces sobre una función de coste J se expresa como $T^k J$, donde

$$T^k J = T(T^{k-1} J) = T(T^{k-1}(T^{k-2} J)) = \dots \quad (43)$$

Igualmente, para una política estacionaria μ , el coste para un problema de k etapas, con un coste final $\alpha^k J$ se expresa como $T_\mu^k J$.

La política óptima se obtendrá mediante el algoritmo **Policy Iteration**. Antes de explicar el funcionamiento de este algoritmo, es necesario introducir una serie de condiciones previas.

1. **Convergencia del algoritmo DP:** para cualquier función acotada $J: S \rightarrow R$, la función de

coste óptimo satisface:

$$J^*(x) = \lim_{N \rightarrow \infty} (T^N J)(x), \quad \forall x \in S \quad (44)$$

Esta proposición se puede formular de forma más específica para una MDP con N estados. La secuencia generada por la iteración

$$J_{k+1}(i) = \min_{u \in U_i} \left[g(i, u) + \alpha \sum_{j=1}^N p_{ij}(u) J_k(j) \right], \quad \forall i=1, 2, \dots, N \quad (45)$$

converge a los costes óptimos $J^*(i)$, $i=1, \dots, N$, empezando a partir de cualquier conjunto de condiciones iniciales $J_0(1), J_0(2), \dots, J_0(N)$.

2. Ecuación de Bellman: la función de coste óptimo J^* satisface

$$J^*(i) = \min_{u \in U_i} \left[g(i, u) + \alpha \sum_{j \in S} p_{ij}(u) J^*(j) \right], \quad \forall i \in S \quad (46)$$

O, equivalentemente,

$$J^* = T J^* \quad (47)$$

Además, J^* es la única solución de esta ecuación entre el conjunto de funciones acotadas. La ecuación de Bellman, que es probablemente uno de los resultados más relevantes, tiene una interpretación muy sencilla, como se avanzaba anteriormente: en el estado i , el coste óptimo $J^*(i)$ es el mínimo de la suma del coste esperado del estado actual y el coste óptimo esperado de todos los estados futuros. Por tanto, para cualquier política estacionaria, los costes $J_\mu(1), J_\mu(2), \dots, J_\mu(N)$ son la única solución de la ecuación

$$J_\mu(i) = g(i, u) + \alpha \sum_{j \in S} p_{ij}(u) J_\mu(j), \quad \forall i \in S \quad (48)$$

O, equivalentemente:

$$J_\mu = T_\mu J_\mu \quad (49)$$

3. **Condición necesaria y suficiente para optimalidad:** una política estacionaria μ es óptima si y sólo si para todos los estados i , $\mu(i)$ alcanza el mínimo de la ecuación de Bellman para todos los estados, es decir:

$$TJ^* = T_{\mu} J^* \quad (50)$$

Esta política implica que si se alcanza el mínimo de la ecuación de Bellman para todos los valores de i , existe una política estacionaria óptima. Si el conjunto de decisiones de control, U , es finito, está garantizado que exista dicha política óptima. Nótese que en ningún momento se garantiza que haya una política estacionaria óptima, sólo se garantiza que haya un coste óptimo, que se podrá alcanzar con una o más políticas estacionarias, siempre que todas ellas cumplan esta condición. Es decir, que si se encuentra una política óptima que, aplicada al vector de coste óptimo J^* dé como resultado ese mismo vector, esa política es óptima.

4. Si μ y $\bar{\mu}$ son dos políticas estacionarias tales que $T_{\bar{\mu}} J_{\mu} = T J_{\mu}$, o lo que es lo mismo:

$$\bar{\mu}(i) = \arg \min_{u \in U_i} \left[g(i, u) + \alpha \sum_{j \in S} p_{ij}(u) J^*(j) \right], \quad \forall i = 1, 2, \dots, n, \quad (51)$$

entonces $J_{\bar{\mu}} \leq J_{\mu}$ para $i = 1, 2, \dots, n$. Además, si μ no es óptima, la desigualdad es estricta para, al menos, un estado i .

Lo que esta condición indica es que una política obtenida tras optimizar una función de coste es mejor que cualquier política asociada a la función de coste original.

Para la implementación del algoritmo Policy Iteration, resulta muy útil introducir una **notación vectorial** para las incógnitas y operadores empleados:

- Función J como vector, en un sistema de n estados:

$$J = \begin{bmatrix} J(1) \\ \vdots \\ J(n) \end{bmatrix} \quad (52)$$

- Matriz de transición P_μ para una política estacionaria, μ :

$$P_\mu = \begin{bmatrix} p_{11}(\mu(1)) & \cdots & p_{1n}(\mu(1)) \\ \vdots & \ddots & \vdots \\ p_{n1}(\mu(n)) & \cdots & p_{nn}(\mu(n)) \end{bmatrix} \quad (53)$$

Nótese que el control $u = \mu(i)$ es igual para todos los elementos de la fila i -ésima.

- Vector de coste g_μ para μ lo podemos escribir como

$$g_\mu = \begin{bmatrix} g(1, \mu(1)) \\ \vdots \\ g(n, \mu(1)) \end{bmatrix} \quad (54)$$

Mediante esta notación, la ecuación que proporciona el coste J_μ correspondiente a una política estacionaria μ es

$$J_\mu = T_\mu J_\mu = g_\mu + \alpha P_\mu J_\mu \quad (55)$$

Dado que no se trata de un problema con descuento, $\alpha=1$. El algoritmo Policy Iteration es una consecuencia directa de la condición 4, es decir, consiste en generar una secuencia de políticas estacionarias, cada una con un coste mejorado respecto a la política precedente.

El algoritmo tomará como entradas las matrices de transición (tantas como controles haya) calculadas en el proceso de uniformización, así como todos los vectores de costes asociados, y devolverá como salida la política óptima obtenida, el coste medio asociado a dicha política y la matriz de probabilidades de transición calculada combinando las matrices anteriores en función de los controles de la política óptima.

A continuación se describen los tres pasos que realiza el algoritmo:

1. **Inicialización.** Se escoge una política estacionaria inicial, μ^0 , aleatoria.
2. **Evaluación de la política.** Dada una política estacionaria, μ^k , se computa su función de coste correspondiente, J_μ^k , resolviendo el siguiente sistema de ecuaciones lineales:

$$(I - P_{\mu^k})J_{\mu^k} = g_{\mu^k} \quad (56)$$

3. **Mejora de la política.** Se obtiene una nueva política estacionaria mediante la iteración siguiente:

$$\mu^{k+1}(i) = \arg \min_{u \in U_i} \left[g(i, u) + \sum_{j \in S} p_{ij}(u) J_{\mu^k} \right] \quad \forall i = 1, 2, \dots, n \quad (57)$$

Nótese que la **condición de terminación** viene dada por la ecuación de Bellman: en el momento en que $J_{\mu^k} = TJ_{\mu^k}$, se habrá encontrado el óptimo y por tanto finaliza el algoritmo.

Puesto que el número de políticas estacionarias es finito, ya que los conjuntos S y U son finitos, y se genera una política distinta y mejorada en cada iteración, se sabe que el algoritmo alcanzará una política estacionaria óptima en un número finito de iteraciones. El problema que se presenta con este algoritmo es que en cada iteración hay que resolver un sistema lineal de ecuaciones, dando lugar a un proceso relativamente lento cuando el sistema tiene muchos estados. No obstante, el número de estados del modelo propuesto será lo suficientemente bajo como para que la convergencia se alcance considerablemente rápido.

Una vez que el algoritmo finaliza se obtiene la política óptima, μ^* , y a partir de ésta, la matriz de transiciones P^* y el coste medio, J^* . Entonces se obtiene el coste (o beneficio) para cada tipo de usuario por separado, para poder enfrentar ambos objetivos en la representación del frente de Pareto. Para ello, se vuelve a resolver la ecuación 56 del segundo paso de Policy Iteration, pero esta vez con los costes individuales, $g_L(i, u)$ y $g_U(i, u)$, en lugar de la combinación de ambos, $g(i, u)$. Dado que estos costes se obtienen de la política óptima, se denotarán por $g_L^*(i, u)$ y $g_U^*(i, u)$.

$$J_L^{opt} = J^{opt} \Big|_{g=g_L} = g_L + P^{opt} J^{opt} \quad (58)$$

$$J_U^{opt} = J^{opt} \Big|_{g=g_U} = g_U + P^{opt} J^{opt} \quad (59)$$

Para comprobar que la política obtenida con Policy Iteration es, en efecto, la óptima, se realizará el procedimiento descrito anteriormente, pero sin hacer uso del algoritmo Policy Iteration. Es decir, **se calcularán y representarán los costes asociados a distintas políticas aleatorias** para verificar la validez de la política óptima. Si el resultado es correcto, los costes de la política óptima representarán el frente de Pareto, y los de las distintas políticas aleatorias serán puntos

dentro del área T encima de éste (ver figura 6, página 14), como muestra la figura siguiente:

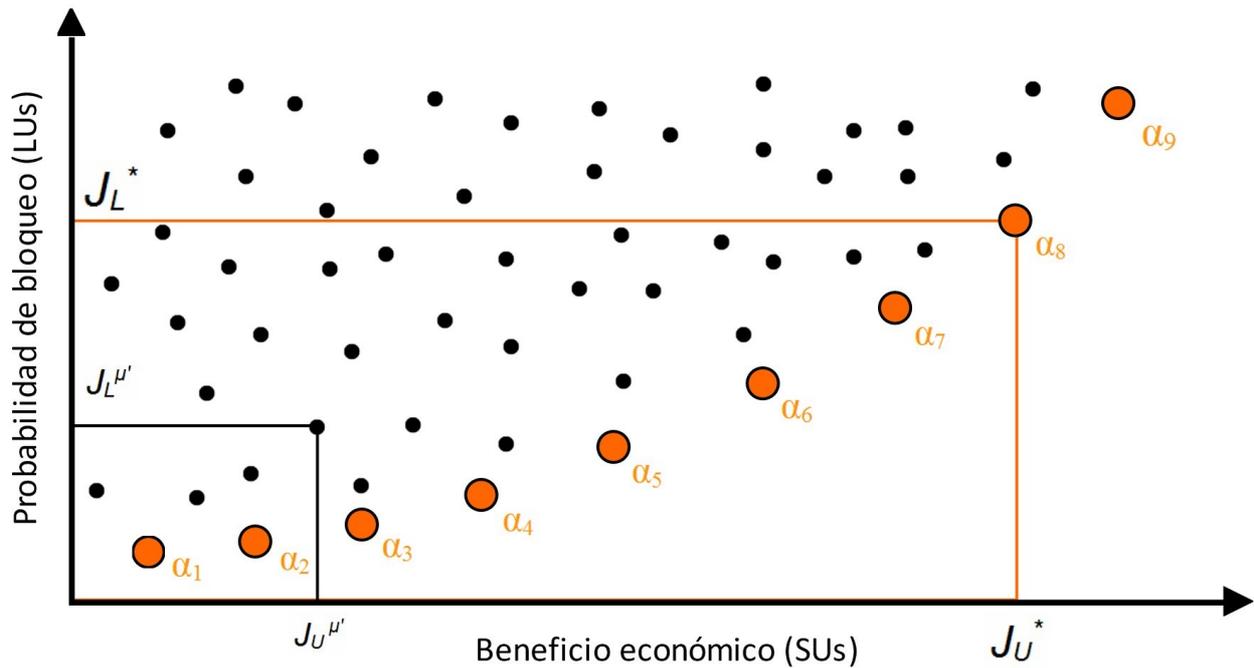


Figura 18: Frente de Pareto esperado a partir de las diferentes políticas

Los resultados obtenidos en Matlab se expondrán con detalle en apartados posteriores.

Para acabar este subapartado, se resume todo el proceso descrito en los siguientes pasos:

1. Cálculo del número de estados a partir del número de canales y mapeo con la ocupación.
2. Cálculo de las matrices de tasas, R_u , asociadas a la CTMC.
3. Uniformización: conversión al problema DTMC asociado.
 1. Cálculo de las matrices de probabilidades de transición M_u a partir de R_u .
 2. Cálculo de los costes asociados a cada tipo de usuario para cada control.
4. Cálculo de la política óptima y su coste/beneficio mediante el algoritmo Policy Iteration.
5. Cálculo de los costes/beneficios asociados a diferentes políticas aleatorias.
6. Representación de los costes de los pasos 4 y 5 en un frente de Pareto.

3.2. Optimización mediante programación lineal.

3.2.1. Introducción a la programación lineal.

La programación matemática y en particular la **Programación Lineal (LP)** y sus extensiones ha adquirido en los últimos años una gran importancia, dada la gran variedad de problemas que pueden ser modelados como problemas lineales. Esto, unido al desarrollo tecnológico experimentado por los ordenadores cada vez más potentes y de menor precio, ha hecho que la programación lineal sea más utilizada, puesto que ya es posible resolver, hasta con un ordenador personal, muchos de los problemas que hasta hace unos años requerían el uso de grandes ordenadores [28].

El planteamiento de un problema de optimización (en este caso, búsqueda de la política óptima), en términos matemáticos, queda definido por:

- a) Las **variables** del problema, que reciben el nombre de “variables instrumentales”. Se denotarán por x_i , o por el vector X .
- b) La **función objetivo**, $F(X)$, que es la descripción en términos matemáticos del objetivo final a conseguir.
- c) Las **restricciones** o ecuaciones de restricción, $h_i(x)$, que condicionarán el valor de las variables.

Particularmente, la denominada “programación lineal” se caracteriza por:

1. **La función objetivo es lineal**, es decir:

$$F(x) = c_1 x_1 + c_2 x_2 + \dots + c_n x_n \quad (60)$$

2. **Las condiciones o restricciones del problema también son lineales**, es decir:

$$\begin{aligned} h_1(X); & a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1 \\ h_2(X); & a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \leq b_2 \\ & \dots\dots\dots \\ h_m(X); & a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \leq b_m \end{aligned} \quad (61)$$

3. **Las variables instrumentales han de tomar necesariamente valores mayores o iguales que cero.** Se trata de una condición de no negatividad:

$$x_i \geq 0 \quad (62)$$

En notación matricial:

$$\begin{array}{ll} \text{Max}(\text{Min}) & F(x) = C^t \cdot X \\ \text{s.a.} & A \cdot X \leq B \\ & X \geq 0 \end{array} \quad (63)$$

Donde:

$$C^t = (c_1, c_2, \dots, c_n) \quad (64)$$

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (65)$$

son vectores de n componentes.

