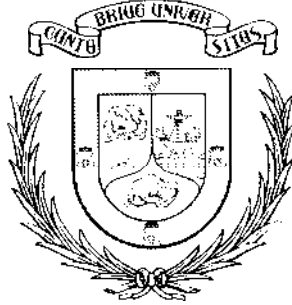# Universidad de Cantabria

Departamento de Ingeniería de Comunicaciones

TESIS DOCTORAL

Análisis y desarrollo de mejoras para la reducción a la exposición a campos electromagnéticos en redes heterogéneas

Reducing the exposure to electro-magnetic fields in heterogeneous networks: design and analysis

Autor: *Luis Francisco Diez Fernández*
Director/Tutor: *Ramón Agüero Calvo*

Santander, Febrero de 2018

Certificado del director de la Tesis

D. Ramón Agüero Calvo, Profesor Titular de la Universidad de Cantabria en el Área de Ingeniería Telemática

HACE CONSTAR

Que la Tesis titulada **"Análisis y desarrollo de mejoras para la reducción a la exposición a campos electromagnéticos en redes heterogéneas"** ha sido realizada por **D. Luis Francisco Diez Fernández** en el Departamento de Ingeniería de Comunicaciones de la Universidad de Cantabria bajo mi dirección, y que reúne las condiciones exigidas a los trabajos de doctorado.

Santander, Febrero de 2018

Fdo. Ramón Agüero Calvo

Declaration

Ramón Agüero Calvo, PhD, Associate Professor at the University of Cantabria in the Area of Telematics Engineering

HEREBY STATES THAT

The Thesis entitled **"Reducing the exposure to electro-magnetic fields in heterogeneous networks: design and analysis"** has been carried out by **Mr. Luis Francisco Diez Fernández** in the Engineering Communications Department of the University of Cantabria, under my supervision, and it meets the requirements for doctoral work.

Santander, February of 2018

Ramón Agüero Calvo

Acknowledgments

Me gustaría dar las gracias a todo el Grupo de Ingeniería Telemática de la Universidad de Cantabria, por acogerme hace unos años y por el apoyo continuo desde entonces.

En especial, quiero agradecer a Ramón Agüero su confianza para entrar a formar parte del grupo, y aceptar dirigir esta Tesis. Desde que fui a tu despacho para preguntarte por algún trabajo de fin de carrera, has sido director (y tutor), mentor, amigo. Así que sólo puedo decirte gracias.

También quiero dar las gracias a todo el laboratorio 21, tanto a los que aún están como a los que no, por compartir el día a día. Gracias Carmen, Pablo (el de *Carblo*, para diferenciar), Juanra, Pablo G. y Nacho. También quiero acordarme de Pablo M. y Javier que, aunque ya hace tiempo que no están, siguen siendo parte del laboratorio. No me puedo olvidar del otro laboratorio, Laura, Chus, Vero, y Jose. Parece que se me olvida, pero no, es que te mereces un agradecimiento en exclusiva: gracias David, por ser el alma del laboratorio durante mucho tiempo, por las charlas a la salida y por todo; se te echa de menos.

También quiero dar las gracias al resto del grupo, especialmente a Johnny y Laura por ayudarme en diferentes partes de la Tesis, y a Luis Muñoz, por hacer posible todo lo que hacemos. Asimismo, quiero dar las gracias a los alumnos con los que he tenido la suerte de trabajar en sus proyectos, y que han ayudado en diferentes aspectos de tesis. Gracias Julián, Juan, Sergio, y las Paulas, espero que os vaya, ahora y en futuro, todo lo bien que os merecéis.

De forma muy especial me gustaría dar las gracias a mi familia, que no ha pasado por los mejores momentos en los últimos tiempos. A mi madre y mi padre, Marisa y Paco, por todo su apoyo siempre y por enseñarme, y demostrarme, que no hay que bajar los brazos, y que todo se puede conseguir con esfuerzo, cada día. Quiero agradecerle a mi hermana, Elena, su apoyo y compresión, y por ser (casi) siempre esa voz de ánimo para el resto. También quiero dar las gracias a quienes ya no están con nosotros, en especial a mi abuela Lupe, por enseñarnos a todos a vivir la vida siempre con una sonrisa.

I would like to also thank the 5G Network R&D group at NEC for the opportunity to stay with them. In particular thank you Xi Li for making it possible and Andrés García-Saavedra for your support during and after my visit.

Por último quiero dar las gracias a todos los amigos que han estado siempre ahí, Javi, Jaime Lore, Aida, etc. A vosotros no os voy a decir nada en concreto, sólo gracias por estar siempre desde hace ya muchos años.

Resumen

Cada pocos años, ante la aparición de nuevos retos y requisitos sociales y técnicos, aparece la necesidad de replantearse las redes de comunicaciones, con la vista puesta en el futuro. Aunque habitualmente no es posible proporcionar una idea precisa, la tendencia actual apunta a escenarios complejos en los que convivirán diferentes tecnologías de manera coordinada.

Hoy en día, las tecnologías de la información y comunicación (TIC), y en especial las comunicaciones inalámbricas, juegan un papel muy relevante en nuestra vida diaria, tanto de forma directa como indirecta. Mientras que en el pasado el uso de las redes de comunicaciones estuvo principalmente dedicado a la interacción entre personas, ahora su ámbito de aplicación se extiende mucho más allá.

Por un lado, la creciente presencia de la tecnología se ha entendido de tal manera que las nuevas generaciones, conocidas ya como nativas digitales, dan por hecho estar siempre conectados. Por otro lado, en los últimos años se han adoptado diferentes soluciones basadas en las TIC, que actúan de catalizador en campos no necesariamente técnicos. En este sentido, agentes públicos están reconvirtiendo servicios tradicionales, ayudados por plataformas digitales, dando lugar a la conocida como "e-administración". Asimismo, bajo el paradigma de la ciudad inteligente, las TIC se emplean para optimizar el uso de servicios urbanos tradicionales, así como para desarrollar otros nuevos. Por otro lado, la industria se esta apoyando en sistemas ciberfísicos para mejorar la eficiencia de los procesos de producción, en lo que ya se conoce como "industria 4.0". Siguiendo esta tendencia, parece razonable pensar que la evolución de las TIC se centrará en sustentar nuevos servicios, tecnológicos o no, más que en la búsqueda de soluciones puramente tecnológicas. Un claro ejemplo puede ser la conducción autónoma, donde las tecnologías de comunicaciones juegan un papel fundamental.

Junto con la adopción de nuevas tecnologías, especialmente diseñadas para responder a las necesidades de servicios concretos, el número de dispositivos con conectividad inalámbrica también está creciendo de manera notable. Por un lado, su variedad se ha incrementado notablemente, pasando del móvil u ordenador a relojes, pulseras y otros tipos de *wearables*. Asimismo, en los últimos años se han popularizado nuevos dispositivos relacionados con el Internet de las cosas o comunicaciones máquina-a-máquina.

Como consecuencia, un amplio conjunto de servicios coexistirán en el mismo escenario, teniendo cada uno sus propias características y requisitos técnicos. Ante la dificultad para definir una única tecnología capaz de satisfacer todo el conjunto de necesidades, parece razonable pensar que varias tecnologías actuarán conjuntamente de manera coordinada. En este sentido, se prevén escenarios donde tecnologías celulares, tales como 2G, 3G, 4G u otras basadas en 5G, se coordinarán tanto con tecnologías de área local (WiFi) como con

otras menos comunes actualmente, especialmente diseñadas para soportar nuevos servicios, tales como LoRa. Esta estrecha integración habilitará un sustrato sobre el que ofrecer, de manera óptima, la gran variedad de servicios TIC que caracterizarán lo que se empieza a llamar *Internet de todo*, o *Internet of everything* (IoE).

Ante este complejo escenario, habrá que hacer frente tanto a retos tradicionales como a otros nuevos. La integración de las múltiples tecnologías, de manera constructiva, supone en sí mismo un reto. Un claro ejemplo es la gran variedad de requisitos de los diferentes servicios. Por un lado, algunos, tales como vídeo o realidad virtual, requerirán una alta capacidad, lo que seguramente precisará el despliegue de un gran número de elementos de acceso que actúen de forma coordinada. Por otro lado, otros no requerirán tanta capacidad, pero precisarán una cobertura ubicua, de forma que nunca se pierda la conectividad. Éste será el caso de algunos servicios críticos relacionados con la Internet de las cosas, tales como la conducción autónoma.

Junto con los retos técnicos, la adopción de las TIC como base que soporte el desarrollo de otros servicios, también requiere tomar en consideración retos de ámbito social. En este sentido, desde hace varios años el ahorro energético ha sido un factor clave en el desarrollo de tecnologías de comunicaciones. Por otro lado, el incremento en el número de elementos de acceso, especialmente estaciones base y puntos de acceso WiFi, ha despertado recientemente cierta preocupación respecto a la exposición a los campos electromagnéticos inducidos por las redes de comunicaciones. Como respuesta a esta preocupación, en varios países de todo el mundo se están llevando a cabo estudios y empezando a aplicar medidas concretas. El ejemplo más claro puede ser Francia, donde se debe indicar los niveles de radiación de los diferentes modelos de teléfonos móviles, y se aplican regulaciones especiales para prohibir redes WiFi en guarderías, e incluso en escuelas de educación primaria. De forma similar, varios países europeos han mostrado su preocupación en este sentido, lo que ha llevado al Consejo de Europa a instar a los gobiernos a tomar todas las medidas razonables para reducir la exposición a campos electromagnéticos, especialmente en los casos que involucren a las personas más jóvenes.

Ante esta situación, resulta necesario considerar la exposición a campos electromagnéticos desde diferentes perspectivas, tanto en sistemas actuales como de cara a las redes de comunicaciones futuras. En primer lugar, se precisa asegurar que las redes inalámbricas no tienen efectos adversos para la salud. En este sentido, el incremento de elementos de acceso para dar soporte a los nuevos servicios estará supeditado a la aceptación de la población. Aunque existen métricas y herramientas capaces de medir la exposición de manera puntual, se precisan nuevos procedimientos que permitan acometer medidas que consideren a la población en su conjunto. Como respuesta a esta necesidad, como parte del proyecto LEXNET, se definió una métrica que caracteriza la exposición del conjunto de la población en un área concreta.

Además de las herramientas de medida, la exposición a campos electromagnéticos (CEM) debe ser tenida en cuenta en el diseño y desarrollo de nuevas soluciones tecnológicas, de forma que la exposición sea integrada como una nueva figura del rendimiento de la red. Entre las diferentes técnicas, aquellas que definen el uso de los recursos radio-eléctricos son, seguramente, las que en mayor medida inciden en la exposición electromagnética. Por ello, en escenarios heterogéneos, tanto la selección de acceso como la gestión de recursos

han de ser cuidadosamente estudiadas.

Por otro lado, también se deben tener en cuenta nuevas topologías de red que pueden ser relevantes para la exposición a los CEM. En concreto, en los últimos años se está proponiendo el uso de redes multi-salto para dar soporte a nuevos servicios de conectividad, tanto en el ámbito del Internet de las cosas, como en el de la comunicación directa entre terminales. Teniendo en cuenta la proximidad de los usuarios con los dispositivos que forman parte de las redes multi-salto, sería beneficioso incluir la exposición electromagnética en las técnicas de encaminamiento.

Desde una perspectiva diferente, también resulta necesario considerar la exposición a los CEM en el diseño de protocolos de transmisión de datos, en especial en aquellos utilizados por los servicios que demandan mayores capacidades de transmisión. En este caso, una mayor eficiencia de estos protocolos se traducirá en una menor necesidad de recursos radio-eléctricos para prestar el servicio y, en definitiva, en una reducción de la exposición a los CEM.

Teniendo en cuenta los aspectos que caracterizarán las redes en el futuro, así como los requisitos que permitan su implementación, esta tesis se articula en base a las siguientes contribuciones:

- El diseño y desarrollo de una herramienta de medida capaz de monitorizar de forma continua la exposición a los CEM sobre áreas extensas. La herramienta está formada por un sistema distribuido de sensores, que se ha integrado en una plataforma de ciudad inteligente, de forma que pueda beneficiarse de las capacidades de gestión proporcionadas por ésta. A fin de optimizar el despliegue de los sensores, se ha caracterizado el área de despliegue de manera exhaustiva y se ha definido un conjunto de fuentes de incertidumbre, consecuencia de diferentes factores (instalación, características de los dispositivos, etc.). Tras estos estudios preliminares se ha procedido a la definición de un protocolo de medida que permite mejorar la exactitud de los las muestras obtenidas, y calcular métricas relevantes. Haciendo uso de esa metodología se ha analizado la exposición a los CEM en en centro de la ciudad de Santander durante un año, prestando especial atención a las variaciones temporal y espacial de los campos inducidos por diferentes tecnologías. Finalmente, se ha desarrollado un portal *web* que permite la visualización en tiempo real de los datos de exposición, así como la descarga de los valores obtenidos por los sensores.

- Diseño de un marco de simulación, *Generic Wireless Network System Modeler* (GWNSyM), capaz de abordar el análisis de sistemas complejos, con un gran número de dispositivos. Basada en los principios de simulación a nivel de sistema, esta herramienta ha sido diseñada para facilitar la agregación de elementos de red, así como cambios en la arquitectura del sistema a simular. Por otro lado, *GWNSyM* se ha diseñado para facilitar la reutilización del código, y de los modelos de simulación, de forma que los desarrollos llevados a cabo en esta herramienta sean fácilmente trasladados a otros entornos.

- Como parte de la validación del entorno *GWNSyM* se ha definido y validado una política de selección de acceso óptima en base a múltiples parámetros. El modelo de

selección de acceso se ha definido como un problema de optimización lineal binaria, que tiene en cuenta tanto preferencias de los operadores como parámetros de calidad de servicio. El diseño de la política de selección de acceso permite adaptarla, en función del escenario concreto que se esté analizando.

- A fin de permitir la evaluación de la exposición a los CEM, de forma precisa, en *GWNSyM*, se han propuesto dos modelos avanzados de recursos radio, tanto para el sentido ascendente como descendente de la comunicación entre terminales y elementos de acceso. A diferencia de otros modelos de simulación usados tradicionalmente, se considera el efecto que la selección de acceso tiene en la eficiencia de los recursos, lo que dota a la simulación de mayor realismo. Los modelos se han definido como problemas de optimización, lineal y geométrica, que han sido adaptados para ser resueltos con herramientas de optimización públicamente disponibles. La comparación de estos modelos con soluciones tradicionales pone de relieve que el hecho de asumir que la eficiencia de los recursos es independiente de la selección de acceso proporciona resultados menos precisos.

- Haciendo uso de los modelos de simulación propuestos, se ha evaluado el potencial beneficio, en términos de exposición, que tiene el descargar tráfico desde las redes celulares a las basadas en *WiFi*. En este sentido, los resultados obtenidos en el marco de esta Tesis, por medio de *GWNSyM*, forman parte de un estudio de mayor envergadura y complementan otros obtenidos mediante herramientas comerciales de planificación de redes. En concreto, el análisis con *GWNSyM* se ha centrado en el estudio de la calidad de servicio, mostrando que, en ciertos escenarios, la descarga de tráfico reduce de forma notable los niveles de exposición, sin deteriorar la calidad de servicio. Sin embargo, dependiendo de las políticas utilizadas por los operadores de red, que pueden impedir que algunos servicios sean movidos a redes *WiFi*, la calidad del servicio percibida por los usuarios puede resultar bastante desigual.