A es una matriz de tamaño $m \times n$, llamada también "matriz técnica":

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = (P_1, P_2, P_3, \dots, P_n) \quad (66)$$

siendo P_j el vector columna asociado a la variable x_j . B es el vector de términos independientes de m componentes.

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad (67)$$

Un problema de programación lineal puede presentarse de tres formas diferentes:

- a) En “forma canónica”, cuando las restricciones son desigualdades en una de las formas siguientes:

$$\begin{array}{ll} \text{a.1)} & \begin{array}{l} \text{Max}(\text{Min}) \quad F(x)=C^t \cdot X \\ \text{s.a.} \quad A \cdot X \leq B \\ X \geq 0 \end{array} \\ \text{a.2)} & \begin{array}{l} \text{Max}(\text{Min}) \quad F(x)=C^t \cdot X \\ \text{s.a.} \quad A \cdot X \geq B \\ X \geq 0 \end{array} \end{array}$$

- b) En “forma *standard*”, cuando las restricciones son igualdades:

$$\begin{array}{l} \text{Max}(\text{Min}) \quad F(x)=C^t \cdot X \\ \text{s.a.} \quad A \cdot X = B \\ X \geq 0 \end{array}$$

3.2.2. Caracterización del modelo propuesto mediante programación lineal.

En la formulación del modelo propuesto basada en la programación lineal, el objetivo consistirá en maximizar el beneficio obtenido con los usuarios secundarios (función objetivo) sujeto a que el coste o bloqueo de los usuarios primarios esté por encima de un valor determinado (restricción).

Al igual que en la aproximación basada en MDP, el software utilizado para el proceso de optimización que aquí se describe será MATLAB. En particular se usará CVX [29], que es un software de modelado basado en MATLAB para resolver problemas de optimización convexa que se describen usando un conjunto predeterminado de reglas sintácticas. Es decir, CVX “convierte” a MATLAB en un lenguaje de modelado de este tipo de problemas, permitiendo especificar las restricciones y objetivos mediante la sintaxis estándar de MATLAB. A continuación se muestra un ejemplo de la sintaxis básica empleada:

```
cvx_begin
    variables VARIABLES
    minimize (maximize) FUNCIÓN OBJETIVO
    subjetc to
        RESTRICCIÓN/ES
cvx_end
```

Así será la caracterización del sistema mediante una formulación de programación lineal:

1. **Variables.** Son las incógnitas que se tratan de resolver con el problema de optimización. Para este caso, se denotarán por la letra φ , y por conveniencia, se representarán en el vector columna Φ de la siguiente forma:

$$\Phi(i, u) = \begin{bmatrix} \varphi(1, u_1) \\ \varphi(2, u_1) \\ \vdots \\ \varphi(N_{st}, u_1) \\ \varphi(1, u_2) \\ \varphi(2, u_2) \\ \vdots \\ \varphi(N_{st}, u_2) \\ \varphi(1, u_3) \\ \vdots \\ \varphi(N_{st}, u_n) \end{bmatrix} \quad \text{con } i \in S, u \in U \quad (68)$$

donde $\varphi(i, u)$ representa la probabilidad de usar el control u en el estado i .

2. **Función objetivo** (a maximizar): beneficio obtenido de los usuarios secundarios.

$$\arg \max_{\Phi} \sum_{i \in S} \sum_{u \in U} r(i, u) \varphi(i, u) \quad (69)$$

donde $r(i, u)$ representa el beneficio obtenido al usar el control u en el estado i . La función objetivo también tendrá una notación vectorial asociada, R (*revenue*):

$$R = [r(1, u_1) \quad \cdots \quad r(N_{st}, u_1) \quad \cdots \quad r(1, u_n) \quad \cdots \quad r(N_{st}, u_n)] \quad , \quad (70)$$

donde los beneficios asociados a cada estado y control, $r(i, u)$, se obtendrán del mismo proceso de uniformización descrito en la aproximación mediante MDP (ecuación 38, página 43).

3. **Restricciones:**

- a) Mantener el bloqueo (estrictamente hablando, la probabilidad de bloqueo) de los usuarios primarios por debajo de un umbral, β_L :

$$\sum_{i \in S} \sum_{u \in U} b_L(i, u) \varphi(i, u) \leq \beta_L \quad (71)$$

donde $b_L(i, u)$ representa el bloqueo de un usuario primario en el estado i . La notación vectorial asociada en este caso será C (*constraints*):

$$C = [b(1, u_1) \quad \cdots \quad b(N_{st}, u_1) \quad \cdots \quad b(1, u_n) \quad \cdots \quad b(N_{st}, u_n)] \quad (72)$$

Como en la función objetivo, los elementos de coste de este vector se calculan con la ecuación 34, descrita en el proceso de uniformización de la página 42.

b) La suma de todos los elementos del vector Φ debe ser la unidad:

$$\sum_{i \in S} \sum_{u \in U} \varphi(i, u) = 1 \quad (73)$$

c) Condición de no negatividad, expuesta en la ecuación 62 (página 51). Las variables $\varphi(i, u)$ deberán ser mayores o iguales que cero:

$$\varphi(i, u) \geq 0 \quad \forall i \in S, u \in U \quad (74)$$

d) La suma de todos los valores de $\varphi(i, u)$ asociados a un estado ha de ser igual al producto de dichos valores por las probabilidades de transición:

$$\sum_{u \in U} \varphi(j, u) - \sum_{i \in S} \sum_{u \in U} p_{ji}(u) \varphi(i, u) = 0 \quad \forall j \in S \quad (75)$$

donde p_{ji} representa la probabilidad de transición entre los estados j e i . También por conveniencia en la notación, se agruparán en vectores $P(i)$, de la siguiente forma:

$$\begin{aligned} P(1) &= [p_{11}(u_1) \quad \cdots \quad p_{1, N_{st}}(u_1) \quad \cdots \quad p_{11}(u_n) \quad \cdots \quad p_{1, N_{st}}(u_n)] \\ P(2) &= [p_{21}(u_1) \quad \cdots \quad p_{2, N_{st}}(u_1) \quad \cdots \quad p_{21}(u_n) \quad \cdots \quad p_{2, N_{st}}(u_n)] \\ &\vdots \\ P(N_{st}) &= [p_{N_{st}, 1}(u_1) \quad \cdots \quad p_{N_{st}, N_{st}}(u_1) \quad \cdots \quad p_{N_{st}, 1}(u_n) \quad \cdots \quad p_{N_{st}, N_{st}}(u_n)] \end{aligned}$$

Cada elemento p_{ij} de estos vectores se obtiene de la matriz de transiciones de

probabilidad, M , calculada durante el proceso de uniformización en la aproximación mediante MDP (ver ecuaciones 29 y 30, en la página 41).

Como se describía anteriormente, las variables $\varphi(i, u)$ que se obtendrán como resultado en el proceso de optimización representarán las probabilidades de usar el control u en el estado i , pero no serán probabilidades en sentido estricto, sino que serán directamente proporcionales a éstas. Por ello será necesario llevar a cabo un proceso de normalización para hablar de probabilidades en sentido estricto:

$$q_{u^{opt}}(i, u) = \frac{\varphi(i, u)}{\sum_{u \in U} \varphi(i, u)} \quad , \quad (76)$$

de forma que se cumplirá que

$$\left. \begin{array}{l} 0 \leq q_{u^{opt}}(i, u) \leq 1 \\ \sum_{u \in U} q_{u^{opt}}(i, u) = 1 \end{array} \right\} \forall i \in S, u \in U \quad . \quad (77)$$

Nótese que si se cumple que

$$q(x, u_j) = 1 \rightarrow q(x, u_i) = 0, \quad \forall x \in S, i \neq j \quad , \quad (78)$$

lo que se traduce en que en el estado x se empleará el control u_i con total seguridad, **la política obtenida será estacionaria**.

Para facilitar la lectura de los valores $q(x, u)$, éstos se agruparán en una matriz, Q , con tantas filas como estados y tantas columnas como controles. Así, a simple vista se podrá observar con qué probabilidad se elige cada control mirando a una fila de la matriz. Obviamente, la suma de todos los elementos de una fila será igual a la unidad.

$$Q = \begin{bmatrix} q(1, u_1) & \cdots & q(1, u_n) \\ q(2, u_1) & \cdots & q(2, u_n) \\ \vdots & \ddots & \vdots \\ q(N_{st}, u_1) & \cdots & q(N_{st}, u_n) \end{bmatrix} \quad (79)$$

A continuación se presenta la formulación descrita en notación vectorial para un ejemplo sencillo, con un único canal ($N = 1 \rightarrow N_{st} = 3$) y dos posibles controles, u_1 y u_2 .

El vector Φ con las variables del problema será:

$$\Phi = \begin{bmatrix} \varphi(1, u_1) \\ \varphi(2, u_1) \\ \varphi(3, u_1) \\ \varphi(1, u_2) \\ \varphi(2, u_2) \\ \varphi(3, u_2) \end{bmatrix} . \quad (80)$$

La función objetivo, representada en el vector R , será:

$$R = [r(1, u_1) \quad r(2, u_1) \quad r(3, u_1) \quad r(1, u_2) \quad r(2, u_2) \quad r(3, u_2)] \quad (81)$$

Los costes asociados al bloqueo de LUs, en el vector C :

$$C = [b(1, u_1) \quad b(2, u_1) \quad b(3, u_1) \quad b(1, u_2) \quad b(2, u_2) \quad b(3, u_2)] \quad (82)$$

Los vectores P con las probabilidades de transición serán:

$$\begin{aligned} P(1) &= [p_{11}(u_1) \quad p_{12}(u_1) \quad p_{13}(u_1) \quad p_{11}(u_2) \quad p_{12}(u_2) \quad p_{13}(u_2)] \\ P(2) &= [p_{21}(u_1) \quad p_{22}(u_1) \quad p_{23}(u_1) \quad p_{21}(u_2) \quad p_{22}(u_2) \quad p_{23}(u_2)] \end{aligned} \quad (83)$$

Por tanto, en este problema, la función a maximizar será

$$\arg \max_{\Phi} F(\Phi) = R \cdot \Phi \quad (84)$$

y estará sujeto a las siguientes restricciones:

a) Bloqueo de LUs por debajo del umbral β_L (ecuación 71):

$$C \cdot \Phi \leq \beta_L \quad (85)$$

b) La suma de los elementos de Φ debe ser la unidad:

$$[1 \ 1 \ 1 \ 1 \ 1 \ 1] \cdot \Phi = 1 \quad (86)$$

c) Condición de no negatividad:

$$\Phi \geq 0 \quad (87)$$

d) La suma de todos los valores de $\varphi(i,u)$ asociados a un estado ha de ser igual al producto de dichos valores por las probabilidades de transición:

$$\begin{aligned} [1 \ 0 \ 0 \ 1 \ 0 \ 0] \cdot \Phi - P(1) \cdot \Phi &= 0 \\ [0 \ 1 \ 0 \ 0 \ 1 \ 0] \cdot \Phi - P(2) \cdot \Phi &= 0 \\ [0 \ 0 \ 1 \ 0 \ 0 \ 1] \cdot \Phi - P(3) \cdot \Phi &= 0 \end{aligned} \quad (88)$$

Volviendo a la descripción general del procedimiento, el resultado inmediato que se obtiene será el vector Φ y, de forma casi inmediata, la matriz Q . Con estos datos, el siguiente paso es confeccionar una nueva matriz de probabilidades de transición, de una forma similar a la realizada en el subapartado anterior.

En la aproximación basada en MDP, una vez obtenida la política óptima, μ^* , se generaba su matriz de probabilidades de transición asociada, que se denotaba por P^* , “escogiendo” aquellos elementos de las matrices asociadas a cada control particular (M_0, M_1, \dots, M_n) en función de lo que dictara la política óptima.

En este caso, el proceso es similar, pero, dado que a priori se desconoce si la política obtenida es estacionaria (si $q(i,u) = 1$ para todo $i \in S$), será necesario generar la matriz P^* **ponderando** las probabilidades de transición de las distintas matrices (M_0, M_1, \dots, M_n) en función de las probabilidades de la matriz Q :

$$p_{ij}^{opt} = M_0(i, j) \cdot q(i,0) + M_1(i, j) \cdot q(i,1) + \dots + M_n(i, j) \cdot q(i, n) \quad (89)$$

En el ejemplo descrito antes, con tres estados y dos controles, se tendría:

$$Q = \begin{bmatrix} q(1,0) & q(1,1) \\ q(2,0) & q(2,1) \\ q(3,0) & q(3,1) \end{bmatrix} \quad (90)$$

La matriz P^* tendría un tamaño de 3x3 (una fila/columna por estado):

$$P^{opt} = \begin{bmatrix} p_{11}^{opt} & p_{12}^{opt} & p_{13}^{opt} \\ p_{21}^{opt} & p_{22}^{opt} & p_{23}^{opt} \\ p_{31}^{opt} & p_{32}^{opt} & p_{33}^{opt} \end{bmatrix}, \quad (91)$$

y, por tanto, el valor un elemento de P^* , por ejemplo p_{23}^* , se calcularía así:

$$p_{23}^{opt} = M_0(2,3) \cdot q(2,0) + M_1(2,3) \cdot q(2,1) + M_2(2,3) \cdot q(2,2) \quad (92)$$

A continuación, de modo análogo, se calculan los costes asociados a la política obtenida. Se parte, en este caso, de los vectores $g_L(i,u)$ y $g_U(i,u)$, calculados también en el anterior subapartado (ver ecuaciones 58 y 59, página 48). Evidentemente, el objetivo de todo el proceso de optimización es representar el coste medio del sistema en un frente de Pareto, y para ello es necesario obtener los costes por separado.

Los vectores g_L^* y g_U^* tendrán tantas posiciones como estados, y también se calcularán ponderando los vectores g_L y g_U con las probabilidades de la matriz Q:

$$\begin{aligned} g_L^{opt}(i) &= g_L(i,0) \cdot q(i,0) + g_L(i,1) \cdot q(i,1) + \dots + g_L(i,n) \cdot q(i,n) \\ g_S^{opt}(i) &= g_S(i,0) \cdot q(i,0) + g_S(i,1) \cdot q(i,1) + \dots + g_S(i,n) \cdot q(i,n) \end{aligned} \quad (93)$$

Para el caso particular del ejemplo:

$$\begin{aligned} g_L^{opt}(1) &= g_L(1,0) \cdot q(1,0) + g_L(1,1) \cdot q(1,1) + g_L(1,2) \cdot q(1,2); \\ g_L^{opt}(2) &= g_L(2,0) \cdot q(2,0) + g_L(2,1) \cdot q(2,1) + g_L(2,2) \cdot q(2,2); \\ g_L^{opt}(3) &= g_L(3,0) \cdot q(3,0) + g_L(3,1) \cdot q(3,1) + g_L(3,2) \cdot q(3,2); \end{aligned} \quad (94)$$

$$\begin{aligned} g_U^{opt}(1) &= g_U(1,0) \cdot q(1,0) + g_U(1,1) \cdot q(1,1) + g_U(1,2) \cdot q(1,2); \\ g_U^{opt}(2) &= g_U(2,0) \cdot q(2,0) + g_U(2,1) \cdot q(2,1) + g_U(2,2) \cdot q(2,2); \\ g_U^{opt}(3) &= g_U(3,0) \cdot q(3,0) + g_U(3,1) \cdot q(3,1) + g_U(3,2) \cdot q(3,2); \end{aligned} \quad (95)$$

Una vez que se han obtenido estos vectores y la matriz de probabilidades de transición P^* , el siguiente paso es obtener el coste medio asociado a cada tipo de usuario para poder llevar a cabo la representación. Para ello, se resolverá la misma ecuación que en Policy Iteration,

$$(I - P_{u^k}) J_{u^k} = g_{u^k} \quad , \quad (56)$$

particularizada para cada coste individual:

$$J_L^{opt} = J^{opt} |_{g=g_L} = g_L + P^{opt} J^{opt} \quad (58)$$

$$J_U^{opt} = J^{opt} |_{g=g_U} = g_U + P^{opt} J^{opt} \quad (59)$$

Con los costes J_U^* y J_L^* ya se puede representar el frente de Pareto que, teniendo en cuenta que se obtiene a partir de un proceso de optimización, estará asociado a la política óptima. Una diferencia respecto al frente de Pareto anterior es que con esta formulación, la representación se realiza para un valor de α fijo. Los distintos puntos de la gráfica obtenida se obtendrán mediante un barrido de la variable β_L , para obtener, fijada una probabilidad de bloqueo para los usuarios con licencia, cuál es el máximo beneficio que es posible obtener de los usuarios secundarios.

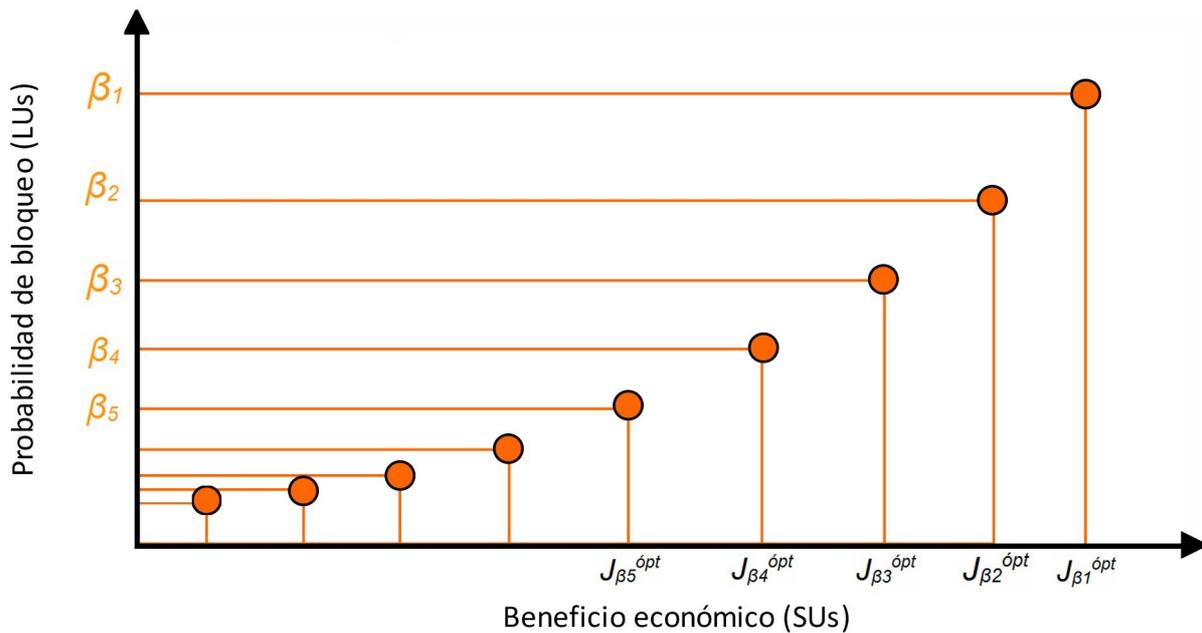


Figura 19: Frente de Pareto esperado mediante la formulación basada en programación lineal

Recapitulando, éstos son los pasos seguidos en este apartado para obtener el frente de Pareto:

1. Obtener las matrices de probabilidades de transición y los costes por estado asociados a cada control, siguiendo el proceso de uniformización descrito a partir de la página 41.

2. Definir los vectores R , para la función objetivo, y C , para la restricción de la probabilidad de bloqueo para LUs, así como los vectores auxiliares necesarios para el programa de optimización utilizado.
3. Resolver el problema de optimización (MATLAB) y obtener, a partir del vector Φ , la matriz Q con las probabilidades asociadas a cada control en cada estado.
4. Generar la matriz de probabilidades de transición óptima y los vectores de coste por estado relativos a la política obtenida.
5. Calcular el valor medio asociado a cada coste individual y representarlo en el frente de Pareto para distintos valores de bloqueo de usuarios primarios (β_L).

En general, en comparación con la asignación de recursos en las redes inalámbricas tradicionales, una red de radio cognitiva tiene más restricciones de diseño debido a la presencia de usuarios primarios. Además, la monitorización del espectro implica una serie de consideraciones adicionales.

La aproximación formulada en este apartado requiere un control centralizado para disponer de toda la información necesaria para la optimización. En comparación con el resto de esquemas de asignación de recursos, el centralizado ofrece el mejor rendimiento pero tiene una sobrecarga considerable debida al intercambio de la señalización necesaria para actualizar la información. Este problema puede suponer una reducción importante en la eficiencia del sistema, por lo que la aproximación centralizada es apropiada para escenarios con una topología simple, con movilidad reducida y con una capacidad computacional elevada.

3.3. Caracterización mediante simulación.

La tercera herramienta para la caracterización del sistema será la **simulación por eventos discretos**. El objetivo de esta parte es modelar el comportamiento real del sistema desde un punto de vista más práctico, observando los distintos “eventos” que ocurren en el sistema, como las llegadas de los distintos tipos de usuarios y la evolución del sistema por los diferentes estados a lo largo del tiempo.

Los resultados obtenidos con cada simulación se almacenarán y analizarán estadísticamente para llegar a un resultado que permita verificar la validez de la política elegida y

los resultados de los apartados anteriores.

3.3.1. Introducción a la Simulación por Eventos Discretos.

La simulación se define como la imitación del comportamiento de un proceso o sistema real a lo largo del tiempo mediante la generación de una historia artificial del mismo y la observación de dicha historia para extraer conclusiones sobre las características del funcionamiento de dicho sistema en la realidad.

La simulación de un sistema requiere, como paso previo, la realización de un “modelo” del sistema. Éste es una representación abstracta del sistema real, y estará basado en los parámetros y reglas de funcionamiento que caracterizan al sistema. Es especialmente importante obtener un modelo con un grado de detalle adecuado, que no contenga detalles irrelevantes que lo compliquen y lo hagan más costoso de lo necesario; del mismo modo que no se deben omitir detalles esenciales, dando lugar a resultados erróneos.

Particularizando, el simulador será, por tanto, un programa de ordenador en MATLAB que representará la evolución del sistema real de acceso al espectro mediante un modelo del mismo.

Existen una serie de circunstancias en las cuales se considera que la simulación no es una herramienta adecuada para estudiar un sistema, por ejemplo, cuando existen modelos analíticos del sistema a simular. En este Proyecto, la simulación se utiliza precisamente para validar los procesos analíticos descritos en los apartados anteriores.

Se define **estado de un sistema** como el conjunto de variables necesarias para describir el sistema en cualquier momento dentro del ámbito estudiado. Las variables que definirán el estado del sistema modelado serán el número de usuarios primarios y el número de usuarios secundarios en el sistema. El resto de variables, como los precios que ofertan los usuarios no licenciados o las tasas de llegada/salida, definirán el escenario particular que se simula, pero no el estado concreto dentro de dicho escenario.

En función de estas variables, los sistemas se pueden clasificar en dos grupos: discretos o continuos. Un **sistema discreto** es aquél en el que las variables de estado sólo toman valores de un conjunto discreto de puntos. En este caso, también se dice que el espacio de estados del sistema es discreto. El tipo de sistema que se va a modelar es discreto, porque su estado vendrá representado por el número y tipo de usuarios que accedan al espectro.

Dado que la simulación representará la ocupación espectral a lo largo del tiempo, será necesaria una gestión temporal en el simulador. En este sentido, existen dos **mecanismos básicos para controlar el avance del tiempo**:

1. **Discretización temporal** (*time slicing*). Consiste en generar un modelo con un estado inicial, asociarlo a un reloj virtual que se incrementaría en intervalos suficientemente pequeños de tiempo y, en cada uno de estos intervalos, se examinaría el sistema para decidir si se ha producido algún evento o suceso que cambie el estado del sistema. Esta aproximación no resulta práctica en la mayoría de las ocasiones, ya que hay que simular todos los intervalos de tiempo de la simulación, tanto aquéllos en los que se producen eventos como en los que no, lo cual puede resultar demasiado ineficiente. Lo más razonable sería detenerse únicamente en los instantes de tiempo en los que realmente se produzca un evento.
2. **Avance por eventos discretos**. En este caso, el reloj avanza directamente al instante temporal en el que realmente ocurre un evento en el sistema. Es decir, el reloj “salta” de un instante de ocurrencia de un evento al siguiente. Ésta será la aproximación utilizada, dada la reducción de carga computacional y rapidez de ejecución con respecto a la alternativa anterior [30].

La simulación por eventos discretos es una herramienta de análisis numérico de los sistemas discretos que se basa precisamente en **simular únicamente los momentos en los que se producen eventos**, adelantando el reloj de simulación directamente a estos momentos.

Se define como **evento** cualquier suceso instantáneo que cambia el estado del sistema. Los eventos pueden ser “endógenos”, si se generan dentro del sistema, o “exógenos”, si se generan fuera (en lo que se conoce como “entorno del sistema”). En este Proyecto habrá cuatro tipos de eventos: llegadas y salidas de usuarios primarios y llegadas y salidas de usuarios secundarios. En realidad el número total de eventos dependerá del tipo de usuario secundarios que exista. Cada tipo de usuario secundario y los primarios tendrán asociados dos eventos, uno de entrada y otro de salida. El simulador deberá generar los eventos que se producirán en el sistema en el futuro, de los cuales se generará un registro (se planificará) que contendrá, al menos, el tipo de evento y el momento en que ocurrirá.

Se define como **entidad** cualquier objeto o componente del sistema que requiere una representación explícita en el modelo. Por ejemplo, en un protocolo DSA: los usuarios, el

espectro, etc. Las entidades tienen **atributos**, que son sus propiedades. En este trabajo, por ejemplo, el tipo de usuario (primario o secundario), el tipo de usuario secundario, el número de canales en el espectro, etc.

La **Lista de Eventos Futuros (F.E.L.)** es una tabla o lista que contiene los registros con los eventos futuros planificados previamente. Estará ordenada según el tiempo de ocurrencia de los eventos.

Se define como **actividad** aquel periodo de tiempo de cierta duración del que se sabe cuándo comienza. Este tiempo puede estar definido por una distribución estadística. En este caso las actividades serán los tiempos de servicio de los usuarios y los tiempos entre llegadas.

Por último, el **reloj** es la variable que representa el tiempo real simulado o el “tiempo de simulación”.

3.3.2. Caracterización del modelo propuesto mediante simulación.

El algoritmo empleado en la Simulación por Eventos Discretos es conocido como **Event-Scheduling / Time Advance** (que se podría traducir como “Programación de Eventos / Tiempo Anticipado”).

Este algoritmo consiste en la generación de los momentos del “tiempo de simulación” en que se producirán los eventos, almacenar el registro del evento generado en la posición adecuada de la Lista de Eventos Futuros, y extraer los eventos cronológicamente, actualizando el reloj al momento de ocurrencia del evento extraído.

En este Proyecto, **la extracción de un evento de la F.E.L. puede implicar un cambio en el estado del sistema y será necesario planificar el momento de ocurrencia de uno o más eventos.**

En un momento cualquiera t del tiempo de simulación, la Lista de Eventos Futuros contiene todos los eventos futuros previamente planificados, con sus tiempos de ocurrencia asociados (t_1, t_2, \dots). La F.E.L. está ordenada cronológicamente, por lo que los tiempos de evento cumplen:

$$t \leq t_1 \leq t_2 \leq t_3 \leq \dots \leq t_n \quad , \quad (96)$$

donde t es el valor del reloj, es decir, el valor actual del tiempo de simulación. El evento asociado a t_i es conocido como “evento inminente”. Una vez que el estado del sistema está actualizado al tiempo de simulación correspondiente a $reloj = t$, el reloj se hace avanzar al momento $reloj = t_i$, el registro del evento inminente es eliminado de la F.E.L. y el evento se ejecuta, lo que implica una variación en el estado del sistema y, en ciertos casos, la generación de al menos un evento futuro, que será insertado en su posición correspondiente de la F.E.L. Este proceso se repite hasta el fin de la simulación.

Cuando se extrae de la F.E.L. el evento inminente, éste se ejecuta y como consecuencia se insertan nuevos eventos en la lista. La implementación de las acciones a realizar cuando se “ejecuta” un evento dependerá de la orientación que se le dé al simulador. Habitualmente se consideran dos modelos de simulador: orientado a eventos u orientado a procesos.

1. **Simulador orientado a eventos.** En este caso, el propio evento incluye la implementación de las acciones a realizar cuando se ejecuta el evento. Es decir, el propio “objeto” evento incluye una función del tipo *ejecutar()* o *procesar()* que se invoca tras extraer el evento y avanzar el reloj de simulación.
2. **Simulador orientado a procesos.** En este caso, las entidades que intervienen en el sistema se encargan de procesar los eventos, es decir, de realizar las acciones adecuadas en función del tipo de evento ocurrido. Así, se incluye información adicional en la F.E.L., además del tipo de evento y del instante de ocurrencia, se almacena el identificador de la entidad sobre la cuál se produce el evento. Tras extraerlo y avanzar el reloj, se invoca una función tipo *procesarEvento(Evento)* que implementa la entidad sobre la que se produce el evento, y a la que se pasa el propio “objeto” evento o el tipo de evento como parámetro. Éste es el modelo adoptado en este Proyecto.