- También se ha analizado el efecto de diferentes técnicas de selección de acceso en redes densas y heterogéneas, compuestas por estaciones base de diferente tamaño, pero compartiendo la misma tecnología de comunicaciones. En este caso se ha prestado especial atención a la comunicación desde el dispositivo de usuario a los elementos de acceso. A la luz de los resultados, se puede concluir que aquellas políticas que realizan la selección de los enlaces ascendente y descendente de manera independiente tienen un mejor comportamiento, tanto en términos de exposición como de la calidad de los servicios. Como contrapartida, estas técnicas causan un mayor número de traspasos, lo que, dependiendo de la arquitectura de red, podría comprometer la calidad de servicio.

- La gestión de recursos en redes heterogéneas también se ha abordado en esta Tesis mediante la definición de un algoritmo de asignación de recursos en escenarios donde los dispositivos pueden hacer uso de varias tecnologías de manera simultánea. El algoritmo propuesto ha dado lugar al registro provisional de una patente europea junto con *NEC Laboratories Europe*, en marzo de 2017. La evaluación del algoritmo propuesto evidencia que es efectivamente capaz de cumplir con los requisitos de

escenarios con conectividad múltiple, siendo uno de los aspectos más destacables el hecho que permite incorporar restricciones prácticas en la asignación de los recursos.

- Se ha definido un marco teórico que permite incluir la exposición a los CEM en el proceso de decisión de encaminamiento en redes multi-salto. En primer lugar se ha propuesto un algoritmo que considera la exposición, tanto instantánea como acumulada, inducida por los nodos que componen la red multi-salto. A continuación, se ha acometido la definición teórica de métricas que permiten evaluar el rendimiento de algoritmos y protocolos de encaminamiento en lo que a exposición se refiere. Finalmente, el desarrollo teórico se ha complementado con la definición e implementación, en entorno de simulación, de un protocolo de encaminamiento, *EMF Aware Routing Protocol* (EARP). Éste ha sido evaluado en base a las métricas mencionadas anteriormente, comparándolo con soluciones tradicionales, en concreto con el protocolo AODV. Se ha puesto de manifiesto que EARP es capaz de modular su comportamiento de acuerdo a eventos pasados. Asimismo, se ha comprobado que el uso de funciones de envejecimiento, para dar mayor peso a las decisiones de encaminamiento recientes, proporciona beneficios en términos de exposición sin comprometer la calidad del servicio.

- Finalmente, se han propuesto dos modificaciones de protocolos que intervienen en la provisión de servicios de vídeo en redes celulares, que consiguen reducir la exposición inducida sin incurrir en pérdidas de calidad. En concreto, las técnicas propuestas buscan optimizar la interacción entre los protocolos de transporte y nivel de enlace, así como de éstos con las propias aplicaciones de vídeo. Los resultados muestran que la selección dinámica de protocolo de transporte (*TCP* o *UDP*) proporciona beneficios tanto en términos de exposición como de calidad de servicio. Adicionalmente, se ha visto que una protección frente a errores adaptativa en el nivel de enlace, basada en la importancia de los datos transportados, puede reducir sustancialmente el número de transmisiones, sin tener efectos adversos en la calidad de los servicios.

El trabajo llevado a cabo durante esta Tesis ha dado lugar a un conjunto de publicaciones relacionadas con las diferentes contribuciones mencionadas anteriormente, recogidas en el Anexo D, donde se indica tanto su tipo como su temática.

Abstract

Information and communication technologies are being widely adopted, as an enabler for new and legacy services, in nearly all realms, both technological and non-technological. This trend is leading to new technical requirements, which are hardly fulfilled by a single technology. In addition, the communication capacity of wireless networks is increasing, to fulfill the requirements of new communication services. In this respect, future wireless network scenarios are believed to embrace multiple technologies and a remarkably higher density of access elements, to cope with the ever increasing capacity demand. Altogether, future scenarios will be characterized by heterogeneous and dense networks, which will be operated in a coordinated manner to optimize the potential communication capacity in an efficient way.

Taking into account the pervasive presence of communication technologies, its deployment and adoption need not only to face technical challenges, but also societal ones. In this sense, along with well known topics such as power saving, in the last years there has been a growing concern regarding the electro-magnetic fields (EMF) exposure induced by wireless communications. In some cases, this concern is hindering the deployment and adoption of new technologies.

In order to ensure the acceptance of new and more dense communications technologies, it seems that EMF exposure should be considered from different angles. On the one hand, measuring tools able to monitor the exposure levels induced by communications technologies should be designed and implemented. Furthermore, new networking solutions and techniques would also need to consider the EMF, including it as another performance indicator.

This dissertation first tackles the design and development of a sensing tool able to continuously monitor the EMF population exposure over large areas. The tools is designed and developed leveraging a smart-city testbed, and it is actually integrated as another urban service, thus easing its adoption. The deployment of the corresponding sensing devices has been carried out considering the scenario characteristics. Furthermore, a number of factors have been identified to improve the accuracy of the values provided by the tool, as well as the integration of extrapolation techniques. Based on such characterization, a complete methodology, providing meaningful EMF population exposure metrics, is proposed, implemented, and validated.

In addition, the dissertation proposes a simulation framework that fosters the systematic assessment of network solutions over large scenarios, paying special attention to the temporal evolution of the scenario. This aspect is particularly relevant to analyze the impact that different solutions may have over the services run by users. In addition, the assessment of EMF exposure requires simulation models able to consider the utilization of

physical radio resources in a realistic way. In this sense, two advanced simulation models are proposed, considering both downlink and uplink communications. Differently to other approaches, the proposed models account for the effect that access selection has over the radio resource efficiency, as a consequence of mutual interference between network elements and terminals. In turn, the simulation models are implemented as optimization problems that are casted so that they can be solved using well known methods. The comparison of these models with legacy approaches shows a relevant difference, which is stronger in dense and heterogeneous scenarios.

Afterwards, and exploiting the aforementioned models, different access selection strategies have been analyzed, in terms of EMF exposure, over large and heterogeneous scenarios. In general, it can be concluded that higher network densities lead to lower exposure, as a consequence of sorter distances between users' devices and network elements. Furthermore, the analysis evinces that tighter cooperation between access elements can also reduce the population exposure.

Besides access selection evaluation, where a user decides the technology to use, this dissertation also tackles the multi-connectivity paradigm, where multiple technologies can be seamlessly used. In this sense, we propose a novel radio resource scheduler able to cover the more relevant requirements of multi-connectivity scenarios. The evaluation carried out yields that the proposed solution ensures system stability regardless of the type of traffic and load regimes. Furthermore, the mathematical framework in which it is based, allows the addition of decision constraints while preserving stability guarantee. This feature is of utter importance, since it enables considering real system constraints, inherent to multi-connectivity scenarios.

Access selection and resource management are likely to be two of the most relevant aspects as far as EMF exposure is concerned, since they rule how the radio resources are used. Nevertheless, other aspects need also to be considered, taking into account the potential importance they will have in envisioned future scenarios. In this regard, multi-hop wireless communications are believed to play a more relevant role either as dedicated connectivity service or as a means to extend the coverage of other systems. Furthermore, it is also important to study how the exposure can be considered from te service perspective. In this sense, this Thesis proposes a novel routing framework to embed the EMF exposure into the design and performance assessment of routing decision procedures for multi-hop networks. In general, the results show that including the EMF exposure as a routing metric yields a fairer exposure distribution without jeopardizing the quality of service.

The dissertation contributions conclude with the proposal of modifications of the legacy protocol stack used for video content delivery service. By using adaptive data protection and transport protocol selection, the proposed solution reduces the number of retransmissions without decreasing the service quality experienced by end users.

Contents

List of Figures

List of Tables

List of Algorithms

Acronyms

Notation	Description
3GPP	3rd Generation Partnership Project
AAA	Authentication, Authorization and Accounting
ADC	Analog to Digital Converter
AF	Antenna Factor
AM	Acknowledge Mode
AMC	Adaptive Modulation and Coding
ANSDF	Access Network Discovery and Selection Function
AODV	Ad-hoc On-Demand Distance Vector
AP	Access Point
aRLC-AM	adaptive RLC-AM
ARQ	Automatic Repeat Request
BLER	Block Error Rate
BLP	Binary Linear Programming
C-RAN	Cloud Radio Access Network
CA	Carrier Aggregation
CBR	Constant Bit Rate
CCA	Cycle Canceling Algorithm
CCDF	Complementary Cumulative Distribution Function
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CEM	campos electromagnéticos
CL	Constant Load
COMP	Cooperative Multi-Point

Notation	Description
CoMP	Coordinated Multipoint
CQI	Channel Quality Indicator
CRE	Cell Range Extension
CT	Content Type
D2D	device-to-device
DASH	Dynamic Adaptive Streaming over HTTP
DCI	Downlink Control Indicator
DUDe	Downlink Uplink Decoupling
E-field	Electric field
EARP	EMF Aware Routing Protocol
EETT	Exclusive Expected Transmission Time
EI	Exposure Index
EMF	Electromagnetic Fields
eNodeB	enhanced NodeB
EPC	Evolved Packet Core
ESS	Experimentation Support Subsystem
ETT	Expected Transmission Time
ETX	Expected Transmission Count
FR	Frame Rate
FTP	File Transfer Protocol
GLPK	GNU Linear Programming Kit
GOP	Group of Pictures
GP	Geometric Programing
GSM	Global System for Mobile Communications
GWNSyM	Generic Wireless Network System Modeler

Notation	Description
HARQ	Hybrid Automatic Repeat Request
HTTP	Hypertext Transfer Protocol
iAWARE	Interference Aware
ICT	Information and Communications Technologies
IoE	Internet of everything
IoT	Internet of Things
IP	Internet Protocol
ISM	Industrial Scientific and Medical
KPI	Key Performance Indicator
LAA	licensed-assisted access
LaSR	Lagrange approximation Supple Radio
LBLAR	Load-Balanced Ad-hoc Routing
LCD	Low Complexity Dosimeter
LENA	LTE-EPC Network Simulator
LEXNET	Low EMF Exposure Networks
LNA	Low Noise Amplifier
LR	Loss Rate
LTE	Long Term Evolution
LTE-U	LTE in unlicensed spectrum
M2M	machine-to-machine
MAC	Medium Access Control
MBR	Mean Burst Length
MCF	Minimum Cost Flow
MCS	Modulation and Coding Scheme
MDP	Markov Decision Process

Notation	Description
MI	Mutual Interference
MIC	Metric for Interference and Channel Switching
mmW	Millimeter Wave
MOS	Mean Opinion Score
MPD	Media Presentation Description
MSS	Management Support Subsystem
NFV	Network Function Virtualization
OFDMA	Orthogonal Frequency-Division Multiple Access
OSI	Open Systems Interconnection
PCB	Printed Circuit Board
PDCP	Packet Data Convergence Protocol
PDF	Probability Density Function
PDU	Protocol Data Unit
PMF	Probability Mass Function
PPP	Poissson Point Process
PRB	Physical Resource Block
QoE	Quality of Experience
QoS	Quality of Service
QSG	Queue Side Greedy
RAN	Radio Access Network
RAT	Radio Access Technology
RBG	resource block group
RF	Radio Frequency
RLC	Radio Link Control
RLC-AM	RLC acknowledged mode

Notation	Description
RLC-UM	RLC unacknowledged mode
RMS	Root Mean Square
RMSE	Root Mean Square Error
RP	Rest Time Period
RSRP	Reference Signal Received Power
RSSI	Received Signal Strength Indicator
SA	Simulated Annealing
SAR	Specific Absorption Rate
SBR	Source Bit Rate
SC-FDMA	Single Carrier Frequency Division Multiple Access
SDN	Software Defined Networks
SDR	Software Defined Radio
SDU	Service Data Unit
SINR	Signal to Interference plus Noise Ratio
SSS	Service Support Subsystem
TCP	Transmission Control Protocol
TDMA	Time-Division Multiple Access
TIC	tecnologías de la información y comunicación
TM	Transparent Mode
TSA	Traffic-Size Aware
TST	Tecnologías, Servicios telemáticos y Sistemas
TTI	Transmission Time Interval
TU	Traffic Unit
UDP	User Datagram Protocol
UE	User Equipment

Notation	Description
UM	Unacknowledged Mode
UMTS	Universal Mobile Telecommunications System
VRB	virtual resource block
WAN	Wide Area Network
WBSAR	Whole Body SAR
WCETT	Weighted Cumulative Expected Transmission Time
WLAN	wireless local area network
WSN	Wireless Sensor Networks
WTP	Working Time Period

Symbols

Symbol	Definition
$1/\tau$	Cut-off frequency of a low pass filter
A	Incidence matrix
AF	Antenna factor
A_e	Equivalent area
BW_{eff}	Effective bandwidth of the LTE physical resource block
BW_{PRB}	Bandwidth of the OFDMA-based physical resource black (e.g. LTE)
b_i	Estimated bandwidth at instant i
C_i	Capacity of access element i
c_0	Speed of the light
$c_{i,j}$	Capacity of the edge from node i to node j
DR	Data rate of WiFi connections
D_i	Demand of service i
d_i	Arbitrary amount of data at instant i
d_i^{DL}	Downlink Exposure Index coefficient associated to configuration i
$d_{i,j}^{UL}$	Uplink Exposure Index coefficient associated to configuration i using device j
E	Electric field
E_i	Electric field component of axis i
$E_r x$	Received electric field
$e[i]$	ith electric field sample
f	Working frequency
fr_i	Fraction of population corresponding to configuration i
$f(x), g(x), h(x)$	General purpose function definition, they are explained where used
G	Antenna gain
$G = (\mathcal{V}, \mathcal{E})$	Graph where $\mathcal{V}$ is the set of vertices and $\mathcal{E}$ is the set of edges
$g(x; \mu, \sigma)$	Standard Cauchy distribution
L	Pathloss
LR	Loss rate
Lf_{low}	Access element free load corresponding to the maximum price
Lf_{high}	Access element free load corresponding to the minimum price
$L(x, \boldsymbol{\psi})$	Lagrange dual function
M_w	Size of the window of the moving average
N, N_i	General purpose symbol for number of elements
$Nei(x)$	Neighboring solutions of x
N_C	Number of cell types in the area
N_E	Number of environments
N_{EC}	Number of Exposure Index components

N_L	Number of user traffic profiles
N_P	Number of population categories
N_R	Radio access technology used by the end-users
N_T	Number of time periods
N_{po}	Number postures
N_u	Number of devices
$Pe(x)$	Perturbation to solution x
P_0	Reference power parameter to control target SINR in uplink LTE power control
P_{PRB_i}	Power allocated to each LTE physical resource block by access element i
Pr_i	Unitary price offered by access element i
Pr_{max}	Maximum allowable price
Pr_{min}	Minimum price
P_{tx}^{ref}	Reference emitted power
P_{rx}	Received power
P_{tx}	Transmitted power
$P_{tx_{max}}$	Maximum transmission power
r_i	Ratio of Exposure Index configuration i
S_{rx}	Received power density
S_{rx}^{ref}	Reference received power density
$SINR_{eff}$	Effective SINR
SAR_i^{DL}	Downlink Specific Absorption Rate value in configuration i
$SAR_{i,j}^{UL}$	Uplink Specific Absorption Rate value in configuration i with device j
St_i	Target SINR of user i for uplink communication
T	General purpose time period
Te	Temperature values in simulated annealing algorithm
$Tr_{i,j}, Tr(i)$	Traffic between nodes i and j and traffic generated or consumed by node i, respectively
T_i^{DL}	Duration of downlink traffic in configuration i
$T_{i,j}^{UL}$	Duration of uplink session in configuration i with device j
U	Utility value in objective functions
U_C	Utility due to connectivity
U_H	Utility due to handover event
U_P	Utility due to price
U_R	Utility due to affinity to radio access technology
Un_i	Uncertainty of extrapolation factor for axis i
W	Matrix of vector solutions
$WBSAR_i^{UL}$	Uplink whole body Specific Absorption Rate for configuration i
x, y	Decision variables
Z_0	Impedance of the vacuum
$\boldsymbol{b}$	Vector of incoming data rate in queue system