La implementación de este simulador se ha llevado a cabo en MATLAB, al igual que los procedimientos analíticos de los subapartados anteriores, lo cual facilitará la comparación y procesamiento de los resultados obtenidos. Los parámetros de este simulador son:

Parámetros de entrada:

- **Tasas de llegada** de cada tipo de usuario, λ_{LU} y λ_{SU} .
- **Tasas de salida** de cada tipo de usuario, μ_{LU} y μ_{SU} .
- **Número de canales** del espectro, N , o número máximo de usuarios en el sistema.

- Un vector con la **probabilidad de cada tipo de usuario secundario**, U_k .
- Un vector con el **precio que ofrece cada tipo de usuario secundario**, R_k .
- El valor para ponderar cada tipo de objetivo, α .
- El número de **iteraciones** o pasos de simulación que realizará el simulador.
- Una variable booleana, **debug**, para activar o desactivar las notificaciones del simulador.
- Una variable entera, **pausa**, que representa el tiempo en segundos entre pasos del simulador. Si su valor es 0, no habrá pausa y la simulación se ejecutará a máxima velocidad. Si su valor está entre 1 y 4, ambos inclusive, éste será el tiempo entre pasos, útil para observar la evolución del sistema. Si su valor es mayor que 4, el avance entre pasos de simulación se producirá cuando se pulse una tecla cualquiera.

Parámetros de salida:

- La probabilidad de bloqueo de los LU obtenida en la simulación, B_L .
- El beneficio medio obtenido de los SU en la simulación, R_U .

Para llevar a cabo la simulación, es preciso, en primer lugar, **definir cuál será la política** por la que se regirá la entidad administradora del espectro en cuanto al acceso de usuarios secundarios. Esta política se obtendrá mediante el algoritmo Policy Iteration, ya descrito para la aproximación mediante MDP, en el subapartado “Cálculo de la política óptima”, página 44. Al usar el mismo algoritmo, se obtendrá la misma política óptima que en la primera aproximación. Así será posible observar el comportamiento “real” del sistema, verificando el cambio de control en cada estado y comprobando el número de bloqueos que se producen. La política óptima se almacena en un vector con tantos elementos como estados, de forma que durante la simulación se irá accediendo a la posición correspondiente para ver qué control seguir.

Una vez obtenida la política óptima, se lleva a cabo el **mapeo ocupación-estados** en función del número de canales disponibles, del mismo modo que se hizo en el análisis mediante MDP (ver página 35).

El siguiente paso consiste en calcular las **tasas de llegada de los distintos tipos de usuarios secundarios**, λ_{U_i} , a partir de su tasa de llegada general, λ_{SU} , en función de la probabilidad de cada tipo de SU, U_k :

$$\lambda_{U_i} = \lambda_{SU} \cdot U_i \quad (97)$$

A continuación **se crea la F.E.L. y se le añaden los primeros eventos de llegada.** Concretamente, para inicializar el sistema, se programará la llegada de un usuario primario y otra llegada por cada tipo de usuario secundario existente. Dado que el tiempo entre llegadas es exponencial, la generación de estos eventos se llevará a cabo mediante la instrucción *random()* de MATLAB:

$$t = \text{random}('exp', 1/\lambda),$$

donde t es un número aleatorio obtenido de una distribución exponencial con media $1/\lambda$. Por tanto, para generar los eventos iniciales, se llamará a la función anterior con $\lambda = \lambda_L, \lambda_{U1}, \lambda_{U2}, \dots, \lambda_m$.

La F.E.L. quedaría, inicialmente, así:

F.E.L.	
t_1	Llegada de un usuario primario
t_2	Llegada de un usuario secundario de tipo 1
t_3	Llegada de un usuario secundario de tipo 3
t_4	Llegada de un usuario secundario de tipo m
t_5	Llegada de un usuario secundario de tipo $m-2$
...	...

Nótese que es imprescindible que los eventos estén ordenados en función de su tiempo de ocurrencia, siendo $t_1 \leq t_2 \leq t_3 \leq t_4 \leq t_5$, independientemente del tipo de evento que sea y del tipo de usuario al que afecte.

Para una explicación más sencilla del funcionamiento del simulador, supóngase a modo de ejemplo un escenario con tres tipos de usuarios y dos canales. La F.E.L. inicial será:

F.E.L.	
$t_1 = 1$	Llegada de un usuario secundario de tipo 1
$t_2 = 3$	Llegada de un usuario primario
$t_3 = 4$	Llegada de un usuario secundario de tipo 3
$t_4 = 7$	Llegada de un usuario secundario de tipo 2

Los eventos existentes para el simulador en este sistema serán:

Número de evento	Descripción
0	Llegada de un usuario primario
1	Salida de un usuario primario
2	Llegada de un usuario secundario de tipo 1
3	Salida de un usuario secundario de tipo 1
4	Llegada de un usuario secundario de tipo 2
5	Salida de un usuario secundario de tipo 2
6	Llegada de un usuario secundario de tipo 3
7	Salida de un usuario secundario de tipo 3

Tabla 3: Tipos de eventos para el caso de 3 tipos de usuarios secundarios

Tras crear la F.E.L., **se define el estado inicial** con el valor “1” (el primer estado es aquél en el que no hay usuarios de ningún tipo en el espectro).

A partir de este momento **se inicia la simulación** propiamente dicha:

1. Se extrae el primer elemento (evento inminente) de la F.E.L. En este caso, la llegada de un usuario secundario de tipo 1.
2. Se avanza el reloj de la simulación al momento del evento inminente: $t = t_1$.
3. Se ejecuta el evento extraído, es decir, se procesa. El procesamiento realizado dependerá del tipo de evento extraído, como muestra la tabla 4.

Las distintas salidas asociadas a los eventos de llegada, se programa de forma análoga a las llegas, pero en este caso se usa la tasa de salida, μ , en lugar de la de llegada λ :

$$t_i = \text{random}('exp', t + 1/\mu),$$

donde t es el tiempo de la simulación en el momento de generar el nuevo evento. Los eventos deben generarse a partir del tiempo actual de la simulación para no programar eventos anteriores al reloj. Lo mismo ocurre con la programación de las nuevas llegadas.

Evento	Procesamiento
0: llega LU	Se incrementa el número de peticiones LU. Si existe algún canal vacío, el usuario accede al sistema, se programa su salida y se incrementan el número de accesos LU y el número actual de LU en el sistema. Si todos los canales están ocupados, se rechaza la petición y se contabiliza el bloqueo. Se programa una nueva llegada de LU.
1: sale LU	Se decrementa el número actual de LU en el sistema.
2: llega SU 1	Si existe algún canal disponible y si el control asociado al estado actual lo permite ($u \geq 1$), el usuario accede al sistema, se programa su salida y se incrementa el número actual de SU en el sistema y el número total de accesos de SU tipo 1 al sistema. Se programa una nueva llegada de SU tipo 1 en el sistema.
3: sale SU 1	Se decrementa el número actual de SU en el sistema.
4: llega SU 2	Si existe algún canal disponible y si el control asociado al estado actual lo permite ($u \geq 2$), el usuario accede al sistema, se programa su salida y se incrementa el número actual de SU en el sistema y el número total de accesos de SU tipo 2 al sistema. Se programa una nueva llegada de SU tipo 2 en el sistema.
5: sale SU 2	Se decrementa el número actual de SU en el sistema.
6: llega SU 3	Si existe algún canal disponible y si el control asociado al estado actual lo permite ($u = 3$), el usuario accede al sistema, se programa su salida y se incrementa el número actual de SU en el sistema y el número total de accesos de SU tipo 3 al sistema. Se programa una nueva llegada de SU tipo 3 en el sistema.
7: sale SU 3	Se decrementa el número actual de SU en el sistema.

Tabla 4: Ejecución de los diferentes eventos

Volviendo al ejemplo, tras procesar el primer evento de la F.E.L., y asumiendo que el control en el estado inicial permitiría el acceso del usuario secundario de tipo 1, el estado del sistema sería el siguiente:

Reloj de simulación	Estado	LU en el sistema	SU en el sistema
t = 1	4	0	1

F.E.L.	
$t_2 = 3$	Llegada de un usuario primario
$t_3 = 4$	Llegada de un usuario secundario de tipo 3
$t_4 = 7$	Llegada de un usuario secundario de tipo 2
$t_5 = 10$	Salida de un usuario secundario de tipo 1.
$t_6 = 11$	Llegada de un usuario secundario de tipo 1.

El valor del estado del sistema es “4” de acuerdo al mapeo realizado (ver tabla 2, página 35). Los dos últimos eventos de la F.E.L. son los que se programan como consecuencia de la ejecución del evento anterior.

Ahora el evento inminente a extraer sería la llegada de un usuario primario. Dado que existe un canal disponible, éste podría acceder al espectro:

Reloj de simulación	Estado	LU en el sistema	SU en el sistema
$t = 3$	5	1	1
F.E.L.			
$t_3 = 4$	Llegada de un usuario secundario de tipo 3		
$t_4 = 7$	Llegada de un usuario secundario de tipo 2		
$t_5 = 10$	Salida de un usuario secundario de tipo 1.		
$t_6 = 11$	Llegada de un usuario secundario de tipo 1.		
$t_7 = 13$	Llegada de un usuario primario.		
$t_8 = 15$	Salida de un usuario primario.		

En cada actualización del estado del sistema se incrementa el número de iteraciones. Cuando se alcanza el máximo (definido por el usuario como parámetro de entrada del simulador), finaliza la simulación y se calculan los parámetros de salida del simulador a partir de las distintas variables.

Para obtener los parámetros de salida del simulador se utilizan algunas variables ya introducidas en la tabla 4:

- El número total de usuarios secundarios de cada tipo que han accedido al espectro a lo largo de la simulación: $N_{U1}, N_{U2}, \dots, N_{Um}$.
- El número total de usuarios primarios que intentan acceder al sistema, N_{LT} .

- El número total de usuarios primarios que no acceden al sistema, es decir, el número total de bloqueos durante la simulación, N_{LB} .

Con estas variables, una vez realizado el total de iteraciones de una simulación, se obtienen los dos parámetros de salida del simulador:

- El **bloqueo de los usuarios primarios** se calculará simplemente con el porcentaje de bloqueos respecto al total de intentos:

$$B_L = \frac{N_{LB}}{N_{LT}} \quad (98)$$

- El **beneficio medio obtenido** se calculará a partir del beneficio de cada tipo de usuario:

$$\begin{aligned} B_{U1} &= b_1 \cdot N_{U1} \\ B_{U2} &= b_2 \cdot N_{U2} \\ &\vdots \\ B_{Um} &= b_m \cdot N_{Um} \end{aligned} \quad (99)$$

$$B_T = B_{U1} + B_{U2} + \dots + B_{Um} \rightarrow \bar{B} = B = \frac{B_T}{T} \quad (100)$$

donde T es el tiempo total de simulación.

3.3.3. Análisis estadístico de los resultados.

Los parámetros obtenidos como resultado de una simulación no serán definitivos, es decir, no serán los que se representen en el frente de Pareto, sino que, previo a la representación, se llevará a cabo un **análisis estadístico** de los resultados obtenidos.

Para cada escenario simulado, se obtendrá la **media** y la **varianza** de la probabilidad de bloqueo y el beneficio obtenidos en diversas réplicas y, a partir de ambos parámetros, sus **intervalos de confianza**.

Para poder calcular el intervalo de confianza de cualquier medida, es necesario obtener ésta varias veces. Por ejemplo, para calcular el intervalo de confianza del beneficio obtenido, hay que obtener varias veces el beneficio medio obtenido en un escenario dado. Cada uno de los

valores obtenidos será, generalmente, una variable aleatoria, pero es muy importante que estas variables sean independientes entre sí para una correcta estimación del intervalo de confianza. Para ello se realiza el **Método de las Réplicas Independientes**.

El método de las réplicas independientes consiste en la realización de réplicas del experimento, es decir, repeticiones del mismo con las mismas condiciones, pero con una semilla distinta para la generación de números aleatorios que **asegure que no existe solapamiento entre las secuencias de número aleatorios** que se obtienen a partir de funciones pseudoaleatorias utilizadas en el simulador para generar nuevos eventos. En consecuencia, la base del método es que las medidas de las observaciones obtenidas con réplicas independientes son independientes, aunque las muestras dentro de una misma réplica estén correladas.

El proceso que se lleva a cabo es el siguiente:

1. Se definen los parámetros de entrada del simulador, dentro de los que están los parámetros del escenario estudiado (tasas, número de canales, etc.) y también la variable N_s , que indica cuántas réplicas se van a simular para cada escenario. El número de escenarios simulados viene dado por la definición de la variable α . Del mismo modo que en la aproximación basada en MDP, se calculará un punto del frente de Pareto para cada valor de α , por lo que esta variable se definirá como un vector, A , con los distintos valores del parámetro.
2. Por cada valor de α definido, se realizará el número de simulaciones (réplicas) previamente especificado. En cada simulación se obtendrá un valor de bloqueo de usuarios primarios y un beneficio de los usuarios secundarios, por lo que, una vez realizadas todas las réplicas, se tendrá un total de N_s valores de bloqueo y otros tantos de beneficio.
3. Se obtienen la media y la varianza de los N_s valores de bloqueo y de beneficio. Las medias se denotarán por \bar{x}_L y \bar{x}_U , y las varianzas por σ_L^2 y σ_U^2 .
4. Con los valores de media y varianza, se calcula el **intervalo de confianza al 95%** asociado a cada variable:

$$P \left[\bar{x}_L - z_{1-\alpha/2} \cdot \frac{\sigma_L}{\sqrt{n}} \leq \mu_L \leq \bar{x}_L + z_{1-\alpha/2} \cdot \frac{\sigma_L}{\sqrt{n}} \right] = 1 - \alpha$$

$$P \left[\bar{x}_U - z_{1-\alpha/2} \cdot \frac{\sigma_U}{\sqrt{n}} \leq \mu_U \leq \bar{x}_U + z_{1-\alpha/2} \cdot \frac{\sigma_U}{\sqrt{n}} \right] = 1 - \alpha$$
(101)

Donde \bar{x} es la media obtenida de las N_s réplicas de simulación del escenario, σ su desviación típica y n el número de réplicas, es decir $n = N_s$. z es el valor de la distribución *t-student* asociado al intervalo de confianza del 95%. Es constante y su valor es 1.96.