$\boldsymbol{Q}$	Vector of queue occupancies
$\boldsymbol{\psi}$	Vector of Lagrange multipliers
$\mathcal{A}$	Set of access elements
$\mathcal{B}$	Set of LTE cells
$\mathcal{E}, \mathcal{U}$	Basis of sets of solutions
$\mathcal{N}$	Set of nodes
$\mathcal{M}$	Set of MCS levels
$\mathcal{S}$	Set of services
$\mathcal{U}$	Set of users
$\mathcal{U}(i)$	Set of users attached to access element i
$\mathcal{W}$	Set of WiFi access points
$\mathcal{X}, \mathcal{Y}$	Sets of possible solutions
$\overline{P}_{tx_i}$	Average transmitted power associated to configuration i
$\overline{S}_{rx_i}$	Average incident power density associated to configuration i
$\overline{T}_i$	Average time in configuration i
α	Uplink path-loss compensation factor
$\beta(i)$	Unary association function that defines the access element to which the user i is attached to
Γ_i	Power strength received by user i from serving access $\beta(i)$
$\gamma_{i,j}$	Power strength received by user i from access element j
Δ_i, ∇_i	Increments and decrements, respectively, of i
δ_{DL}^i	Constant downlink exposure coefficient for configuration i
δ_{UL}^i	Constant uplink exposure coefficient for configuration i
$\delta_{mcs}, f(\Delta)$	Closed loop factors for uplink LTE power control
ϵ	Exposure sample
ζ	Size of the virtual resource block as a number of LTE physical resource block
η	Spectral efficiency
θ	Polar angle in spherical coordinate system
λ	Wave length
μ	Mean value
ν_i	Single-axis extrapolation factor for axis i
ξ_c	Configuration parameter for connectivity utility
ξ_h	Configuration parameter for handover utility
ξ_r	Configuration parameter for RAT affinity utility
ρ	Weight assigned to components of objective function
σ	Standard deviation
ϕ	Azimuth angle in spherical coordinate system
ω	Cost value in objective functions
ω_P	Cost associated to the transmitted power
ω_{acc}	Cost associated to the node accumulated exposure

Software

Tool	Description
TSmarT	IoT Development Platform: TSmarT `http://api.tst-sistemas.es/`
Google Charts	Google Chart API wrapper `https://developers.google.com/chart/`
GWNSyM	Generic Wireless Network System Modeler `https://github.com/ldiez/GWNSyM`
GLPK	Linear Programming Kit `http://www.gnu.org/software/glpk/glpk.html`
GPKIT	GPkit software for geometric programming `https://github.com/hoburg/gpkit`
Ipopt	Interior Point Optimizer `https://projects.coin-or.org/Ipopt`
NS3	Network Simulator 3 `https://www.nsnam.org/`

CHAPTER 1

Introduction

Every few years, with the appearance of new challenges and requirements, both societal and technological, a recurring question arises: how will communication networks look like in the future? Although providing a definitive answer is not possible, the evolution of the last years draws complex scenarios where different technologies will need to coexist in a coordinated manner.

Nowadays, in the so-called digital society, the Information and Communications Technologies (ICT), and particularly wireless communication technologies, play an important role in people's daily life, both directly and indirectly. While in the past the usage of communication networks was mainly focused on the interaction among people, now its applicability scope goes far beyond.

On the one side, the growing presence of the technology for social interactions has spread up to the point that new generations, usually referred to as digital natives, take the fact of being always connected for granted. But it is not only human interactions where ICT play a prevalent role. Technological solutions have been also adopted as a catalyst or enabler in realms that are not necessarily technological. In this regard, public bodies are enforcing the usage of technology, both to hasten administrative procedures (e-administration), and to generate and optimize urban services under the umbrella of the smart-city paradigm. Also, in the private sector cyber-physical systems are being implemented, promoting the so-called industry-4.0. Following the current trend, it seems sensible to think that the evolution of ICT will not only focus on the search for disruptive and purely technological solutions, but on the assimilation of such technology in new user cases, being autonomous driving or proximity services clear examples.

Along with the adoption of new technologies, specially designed to fulfill the requirements of specific service niches, the number of end-devices equipped with wireless technologies is also increasing. This embraces personal devices, such as smartphones, tablets or wearables, and also other wireless equipments related to Internet of Things (IoT) or machine-to-machine (M2M) communications.

As a result, a large set of heterogeneous services will coexist in the same scenario, each having its own goals, features, and technical requirements. This heterogeneity hinders the definition of one communication technology able to cope with all the different requirements, but an environment where different technologies coexist in a integrated manner seems more reasonable. In this sense, it is foreseen that different types of cellular technologies, embracing 2G, 3G, 4G and 5G-based ones, wireless local area network (WLAN) communication systems, and less common technologies, such as Lora, will exist side-by-side

in tightly coordinated scenarios. It is believed that such integrated scenarios will enable the coordination of the different services in the so-called Internet of everything (IoE).

In this scenario, traditional and new challenges need to be tackled. From the technological perspective, the integration of multiple technologies in a constructive way, so that the whole is greater than the sum of the parts, is itself a challenge. One example is the different service capacity requirements. For instance, high-quality video, or virtual reality, will require the deployment of a very large number of coordinated radio access elements to increase system capacity, leading to ultra dense scenarios. Other services, however, will be less capacity demanding, but full spacial and temporal coverage will be an essential requirement. This is the case of some IoT and M2M services, such as autonomous driving.

Together with these technical challenges, the pervasive presence of ICT, as a infrastructure on top of which traditional and new services will be developed, requires considering additional challenges. In this sense, special attention has been paid to power saving in communication networks, refereed to as green networking. More recently, due to the increase of access element density, specially cellular base stations and WiFi hot-spots, the population concern about Electromagnetic Fields (EMF) exposure has grown. Examples can be seen in many countries, where studies are being done and special regulations apply. One of the countries where this concern is most visible is France, where Specific Absorption Rate (SAR) radiation levels must be indicated for all cell phone models and a special regulation is applied to ban WiFi from nursery schools, and to avoid the usage of mobile phones by students in primary school.

Similarly, other European countries have also shown their concern about the use of mobile phones, specially by children. For instance, Belgium approved, in 2013, some rules to avoid the selling and advertising to children, and to add warning labels on mobile phones about their potential negative health effects. In a broader scope, the Council of Europe has also asked the governments to "take all the reasonable measures" to reduce the EMF exposure, particularly in the case of children and young people. But it is not only Europe, where EMF exposure is causing concern. In countries such as India, Israel, Canada, Russia or the United States, different initiatives have appeared in the last years.

1.1 Motivation and objectives

Considering the situation described above, it seems necessary to consider EMF exposure from different angles, both in legacy and future wireless communication systems.

First and foremost, it is crucial to ensure that wireless networks have no damaging effect on human health. In this sense, the future deployment of dense and ultra-dense access networks, able to tackle the requirements of new services, will be conditioned to people acceptance. Although there exist metrics and tools to measure punctual exposure (for instance on a tissue or the whole body), the effect that a communication system may have over the population would require both metrics and tools that consider large areas and time frames. In response to this need, the Low EMF Exposure Networks (LEXNET) project developed a novel metric, Exposure Index (EI), which effectively accounts for the overall population exposure over an area.

Besides measurement tools, the EMF exposure has to be considered as well in the definition and evaluation of technical solutions to reduce it. This way, the exposure could be included as another Key Performance Indicator (KPI) or figure of merit like energy efficiency or Quality of Service (QoS) parameters. Among the different techniques, those that determine radio resources usage are probably the most relevant in terms of exposure. Considering the heterogeneity foreseen in future network scenarios, access selection and resource management techniques are likely to have a very high impact on the overall exposure.

Furthermore, new network topologies need also to be considered. In particular, the appearance of new applications (for instance IoT) and the advanced communication capabilities of user devices, such as device-to-device (D2D), are causing multi-hop topologies to recover their popularity. Hence, it is also convenient to include the EMF in the routing decision process.

As mentioned earlier, capacity requirement will be very different in the near future, where a few services will generate most of the traffic demand (i.e. video), while others will be less capacity demanding. In this sense, considering the EMF exposure in the management procedures of high capacity services would yield a remarkable overall exposure reduction.

Based on the generic goals and characteristics of future networks, the objectives of this dissertation are summarized as follows:

- Design, development and deployment of am EMF exposure monitoring tool, capable of assessing the population exposure according to the metrics defined in the LEXNET project. In particular, this tool should be able to characterize the exposure evolution over large areas in a continuous way.

- Related to the exposure evaluation, but with a broader scope, the second objective is the definition and implementation of a tool that allows the evaluation of future scenarios, able to embed different technologies but, at the same time, to provide certain flexibility to include new paradigms. In particular, we will focus on system level simulation tools to evaluate access selection and resource management solutions.

- Definition of simulation models that accurately mimic the behavior of real systems. In particular, the modeling of radio resources is of utmost importance as far as EMF
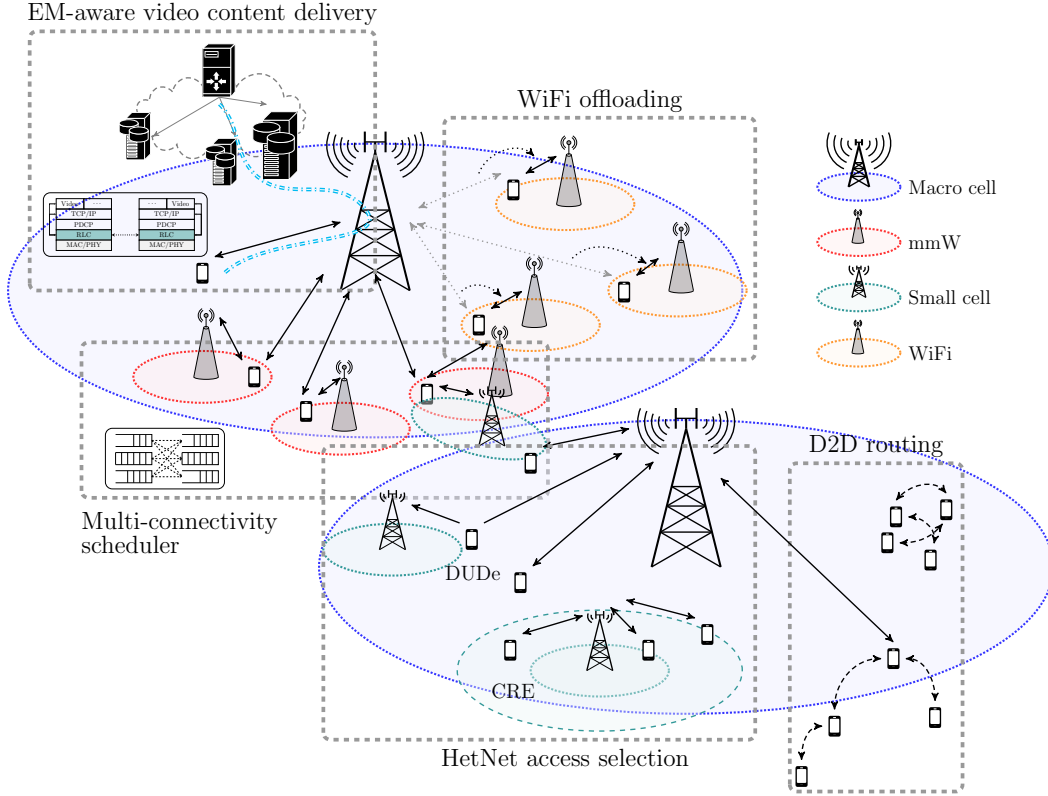
Figure 1.1: Overarching description of the scenario tackled in this Thesis. In some cases the contribution aims at providing alternative solutions, while in other cases the goal is to evaluate already defined solutions for future networks

exposure is concerned.

- Evaluation and proposal of EMF-aware procedures and solutions for future scenarios, as depicted in Figure 1.1. The thesis will cover the following topics:

  - Evaluation of access selection techniques in heterogeneous networks, considering both multiple technologies or a single technology with different capabilities. In the former, we will evaluate traffic offloading to WiFi deployments. The latter case involves studying the impact of access selection techniques over cellular scenarios with access elements of different sizes (resources and coverage).

  - Definition of a scheduler for heterogeneous scenarios where users are able to simultaneously use multiple technologies (multi-connectivity).

  - Design and evaluation of EMF-ware routing mechanisms for multi-hop wireless networks.

  - Proposal and evaluation of EMF-aware techniques for video content delivery

1.2 Structure of the Thesis

The document has been structured according to the aforementioned objectives, as it is depicted below.

- Chapter 2 addresses the first objective. It will be devoted to the design and implementation of an EMF measurement tool able to fulfill the requirements of the LEXNET EI metric. It consists of a distributed sensing system, integrated within a smart-city testbed, so that it can be actually seen as another city service. After describing the smart-city testbed and the sensor device, we will depict the measurement procedure, envisaged to obtain an accurate characterization. Finally, the results provided by the measurement tool over a large period of time will be studied.

- In Chapter 3 we will present the design and evaluation of a simulation framework, which aims at the evaluation of large and complex scenarios. It is a generic framework, based on system-level simulations, specially designed to be highly configurable. In particular, it eases the addition of new network solutions and the integration of architectural changes. We will describe the implementation and design principles of the simulation tool. Finally, the tool will be assessed by means of a heterogeneous scenario, over which abstract models are used for radio resources.

- Chapter 4 will cover the evaluation of access selection and resource management techniques, exploiting the aforementioned simulation tool. First, advanced models for radio resources will be developed and integrated in the simulation tool. Later on, we will evaluate different access selection scenarios, as shown in Figure 1.1, taking advantage of such models. While the evaluation of access selection techniques exploits relatively complex models for the radio resources, simplifications will be assumed for the underlaying scheduling system. In this sense, the resource management analysis will be complemented with the design and evaluation of a scheduler specifically tailored for heterogeneous scenarios. The scheduler is conceived for future scenarios, where user terminals equipped with multiple technologies can use them simultaneously.

- The potential benefits of EMF-aware management techniques, both for routing and video content delivering, will be jointly analyzed in Chapter 5. First, we will propose a theoretical framework to consider the EMF exposure within the routing decision process. Afterwards, a complete routing protocol will be defined and evaluated, according to the theoretical definition. Finally, two different proposals for video content delivery in cellular networks will be discussed. In particular, these solutions are based on modifications of link and transport layer protocols, to improve communication efficiency.

- Finally, Chapter 6 concludes the document, summarizing the main contributions of the Thesis, and outlining some research aspects that remain open for future work.

CHAPTER 2

EMF sensing testbed

Traditionally, the assessment of EMF due to Radio Frequency (RF) communications has focused on individual users, separately considering the exposure induced by personal devices and network elements [Tes14]. However, various studies have demonstrated that there is a high correlation between the power emitted by users' devices and that received from base stations [Gat10]. In order to provide a more accurate way to estimate the population exposure, LEXNET project proposed a novel metric, coined EI, which is associated to a wireless telecommunication network, considering large areas and prolonged periods of time. Furthermore, the EI jointly combines both the exposure induced by personal devices with that coming from network access elements, leading to new KPIs usable during network planning and operation.

In the past, far-field exposure, due to access elements, has been measured by means of sophisticated devices, such as spectrum analyzers, and carrying out individual measurement campaigns, yielding accurate EMF characterization. On the other hand, the evaluation of the EI requires both far-field and near-field, pertaining to users' devices, assessment over large scenarios (for instance a city district) and extended periods of time, which cannot be accomplished by traditional approaches.

In recent years a different approach, fostering the deployment of sensing devices has been proposed, so that the EMF can be continuously monitored over large spatial areas, better suited for the evaluation of the EI. This chapter describes the design, implementation and assessment of a sensing tool that follows this approach. It is based on the deployment of a large number of measuring devices, which are continuously sensing the EMF, on top of a smart-city testbed, in particular the *SmartSantander* [San14] IoT facility.