La ecuación 101 indica que la probabilidad de que **la media real, μ** , esté dentro del

intervalo $\left[\bar{x} - z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right]$ es igual a $1-\alpha$. Evidentemente, en este caso α no tiene nada que ver con el valor de ponderación de los objetivos bloqueo/beneficio descrito en los parámetros de entrada del simulador. $1-\alpha$ es el coeficiente de confianza, y $100 \cdot (1-\alpha)$ es el nivel de confianza que, como se indicaba antes, será del 95%.

5. Una vez finalizadas todas las réplicas de todos los valores de α , **se representa en el frente de Pareto la media** obtenida para el bloqueo y el beneficio asociados a cada valor de α .

La figura 20 muestra el flujograma que representa los 5 pasos del proceso descrito.

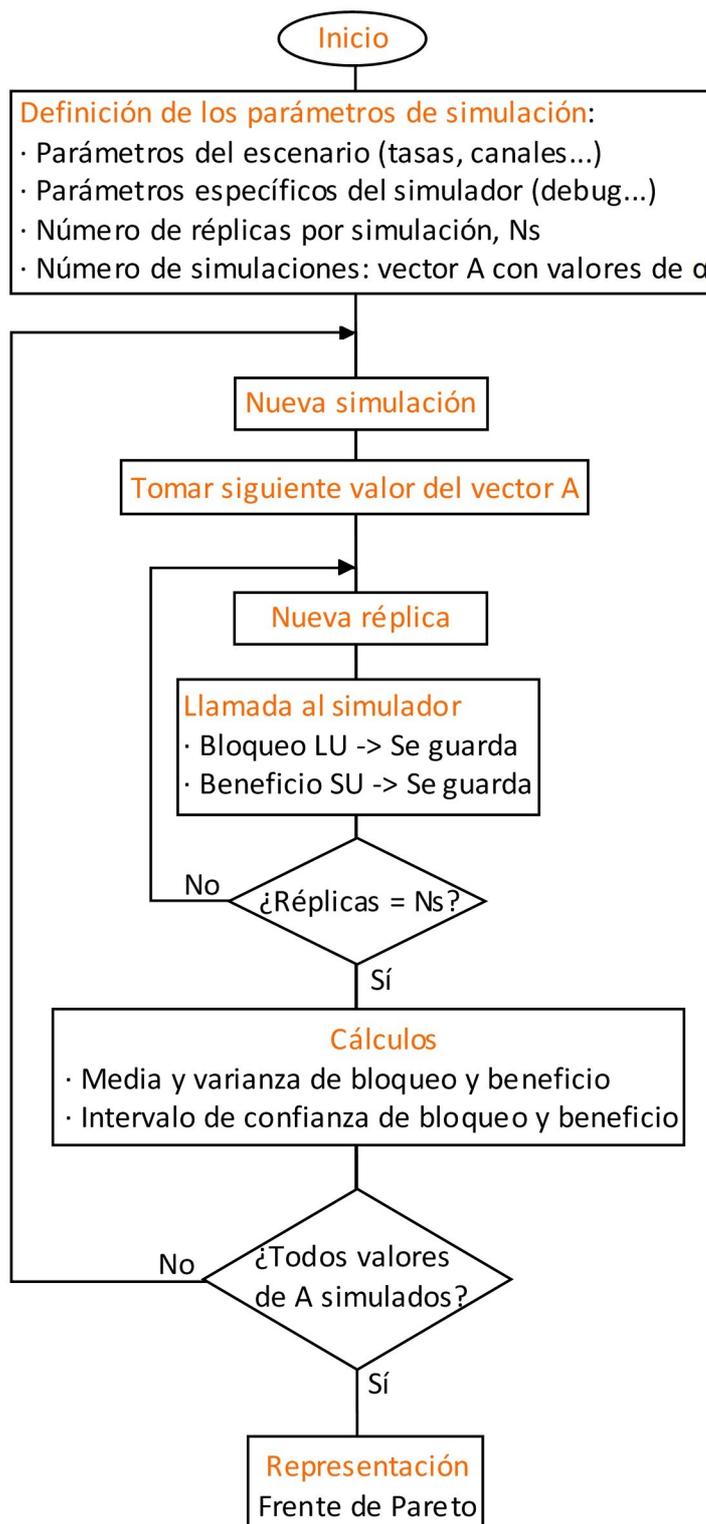


Figura 20: Flujograma del proceso de simulación

4. RESULTADOS.

En este apartado se presentan los distintos escenarios propuestos para probar el rendimiento del modelo expuesto en el tercer apartado. Para cada escenario, se muestran los resultados obtenidos en cada uno de los procesos de caracterización descritos anteriormente: Procesos de Decisión de Markov, optimización mediante programación lineal y la caracterización y validación mediante simulación.

4.1. Escenarios propuestos.

Los resultados asociados al modelo propuesto se obtendrán a partir de **cuatro escenarios diferentes** en términos de tráfico ofrecido, para observar el bloqueo y el beneficio obtenidos. Como se indicó anteriormente en el apartado 3.1.2, “Parámetros del escenario”, en la página 33, en cada uno de los escenarios formulados intervendrán los siguientes parámetros:

- λ_{LU} (llegadas/minuto): Tasa de llegada de los usuarios primarios al sistema.
- λ_{SU} (llegadas/minuto): Tasa de llegada de los usuarios secundarios al sistema.
- μ_{LU} (salidas/minuto): Tasa de salida de los usuarios primarios del sistema.
- μ_{SU} (salidas/minuto): Tasa de salida de los usuarios secundarios del sistema.
- **N** : Número de canales disponibles en el espectro.
- **Tipos de usuarios secundarios**. En función de su “puja” por el acceso.
- **Cantidad ofrecida** por cada tipo de usuario (euros/minuto).
- P_j . Probabilidad de llegada de cada tipo de usuario secundario.
- **Controles**. Los niveles de restricción al acceso de usuarios secundarios.

De estos parámetros, los cinco últimos serán comunes a los cuatro escenarios propuestos.

En cada uno de ellos **habrá 10 canales**, es decir, en un momento dado, podrá haber hasta 10 usuarios en el sistema. Habrá **tres tipos de usuarios secundarios**, y las **cantidades ofrecidas** por cada uno de ellos serán **0.3€, 0.2€ y 0.1€**. Además, se asume que las **probabilidades de llegada** de cada tipo de usuario serán del **50%, 40% y 10%**, respectivamente. Evidentemente, el **número de controles será igual a cuatro** (uno por cada tipo de usuario secundario más el control '0', que inhabilita su acceso al espectro).

Además, se asume que **la duración de las llamadas o accesos a la banda es la misma**, independientemente del tipo de usuario que la utilice. Es decir, $\mu_{LU} = \mu_{SU}$ en todos los escenarios.

A continuación se describe cada escenario en particular.

4.1.1. Escenario 1.

Se trata del escenario más común que se podría encontrar en una situación real. En él, hay un **mayor número de llegadas por minuto de usuarios primarios** que de secundarios, lo cual es razonable, ya que en la mayoría de situaciones, los usuarios secundarios haría uso del espectro licenciado de forma puntual, dado que conlleva un mayor coste que el acceso a la banda del propio operador.

Parámetro	Valor
λ_{LU}	30 llamadas/minuto
λ_{SU}	20 llamadas/minuto
$\mu_{LU} (\mu_{SU})$	5 llamadas/minuto

Tabla 5: Parámetros del escenario 1.

Las tablas 6, 7 y 8 muestran los resultados obtenidos para los distintos valores del parámetro α de ponderación, para el caso de las aproximaciones mediante MDP y simulación. Recuérdese que el parámetro que se somete a variación en el caso de la formulación mediante programación lineal es β , que representa el umbral de probabilidad de bloqueo admisible para los usuario primarios.

Evidentemente, al tratarse de procesos totalmente distintos, los resultados obtenidos en cada formulación difieren ligeramente entre sí. A pesar de ello, sí que se aprecia un parecido considerable en los tres frentes de Pareto obtenidos, con una probabilidad de bloqueo máxima de entre el 12 % y el 14 %, más que razonable si se tiene en cuenta que los beneficios derivados de este bloqueo alcanzan casi 75 céntimos de euro por minuto (45 €/hora).

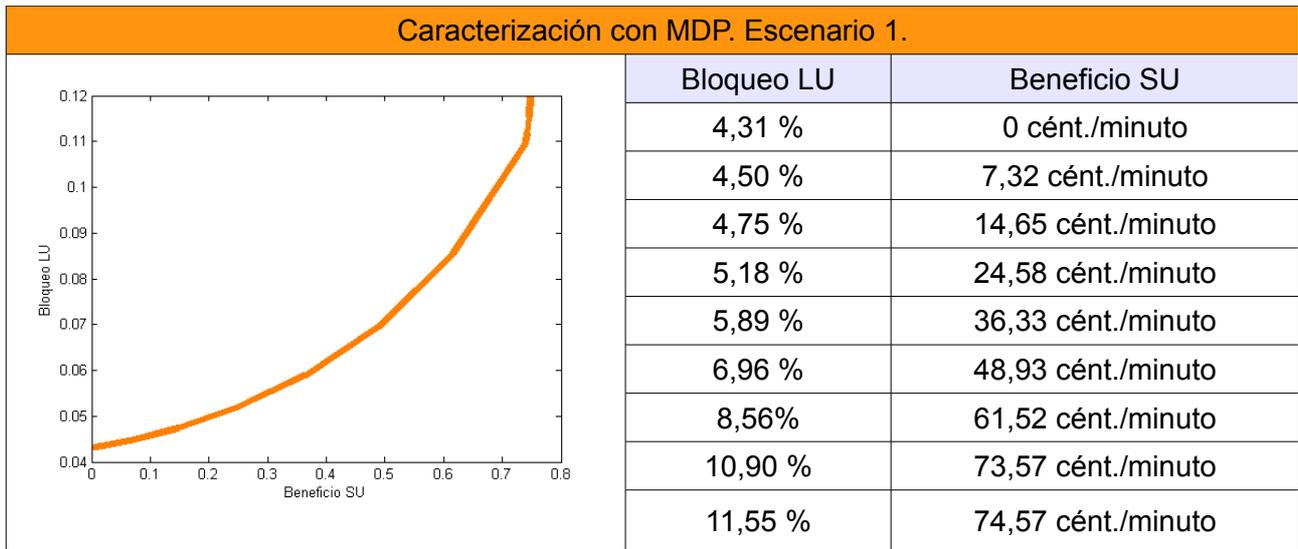


Tabla 6: Resultados de caracterización mediante MDP. Escenario 1.

De acuerdo con los resultados obtenidos en la tabla anterior, en la formulación basada en la optimización con CVX se ha realizado un barrido del parámetro β dentro de un rango similar al obtenidos para el bloqueo de usuarios primario en la caracterización con MDP (primera columna de la tabla 6). La tabla siguiente muestra estos resultados.

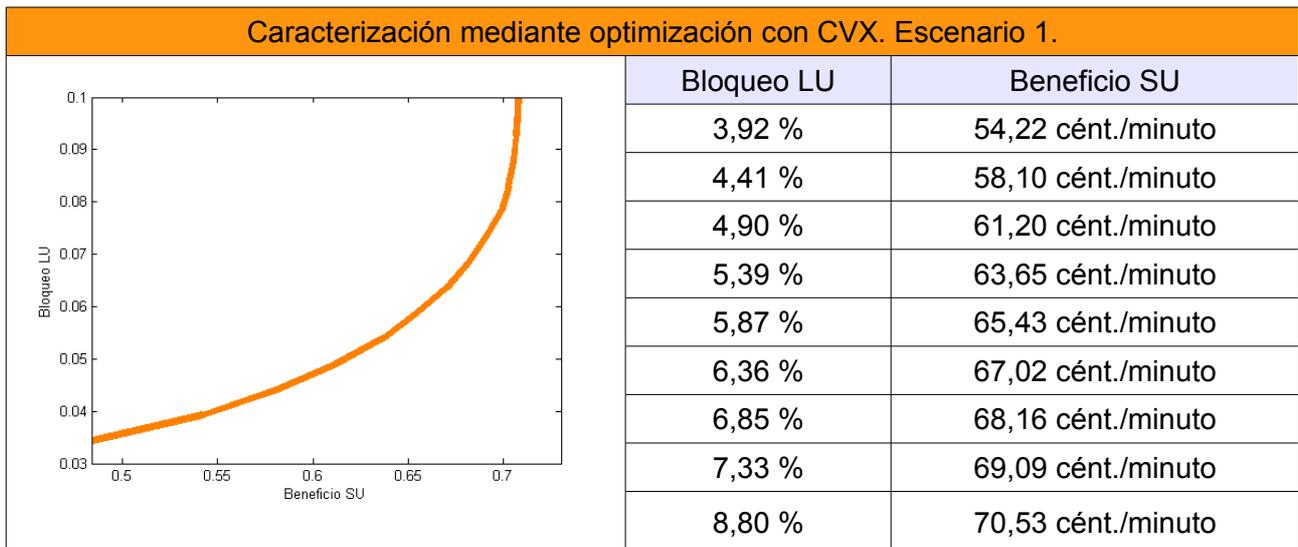


Tabla 7: Resultados de la caracterización mediante optimización con CVX. Escenario 1.

Puede llamar la atención la oscilación que existe en los resultados obtenidos mediante la simulación cuando los valores de α están por debajo de 0.8. Esto se debe al carácter aleatorio de la simulación, por lo que se realiza el análisis estadístico de los resultados, detallado en el apartado 3.3.3., en la página 70.

Caracterización mediante simulación. Escenario 1.

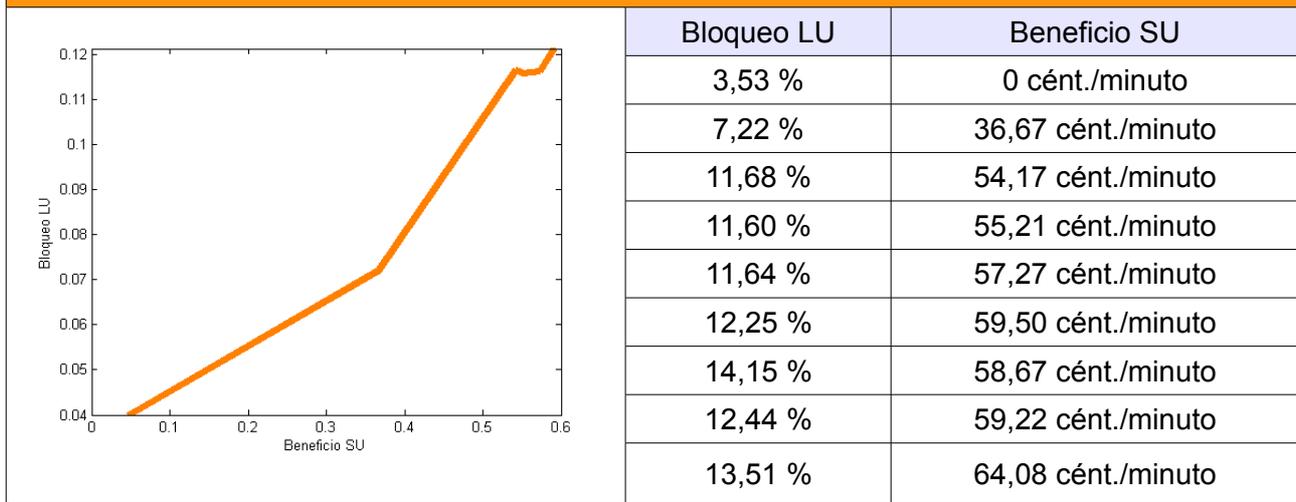


Tabla 8: Resultados de la caracterización mediante simulación. Escenario 1.

En la tabla 9 se muestran muestran estos resultados, en los que se incluyen la media de bloqueo y beneficio obtenidos, así como los intervalos de confianza y la varianza asociada a cada uno de ellos.

Análisis estadístico de los resultados obtenidos mediante simulación. Escenario 1.

Bloqueo LU				Beneficio SU			
Intervalo de confianza			Varianza	Intervalo de confianza			Varianza
0,1222 ←	0,1351	→ 0,1479	$0,9 \cdot 10^{-3}$	0,6218 ←	0,6408	→ 0,6598	$1,9 \cdot 10^{-3}$
0,1121 ←	0,1244	→ 0,1368	$0,8 \cdot 10^{-3}$	0,5642 ←	0,5922	→ 0,6202	$4,1 \cdot 10^{-3}$
0,1267 ←	0,1415	→ 0,1562	$1,1 \cdot 10^{-3}$	0,5583 ←	0,5867	→ 0,6150	$4,2 \cdot 10^{-3}$
0,1082 ←	0,1225	→ 0,1369	$1,1 \cdot 10^{-3}$	0,5706 ←	0,5950	→ 0,6194	$3,1 \cdot 10^{-3}$
0,1030 ←	0,1164	→ 0,1299	$0,9 \cdot 10^{-3}$	0,5572 ←	0,5727	→ 0,5881	$1,2 \cdot 10^{-3}$
0,1024 ←	0,1160	→ 0,1295	$1 \cdot 10^{-3}$	0,5300 ←	0,5521	→ 0,5741	$2,5 \cdot 10^{-3}$
0,1059 ←	0,1168	→ 0,1276	$0,6 \cdot 10^{-3}$	0,5095 ←	0,5417	→ 0,5739	$5,4 \cdot 10^{-3}$
0,0653 ←	0,0722	→ 0,0790	$0,2 \cdot 10^{-3}$	0,3411 ←	0,3667	→ 0,3923	$3,4 \cdot 10^{-3}$
0,0276 ←	0,0353	→ 0,0430	$0,3 \cdot 10^{-3}$	0 ←	0	→ 0	0

Tabla 9: Análisis estadístico de los resultados obtenidos mediante simulación. Escenario 1.

En la tabla anterior se observa que la varianza obtenida para los dos parámetros de interés es muy baja, lo que indica una gran precisión en los resultados obtenidos, como muestra la poca amplitud de los intervalos de confianza calculados.

4.1.2. Escenario 2.

Se trata de un escenario en el que sigue existiendo un tráfico ofrecido mayor por parte de los usuarios con licencia con respecto a los usuarios secundarios, sin embargo, en este caso, a diferencia del anterior, ambos valores son considerablemente menores.

Parámetro	Valor
λ_{LU}	10 llamadas/minuto
λ_{SU}	5 llamadas/minuto
$\mu_{LU} (\mu_{SU})$	5 llamadas/minuto

Tabla 10: Parámetros del escenario 2.

El interés de este escenario se basa en comprobar los beneficios que se podrían obtener con la implantación de un esquema de reparto espectral de este tipo en áreas con baja densidad de población.

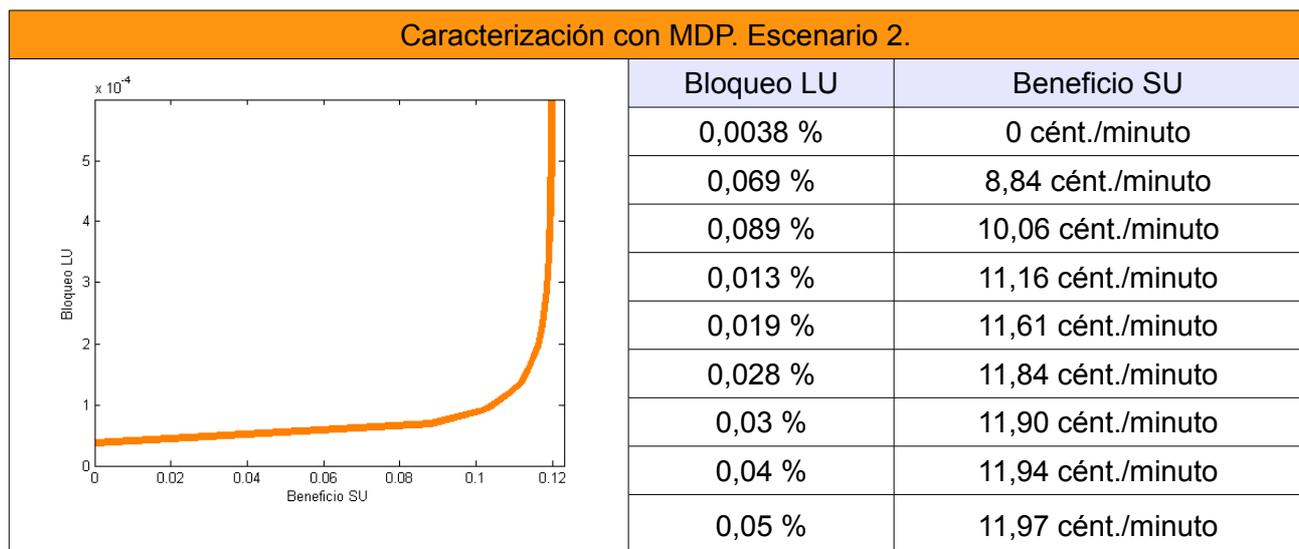


Tabla 11: Resultados de la caracterización con MDP. Escenario 2.

En la tabla 11 se puede observar el bloqueo obtenido. En la primera fila se muestra el bloqueo intrínseco de la banda licenciada, que sin haber usuarios secundarios (beneficio cero), es prácticamente nulo.

Sin embargo, a medida que varía el peso de los usuarios secundarios en el sistema, se observa cómo el bloqueo se mantiene en porcentajes muy bajos, con un beneficio económico de casi 12 céntimos por minuto.

Caracterización mediante optimización con CVX. Escenario 2.

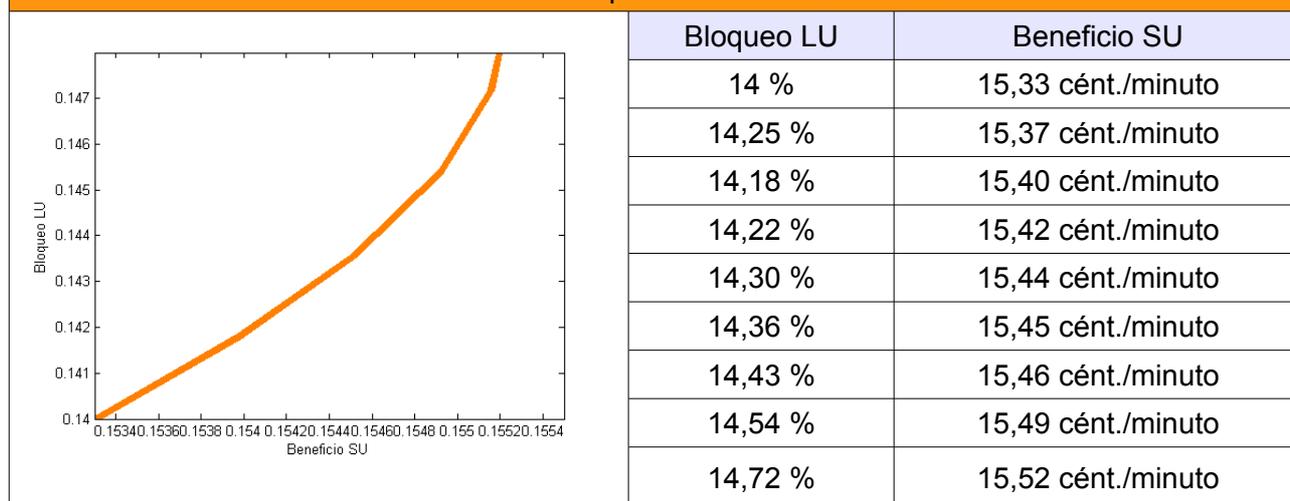


Tabla 12: Resultados de la caracterización mediante optimización con CVX. Escenario 2.

La tabla 12 muestra los resultados obtenidos mediante la optimización basada en programación lineal. En este caso, se obtienen bloqueos considerablemente mayores, lejos de ser despreciables, como ocurre en la aproximación basada en MDP. Por otro lado, se observa que los beneficios reportados son también mayores, aunque crecen en menor medida que el bloqueo.

El motivo por el que con CVX no se alcanzan valores de bloqueo tan bajos como con los procesos de decisión de Markov se debe a limitaciones del propio software utilizado, que no es capaz de resolver el problema con restricciones "límite", con probabilidades muy cercanas al cero.

Caracterización mediante simulación. Escenario 2.

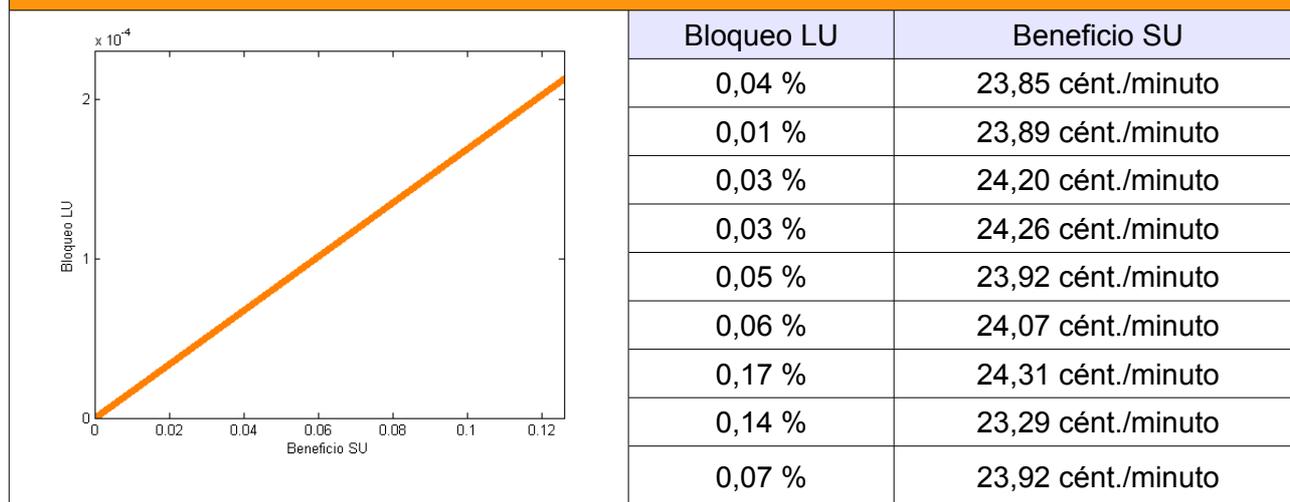


Tabla 13: Resultados de la caracterización mediante simulación. Escenario 2.

En la tabla 13 se observa que el simulador, al igual que el MDP, obtiene probabilidades de bloqueo muy bajas, aunque da unos beneficios superiores a los anteriores.

En este caso el Frente de Pareto obtenido se asemeja a una recta por que para la representación se lleva a cabo una interpolación lineal entre dos resultados obtenidos relativamente distanciados en el eje X de la figura.

La tabla 14 muestra que los resultados obtenidos son más precisos que en la misma formulación del escenario 1, como demuestra la varianza, prácticamente nula.

Análisis estadístico de los resultados obtenidos mediante simulación. Escenario 1.							
Bloqueo LU				Beneficio SU			
Intervalo de confianza			Varianza	Intervalo de confianza			Varianza
0 ←	0,0007	→ 0,0014	$0,351 \cdot 10^{-5}$	0,2322 ←	0,2392	→ 0,2461	$0,3810 \cdot 10^{-3}$
0,0003 ←	0,0014	→ 0,0024	$0,8508 \cdot 10^{-5}$	0,2268 ←	0,2329	→ 0,2389	$0,2836 \cdot 10^{-3}$
0,0006 ←	0,0017	→ 0,0028	$0,9427 \cdot 10^{-5}$	0,2359 ←	0,2431	→ 0,2504	$0,4131 \cdot 10^{-3}$
0,0001 ←	0,0006	→ 0,0011	$0,1963 \cdot 10^{-5}$	0,2334 ←	0,2407	→ 0,2481	$0,4211 \cdot 10^{-3}$
0,0001 ←	0,0005	→ 0,0009	$0,1271 \cdot 10^{-5}$	0,2309 ←	0,2392	→ 0,2475	$0,5384 \cdot 10^{-3}$
0 ←	0,0003	→ 0,0006	$0,1425 \cdot 10^{-5}$	0,2374 ←	0,2426	→ 0,2477	$0,2090 \cdot 10^{-3}$
0 ←	0,0003	→ 0,0006	$0,1460 \cdot 10^{-5}$	0,2345 ←	0,2420	→ 0,2496	$0,4424 \cdot 10^{-3}$
0 ←	0,0001	→ 0,0002	$0,0308 \cdot 10^{-5}$	0,2314 ←	0,2389	→ 0,2464	$0,4371 \cdot 10^{-3}$
0 ←	0,0004	→ 0,0003	$0,1091 \cdot 10^{-5}$	0,2313 ←	0,2385	→ 0,2456	$0,3982 \cdot 10^{-3}$

Tabla 14: Análisis estadístico de los resultados obtenidos mediante simulación. Escenario 2.