The sensing tool consists of a set of distributed Low Complexity Dosimeters (LCDs) integrated into *SmartSantander* as regular IoT devices, thus allowing the use of the testbed management capabilities and data processing for EMF sensing. In this sense, it is the first initiative that has merged the EMF sensing and the smart-city ecosystem [Die17a]. The data obtained with this tool will allow the study of the relationship between EMF levels and different parameters, such as time of the day, or downlink-uplink correlation. Besides, the knowledge acquired can be afterwards exploited to fine tune simulation platforms, which can be then used to assess the impact that network management techniques may have over the EMF exposure, for instance small-cells network densification, or the traffic offloading to WiFi hot-spots.

Although the sensing tool discussed herewith is conceptually new, since it leverages smart-city capabilities, there exist initiatives that share some of its features. One of them

is a deployment promoted by the Government of Catalonia, which comprises more than 300 fixed probes, 22 of them placed in the city of Barcelona. The monitoring equipment measures the EMF at different frequencies, and data has been taken since 2013[1]. The main objective of the project was to assess that the measured values respect the legal limits. As a consequence, they have focused on particularly sensitive places, such as kinder-gardens, schools, or hospitals, and the corresponding density is therefore rather low. In addition, the deployed devices (according to the data that can be seen in the available web page) do not distinguish between different bands, providing a single value, which corresponds to a rather wide spectrum range. Hence, it is quite likely to include the exposure induced by other wireless services, such as radio and television broadcasting. As will be seen later, the information provided by the proposed platform differentiates between four bands, and temporary variations, within different time intervals, can be easily obtained.

Another initiative aiming to provide continuous EMF monitoring is the one proposed by the Hermes project [Her17], which provides both wide and narrow band measurements. While this solution might provide a functionality similar to the tool presented in this chapter, it does not exploit a smart-city infrastructure, but rely on dedicated devices (as was the case of the Catalonia network), thus requiring dedicated management of the infrastructure and data. Likewise, moniT [Oli07] and SEMONT [Dju16] projects also proposed solutions to foster a continuous EMF monitoring, exploiting ad-hoc deployments with dedicated devices. It is also worth highlighting the work reported in [Sir12], where authors exploited a wireless mesh network to deploy sensor devices for continuous EMF monitoring. The proposed system was based on a dedicated testbed, specifically deployed for this purpose, and featured advanced measuring devices (Aaronia Spectran Analyzer, HF 6060), more expensive (around 1000€) than the ones deployed in our platform. Hence, the feasibility of this solution to cover wide areas might be questionable due to the monetary cost.

As has been mentioned, these testbeds are dedicated infrastructures, in which devices take measurements and send it to a control center. Hence, it has been necessary to implement communication systems and tailored management functions (e.g. reconfiguration) to operate such infrastructures. Opposed to that, the sensing tool described here exploits the functionality of a generic smart-city platform, where the LCDs are seen as regular IoT sensing devices, so that they can be remotely managed by the control operations offered by the platform, without implementing tailored functions.

Furthermore, using low-complexity, low-cost, devices allows us to afford massive deployments, enabling broader assessment (both in temporal and spatial terms) or higher sensing device density, although some precision is eventually lost. Figure 2.1 provides an overall comparison of different types of commercial EMF sensing devices according to their capabilities and cost. As can be observed, those able to provide frequency selectivity are relatively expensive for a massive deployment. As for the device used in our deployment, it is able to differentiate between the most relevant frequency bands with a rough cost of

---

1   The values and deployment characteristics can be found in `http://governancaradioelectrica.gencat.cat/web/guest/visor`

300€. This combination makes it quite unique, compared to existing solutions, and brings the possibility to use a dosimeter as an IoT device. In fact, using the IoT infrastructure enables embedding continuous EMF monitoring as another smart-city service. In this sense, leveraging existing smart-city deployments may facilitate the adoption of EMF sensing tools by relevant stakeholders (e.g. such as local authorities), since it can be easily integrated as another urban service.

All in all, the smart-city integration offers three main advantages: (1) the devices are cheaper, since they do not need to provide the means to send the information to a remote location, functionality which is taken care of by the smart-city platform. (2) All the functionalities of the platform (data storage, data mining, external access, etc.) can be straightforwardly used to process the information gathered by the dosimeters. (3) Some configuration parameters of the probes, such as the measuring frequency, can be remotely reprogrammed.

This chapter thoroughly describes the EMF sensing tool, and the deployment of the LCDs, as well as all the aspects considered to calculate the LEXNET EI from the measurements gathered by the tool. In particular, the following aspects will be discussed:

- Characterization of the target scenario, to optimize the deployment of dosimeters and to consider its characteristics in the EI computation. In particular, this analysis will provide better understanding of the communication systems operated in the scenario, and will allow us to extrapolate exposure contribution from user devices.

- Calibration of the measurements obtained by the LCD, considering its properties
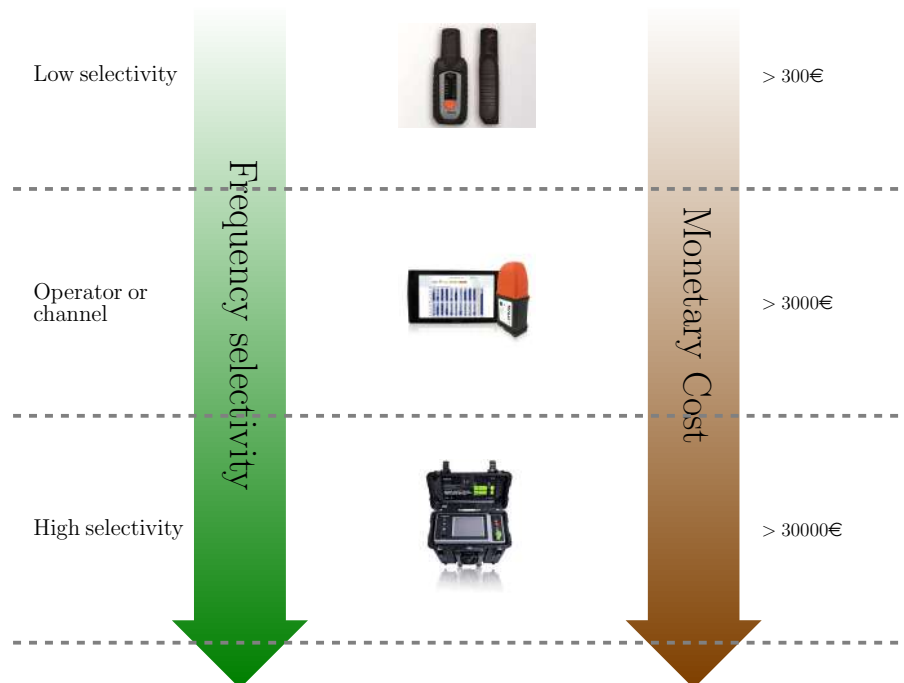
Figure 2.1: Comparison of EMF sensing devices according their capabilities and prices

and the deployment. This will enhance the accuracy of the samples obtained by the sensing tool to better estimate the EI.

- Detailed description of the integration of the LCD into *SmartSantander*, highlighting how the EMF samples are processed and gathered.

- Characterization of the EMF in the target area, and subsequent calculation of the EI.

2.1 Preliminaries

Before depicting the deployment and integration of the EMF sensing tool, this section provides background information that is required to understand the deployment methodology and analysis that will be discussed afterwards. In particular, it introduces the overall architecture of *SmartSantander*, that is relevant for the integration of the LCD. Furthermore, the design of the LCD is also described in this section, highlighting its purpose and limitations. Finally, the LEXNET EI will be formally formulated, indicating its different components.

2.1.1 SmartSantander

The *SmartSantander* [San14] testbed is conceived to satisfy a set of requirements from different stakeholders, embracing from researchers to service providers, including facility providers and local authorities. In the following an overarching description of this testbed will be presented, defining its logical functionalities and architectural concepts, and highlighting how the sensing tool interacts with it.