4.1.3. Escenario 3.

En el tercer escenario se invierten los parámetros del escenario 1. En este caso el tráfico ofrecido por los usuarios secundarios es mucho mayor que el de los usuarios con licencia. Además ambos valores son considerablemente elevados. Se trata de un escenario que sería poco frecuente en situaciones reales, pero es interesante observar cómo se comporta el sistema en esta situación, y cuáles son los beneficios obtenidos, que en este caso sí pueden llevar un coste importante en términos de bloqueo para los usuarios primarios. Éstos son los parámetros de este escenario:

Parámetro	Valor
λ_{LU}	20 llamadas/minuto
λ_{SU}	50 llamadas/minuto
$\mu_{LU} (\mu_{SU})$	5 llamadas/minuto

Tabla 15: Parámetros del escenario 3.

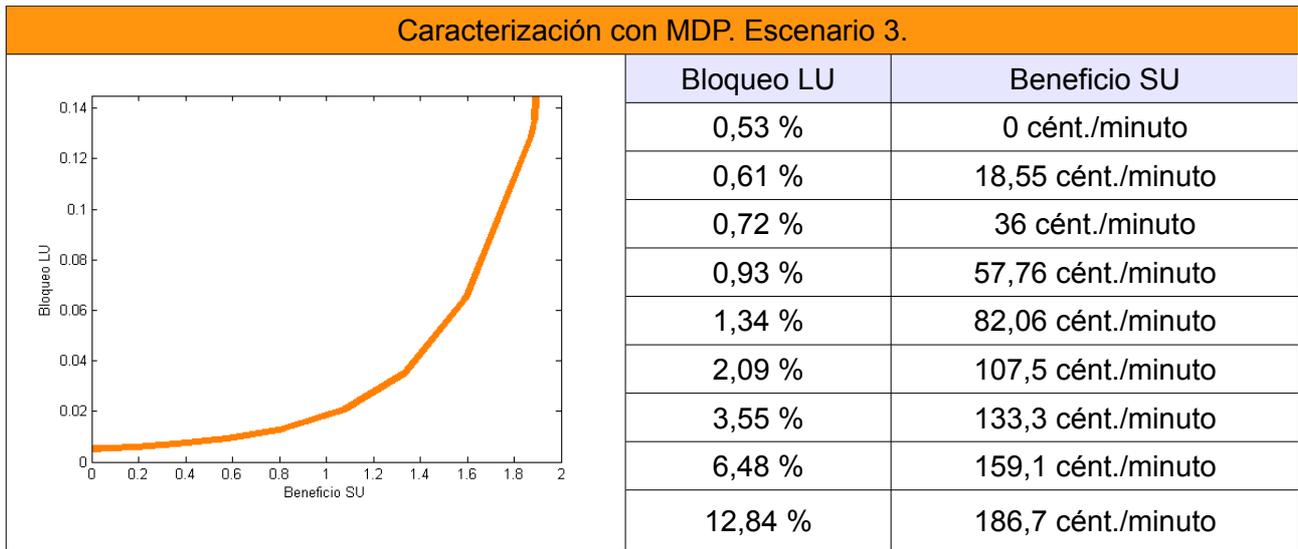


Tabla 16: Resultados de la caracterización con MDP. Escenario 3.

La tabla 16 muestra que en esta situación se pueden obtener beneficios considerablemente superiores a los de los escenarios anteriores, dada la gran demanda de canales por parte de los usuarios secundarios. Por otra parte, incluso para beneficios cercanos a los 2 €/minuto, la probabilidad de bloqueo de usuarios con licencia no llega al 13%, un valor que podría ser aceptable en función de los requisitos de la red y del operador.

Se trata de un valor bajo ya que no existen muchos usuarios primarios que intenten acceder al espectro (20 llamadas/minuto frente a las 50 llam./minuto de los usuarios secundarios). En otras condiciones, el bloqueo obtenido sería superior para valores similares de beneficio.

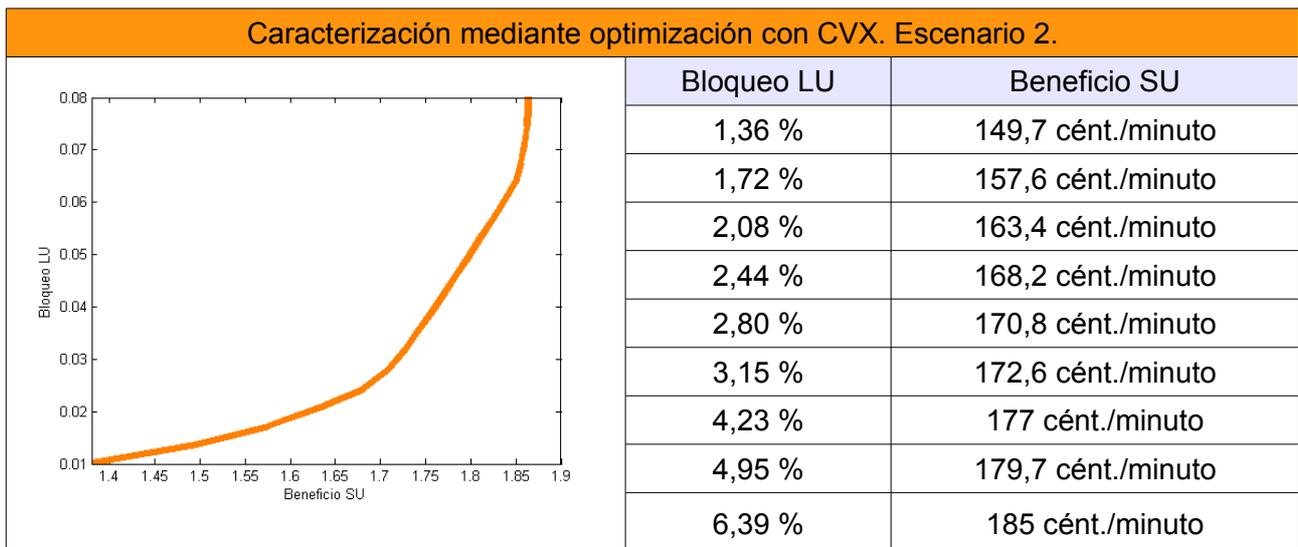


Tabla 17: Resultados de la caracterización mediante optimización con CVX. Escenario 3.

La tabla 17 muestra los resultados de la optimización con CVX. El beneficio obtenidos es muy parecido al anterior, pero en este caso se alcanza con valores menores de bloqueo. Se trata de resultados probablemente demasiado optimistas para una situación real, como se adelantaba en la introducción de este escenario.

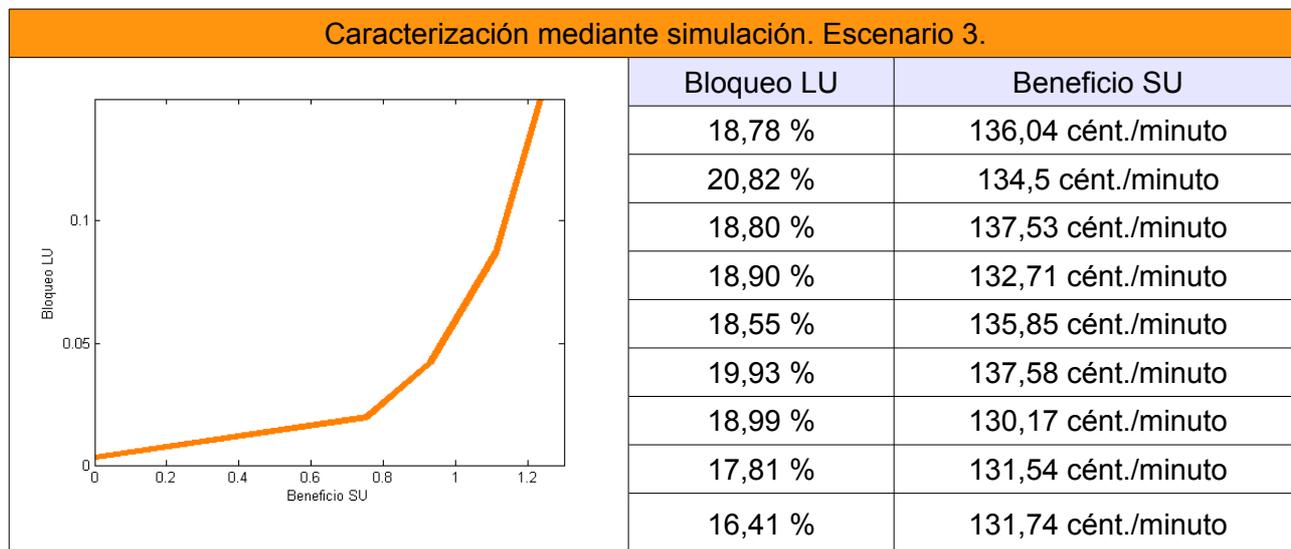


Tabla 18: Resultados de la caracterización mediante simulación. Escenario 3.

Los resultados de las simulaciones por eventos discretos arrojan unos valores más pesimistas, aunque quizás más realistas, en ambos parámetros.

En este escenario se observa un claro aumento en los beneficios obtenidos al haber muchos usuarios secundarios en el sistema.

Análisis estadístico de los resultados obtenidos mediante simulación. Escenario 3.

Bloqueo LU				Beneficio SU			
Intervalo de confianza			Varianza	Intervalo de confianza			Varianza
0,1735 ←	0,1878	→ 0,2021	0,0011	1,3258 ←	1,3604	→ 1,3951	0,0063
0,1890 ←	0,2082	→ 0,2275	0,0019	1,3121 ←	1,3435	→ 1,3750	0,0051
0,1725 ←	0,1880	→ 0,2034	0,0012	1,3349 ←	1,3753	→ 1,4156	0,0085
0,1705 ←	0,1890	→ 0,2074	0,0018	1,2858 ←	1,3271	→ 1,3684	0,0089
0,1682 ←	0,1855	→ 0,2028	0,0015	1,3229 ←	1,3585	→ 1,3940	0,0066
0,1861 ←	0,1993	→ 0,2126	0,0009	1,3312 ←	1,3758	→ 1,4203	0,0103
0,1751 ←	0,1899	→ 0,2047	0,0011	1,2724 ←	1,3017	→ 1,3311	0,0045
0,1612 ←	0,1781	→ 0,1949	0,0015	1,2814 ←	1,3154	→ 1,3494	0,0060
0,1457 ←	0,1641	→ 0,1826	0,0018	1,2835 ←	1,3174	→ 1,3513	0,0060

Tabla 19: Análisis estadístico de los resultados obtenidos mediante simulación. Escenario 3.

4.1.4. Escenario 4.

En el cuarto y último escenario, se estudia la siguiente situación:

Parámetro	Valor
λ_{LU}	50 llamadas/minuto
λ_{SU}	50 llamadas/minuto
$\mu_{LU} (\mu_{SU})$	5 llamadas/minuto

Tabla 20: Parámetros del escenario 4.

Se mantiene la abundancia de usuarios secundarios en el sistema (solicitando acceso), y además aumenta el tráfico ofrecido por los usuarios con licencia. El objetivo de este escenario es comprobar la evolución del sistema cuando está sometido a altas cargas de tráfico, y en el que los usuarios de ambos tipos (primarios y secundarios), están presentes con la misma frecuencia.

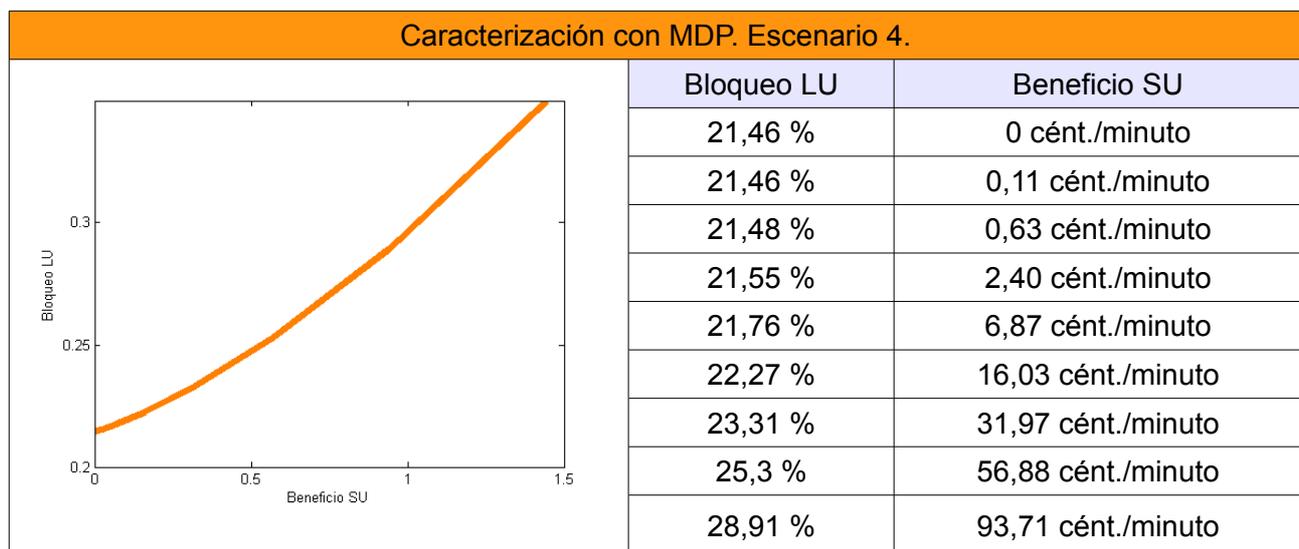


Tabla 21: Resultados de la caracterización con MDP. Escenario 4.

Los resultados de la tabla 21 muestran el beneficio obtenido en un punto medio entre el de los escenarios 2 y 3, sin embargo se aprecia un aumento muy importante en el bloqueo sobre los usuarios primarios, que alcanza casi el 30 % para obtener el máximo beneficio.

Se trata, evidentemente de una situación inaceptable, pues se perdería prácticamente 1 de cada 3 llamadas por parte de los usuarios que poseen la banda.

Caracterización mediante optimización con CVX. Escenario 4.

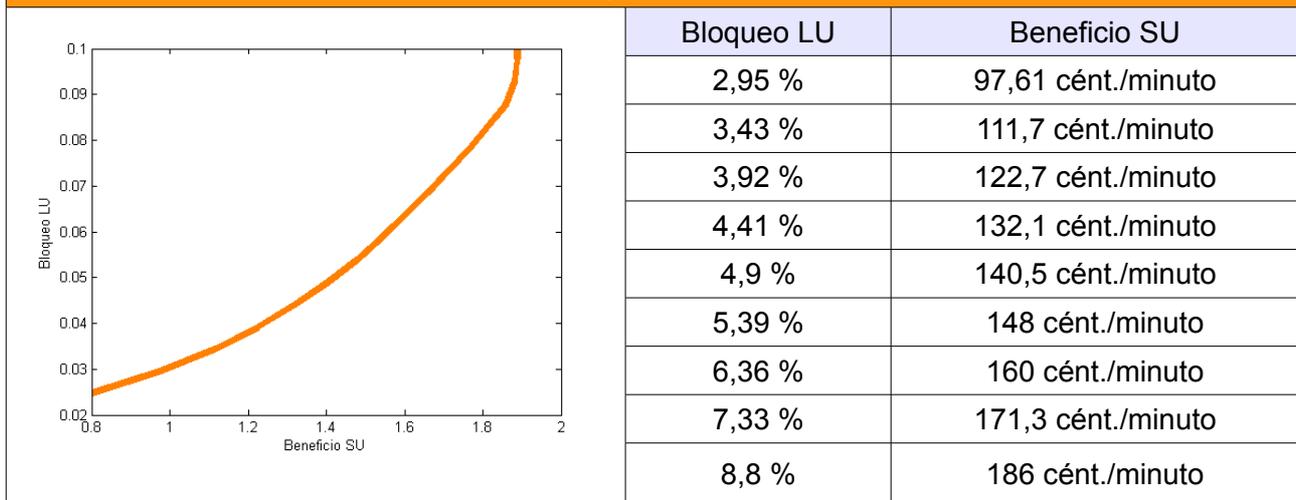


Tabla 22: Resultados de la caracterización mediante optimización con CVX. Escenario 4.

Los resultados de la tabla 22 parecen demasiado optimistas para una situación con tanto tráfico como la de este escenario, ya que el beneficio es muy alto pero el bloqueo se mantiene en niveles muy aceptables. No obstante, al tratarse de un proyecto de investigación, estos resultados no se pueden descartar sin llevar a cabo pruebas con sistemas reales de este tipo.

Caracterización mediante simulación. Escenario 4.

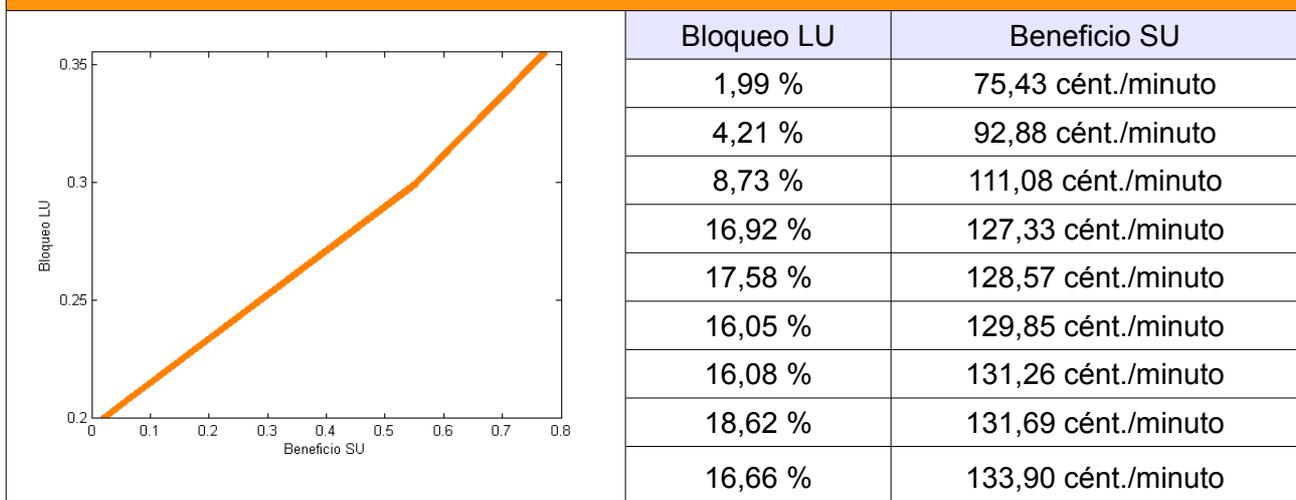


Tabla 23: Resultados de la caracterización mediante simulación. Escenario 4.

Los resultados obtenidos del proceso de simulación son más coherentes con los de la formulación primera, aunque no por ello más válidos que los que se muestran en el proceso de optimización.

En este caso, los valores obtenidos para el bloqueo de usuarios primarios pueden variar hasta en un 2%, como muestran los intervalos de confianza de la tabla 24.

Análisis estadístico de los resultados obtenidos mediante simulación. Escenario 4.							
Bloqueo LU				Beneficio SU			
Intervalo de confianza			Varianza	Intervalo de confianza			Varianza
0,0138 ←	0,0199	→ 0,0259	$2 \cdot 10^{-3}$	0,7184 ←	0,7543	→ 0,7903	0,0067
0,0344 ←	0,0421	→ 0,0498	$3 \cdot 10^{-3}$	0,8878 ←	0,9288	→ 0,9697	0,0087
0,0771 ←	0,0873	→ 0,0974	$5 \cdot 10^{-3}$	1,0795 ←	1,1108	→ 1,1420	0,0051
0,1475 ←	0,1692	→ 0,1910	$2,5 \cdot 10^{-2}$	1,2457 ←	1,2733	→ 1,3010	0,0040
0,1597 ←	0,1758	→ 0,1919	$1,4 \cdot 10^{-2}$	1,2506 ←	1,2857	→ 1,3208	0,0064
0,1460 ←	0,1605	→ 0,1812	$2,2 \cdot 10^{-2}$	1,2632 ←	1,2985	→ 1,3337	0,0065
0,1460 ←	0,1608	→ 0,1755	$1,1 \cdot 10^{-2}$	1,2797 ←	1,3126	→ 1,3455	0,0056
0,1733 ←	0,1862	→ 0,1991	$9 \cdot 10^{-3}$	1,2726 ←	1,3169	→ 1,3611	0,0102
0,1532 ←	0,1666	→ 0,1800	$9 \cdot 10^{-3}$	1,3057 ←	1,3390	→ 1,3722	0,0058

Tabla 24: Análisis estadístico de los resultados obtenidos mediante simulación. Escenario 4.

5. CONCLUSIONES.

Como se comentó en la introducción, y a la vista de la ocupación actual existente en el espectro radioeléctrico, resulta evidente que la cesión estática de espectro que se utiliza hoy en día provoca un uso muy ineficiente de dichos recursos. Parece obvio que la mejor alternativa a este sistema es un nuevo esquema de compartición del espectro que dote a los sistemas de radiocomunicaciones de la flexibilidad necesaria para mejorar el rendimiento de los mismos.

La flexibilidad y el dinamismo de este nuevo esquema pasa por el uso de sistemas de radio definida por software, más concretamente, de radio cognitiva. En este sistema, el transceptor inalámbrico puede cambiar sus parámetros de transmisión de acuerdo a las necesidades del entorno, impuestas por las necesidades de los usuarios con licencia. Los equipos implicados deberán tener la capacidad de observar, aprender, programar y optimizar el acceso al espectro para obtener unos resultados mínimos en cuanto a coste a los usuarios primarios y máximos en cuanto a beneficio para el operador propietario del espacio radioeléctrico.

En este trabajo se han considerado los aspectos económicos del acceso dinámico al espectro en radio cognitiva. Los factores económicos adquieren una especial relevancia en este tipo de sistemas si se tiene en cuenta que el espectro es un recurso muy valioso que requiere un uso muy eficiente para dar respuesta a las necesidades de todos los usuarios implicados.

En resumen, la radio cognitiva es la mejor opción ante la escasez de recursos espectrales, aunque su desarrollo supone un gran desafío en el que se verán implicados factores regulatorios, políticos y económicos. Además, el sistema resultante deberá ser robusto y fiable, y tendrá que ser sometido a todo tipo de procesos de prueba y certificación antes de su total implantación.

Los resultados obtenidos y mostrados en el cuarto apartado del presente documento arrojan **dos conclusiones fundamentales**, que son:

1. **es posible hacer un uso mucho más eficiente del espectro radioeléctrico** causando una interferencia más que aceptable sobre el servicio a los usuarios primarios, y
2. **se puede obtener un beneficio económico** con dicho uso.

BIBLIOGRAFÍA

- [1] R. BRODERSEN, A. WOLISZ, D. CABRIC, S. MUBARAQ MISHRA, D. WILLKOMM. *Corvus: A Cognitive Radio Approach for Usage of Virtual Unlicensed Spectrum*. Universidad de Berkeley y Universidad de Berlín. Julio, 2004.
- [2] F. Ros. *Noticia en Hoy Tecnología*.. URL: <http://www.hoytecnologia.com/noticias/Espana-quiere-eficiente-espectro/163171>. Marzo, 2010.
- [3] S. HAYKIN. *Cognitive radio: brain-empowered wireless communications*. IEEE Communications Magazine, vol. 45, nº 2, pp. 67-73. 2007.
- [4] F. K. JONDRAL. *Software-defined radio prospects for multistandards mobile phones*. IEEE Computer, vol. 40, nº 10, pp. 62-69. 2007.
- [5] E. HOSSAIN, D. NIYATO, Z. HAN. *Dynamic Spectrum Access and Management in Cognitive Radio Networks*. Editorial Cambridge. 2009. ISBN: 978-0-521-89847-8.
- [6] FCC SPECTRUM POLICY TASK FORCE. *Report on the spectrum efficiency group*. FCC Report. Noviembre, 2002.
- [7] FCC. *Notice of Proposed Rule Making and Order*. Diciembre, 2003.
- [8] IEEE. *IEEE 1900 standards committee, IEEE SCC 41*. URL: <http://www.scc41.org>
- [9] JOSEPH MITOLA III, GERALD MAGUIRE JR.. *Cognitive radio: Making software radio more personal*. IEEE Personal Communications, 6. Agosto, 1999.
- [10] JOSEPH MITOLA III. *Cognitive Radio An Integrated Agent Architecture for Software Defined Radio*. Tesis Doctoral. KTH Royal Institute of Technology. Estocolmo, Suecia. 2000.
- [11] G. FETTWEIS, K. IVERSEN, M. BRONZEL, H. SCHUBERT, V. AUE, D. MAEMPEL, J. VOIGT, A. WOLISZ, G. WALF, J. EBERT. *A Closed Solution for an integrated Broadband Mobile System (IBMS)*. En "Proc. of Intl. Conf. on Universal Personal Communications (ICUPC'96)", pp. 707-711. Cambridge, Massachusetts, USA. Octubre, 1996.
- [12] M. BRONZEL, D. HUNOLD, G. FETTWEIS, T. KONSCHAK, T. DOELLE, V. BRANKOVIC, H. ALIKHANI, J. EBERT, A. FESTAG, F. FITZEK, A. WOLISZ. *Integrated Broadband Mobile System (IBMS) featuring Wireless ATM*. En "Proc. of ACTS Mobile Communication Summit 97", pp. 641-646. Aalborg, Dinamarca. Octubre, 1997.
- [13] J. MITOLA III. *Cognitive radio for flexible mobile multimedia communications*. En "Sixth International Workshop on Mobile Multimedia Communications (MoMuC'99)", San Diego, California, USA. 1999.
- [14] F. CAPAR, I. MARTOYO, T. WEISS, F. JONDRAL. *Comparison of bandwidth utilization for controlled and*
-

- uncontrolled channel assignment in a spectrum pooling system*. En "Proceedings of the IEEE 55th Vehicular Technology Conference VTC Spring 2002", páginas 1069-1073, Birmingham, AL, USA. 2002.
- [15] T. WEISS, F. JONDRAL. *Spectrum pooling: An innovative strategy for the enhancement of spectrum efficiency*. IEEE Communications Magazine, 42: S8-S14. Marzo, 2004.
- [16] DARPA XG WORKING GROUP. *The XG architectural framework rfv c1.0*. 2003.
- [17] J. HOU, J. YANG, S. PAPAVALASSIOU. *Integration of pricing with call admission control to meet QoS requirements in cellular networks*. IEEE Transactions on Parallel and Distributed Systems, vol. 13, nº 9, pp. 898-919. 2002.
- [18] D. P. BERTSEKAS. *Dynamic Programming and Optimal Control, vol. 2*. Tercera edición. Editorial Athenea Scientific. 2007.
- [19] L. KLEINROCK. *Packet Switching in a Multiaccess Broadcast Channel: Dynamic Control Procedures*. IEEE Transactions On Communications. Septiembre, 1975.
- [20] H. WANG, N. B. MANDAYAM. *Opportunistic File Transfer Over a Fading Channel Under Energy and Delay Constraints*. IEEE Transactions On Communications, vol. 53, nº 4. Abril, 2005.
- [21] B. WANG, K. J. RAY LIU. *Advance in Cognitive Radio Networks: A Survey*.
- [22] S. ZHANG, F. RICHARD YU, V. LEUNG. *Joint Connection Admission Control and Routing in IEEE 802.16-based Mesh Networks*. IEEE Transactions On Wireless Communications, vol. 9, nº 4. Abril, 2010.
- [23] J. CARLSTROM, R. ROM. *Application-aware Admission Control and Scheduling in Web Servers*. INFOCOM 2002.
- [24] I. PASCHALIDIS, J. TSITSIKLIS. *Congestion-Dependent Pricing of Network Services*. IEEE/ACM Transactions On Networking, vol. 8, nº 2. Abril, 2000.
- [25] J. TSITSIKLIS, B. VAN ROY. *Regression Methods for Pricing Complex American-Style Options*. IEEE Transactions On Neural Networks, vol. 12, nº 4. Julio, 2001.
- [26] H. WANG, Y. GAO, X. CHEN. *RL-DOT: A Reinforcement Learning NPC Team for Playing Domination Games*. IEEE Transactions On Computational Intelligence And AI In Games, vol. 2, nº 1. Marzo, 2010.
- [27] S. FAKIH, T. DAS. *LEAD: A Methodology for Learning Efficient Approaches to Medical Diagnosis*. IEEE Transactions On Information Technology In Biomedicine, vol. 10, nº 2. Abril, 2006.
- [28] M. MOCHOLI ARCE, R. SALA GARRIDO. *Programación lineal. Metodología y problemas*. Editorial Tébar Flores. ISBN: 84-7360-134-3. 272 páginas. 1993.
- [29] M. GRANT, S. BOYD. *CVX User's Guide v. 1.21*. Universidad de Stanford. Abril, 2011.
- [30] J. BANKS. *Discrete-event system simulation*. Editorial: Prentice Hall. 2001.