In order to ease the integration, operation and management of services, the facility provides a set of functionalities that are divided in the following subsystems:

- Authentication, Authorization and Accounting (AAA) subsystem: it ensures that only authorized actions are performed on the platform.

- Management Support Subsystem (MSS): this subsystem provides the functionalities required to add, remove, configure and monitor the resources in the testbed, such us sensors and gateways.

- Experimentation Support Subsystem (ESS): it provides experimentation support over the platform, including reservation of nodes, configuration, running experiments and results gathering.

- Service Support Subsystem (SSS): this subsystem offers functionalities to facilitate the development of services, either for experimentation or final service provisioning.

Regarding the nodes of the facility and communication among them, they are deployed in a 3-tier architecture, according to the following roles:

- IoT node: it is responsible for data acquisition, such as sensing. Most of them are integrated within a repeater, which provides communication capabilities (for instance temperature sensors), while others are deployed in a standalone manner and are able to communicate with the repeaters (e.g. buried parking sensors) in a wireless manner. This is the role that the LCD will take from the *SmartSantander* perspective.

- Repeaters: these nodes provide the first layer of communication and create a capillary network able to gather the information from IoT nodes. The repeaters are deployed above ground (e.g. walls or lampposts) and form a multi-hop network based on 802.15.4 technology. The repeaters are grouped into 25 clusters following a star topology.

- Gateways: these devices are in charge of gathering information from the repeaters and connect the capillary network with the *SmartSantander* backbone, where the information is stored and processed.

Within the *SmartSantander* testbed, devices used for new services are registered by using a logical identifier through the MSS. Once a new device is installed and attached to the system, the registration procedure is automatically performed, by mapping the device to a logical identifier. This mechanism makes use of the AAA subsystem. By this way, the information coming from a particular device can be uniquely identified, along with context metadata. The reader may refer to [San14] for a detailed description of the functionalities provided by *SmartSantander*. From the testbed viewpoint, the EMF sensing tool is seen as another service, and the LCDs are perceived as regular IoT devices, so that no additional setup is required.

The data generated by the dosimeters are encapsulated into proprietary frames following the format shown in Figure 2.2. As can be observed, the frame carries information to identify the node and sensor device providing the data. In this sense, the integration of the LCDs required the definition of a new sensor type, making use of the MSS functionalities. Besides, the subscription to the data gathered by the dosimeters was performed exploiting the functionalities offered by the SSS

2.1.2 Low complexity dosimeter

The low-complexity dosimeter was designed aiming to a large scale deployment. In order to fulfill such goal, the dosimeter design focuses on a low-cost solution while ensuring an appropriate operation and the accuracy of its measurements. In brief, the device is powered and controlled externally, and provides a voltage level as output, which can be converted to the Electric field (E-field) (in V/m) by using the Antenna Factor (AF) and RF chain parameters (gains and losses). The main specifications for the dosimeter are outlined in Table 2.1.

The dosimeter monitors the main frequency bands used for wireless communications,

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15		16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31	
Smart ID		NodeId	
SeqNum	Sensor Type	Frame Type	Length
Payload (3 – 64Bytes)			
...			

Figure 2.2: Format of the *SmartSantander* frame

which are enumerated in Table 2.2. As for the cellular technologies, the device is able
to measure E-field levels corresponding to downlink communications of Global System
for Mobile Communications (GSM), both in 900 and 1800 MHz, and Universal Mobile
Telecommunications System (UMTS) (at 2.1 GHz). The band of 2.4 GHz was included to
estimate the exposure from outdoor WiFi hot-spots together with indoor WiFi sources. It
is worth highlighting that typical bands for Long Term Evolution (LTE) standard were
initially discarded as they were not in use at the target area (Santander city) at the time
of design, February 2014.

Regarding the sensing capabilities, a printed monopole antenna was selected as probe to
reduce the cost of the dosimeter. As a single polarization probe, it can be easily integrated
with the rest of electronics on a Printed Circuit Board (PCB). The main challenge of
this solution was to find an appropriate way to estimate the isotropic exposure from
mono-polarized measurements, as will be discussed in Section 2.3. Besides, the sensitivity
and dynamic range were chosen according to the findings of several studies [Efh14; Neu05;
Vie09], which show that most of the E-field from different telecommunication standards is
below 1 V/m.

All in all, the block diagram of the low-complexity dosimeter is presented in Figure 2.3a.
It consists of a simple PCB with a printed monopole antenna probe at the front to receive
the incoming E-field. The antenna is followed by an RF switch, which is used for calibration
purposes, and a Low Noise Amplifier (LNA) to amplify the incoming signal. After the
amplification, the signal is switched between four RF paths, one per frequency band. The
filters are chosen to efficiently reject the signals from adjacent and other interfering bands.
After the signal is filtered, it passes through the Root Mean Square (RMS) power detector.
Based on the characteristics of the antenna probe and the gain of the RF chain (amplifier
gain and losses in switches and filters), both the received power (in dBm) and the E-field
(V/m) at the antenna can be calculated.

The switches are controlled by digital input signals coming from the multi-point connector
attached with a cable to the repeater. A voltage regulator is also included on the PCB,
to convert the 3.3 volts coming from the battery source to a stable 5 V source, driving
the amplifier and power detector. Then, the output of the power detector is connected

Table 2.1: Summary of the main mechanical and electrical features of the low-complexity
dosimeter

Parameter	Requirement
Dimensions (mm)	$189 \times 80 \times 57$
Dynamic range	60 dB
Sensitivity	5 mV/m
Polarization	Vertical only
Power supply	3.3 V / 300 mA
Power supply type	External
Output type	DC Voltage level

Table 2.2: Frequency bands measured by the low-complexity dosimeter. For cellular technologies, the bands correspond only to downlink communications

Application	3GPP band	Freq. band (MHz)
GSM 900	8	925-960
GSM 1800	3	1805-1880
UMTS	1	2110-2170
WiFi 2.4 GHz	–	2400-2483.5
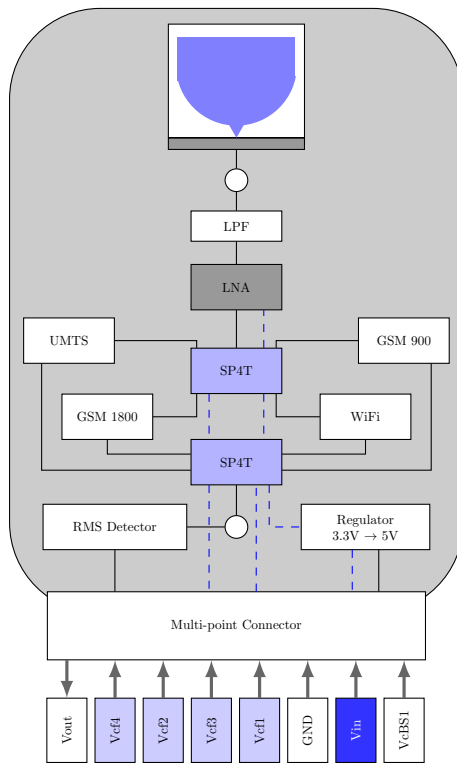
(a) Dosimeter block diagram

(b) Dosimeter prototype

Figure 2.3: Design and implementation of the low-complexity dosimeter

with the same cable back to the IoT controller, where it is converted to E-field using the AF and RF gain. The final prototype is shown in Figure 2.3b, were we can identify the printed monopole antenna, the RF and electronics supply components, and the control and interface part.

Once the prototype was finished, it was characterized using a test setup at Satimo industries[1], by comparing the AF measured with that obtained during the design simulations. The AF is calculated from the real gain (denoted by G) of the antenna: $AF = 9.73/\lambda\sqrt{G}$. The results presented in Figure 2.4 show a reasonable match between the two values. The slight difference is mainly due to the presence of the electrical wire, which was not taken into account during the design simulations.

Next, the RF chain of the dosimeter was characterized. A continuous RF signal was injected through the switch, just before the LNA, as depicted in Figure 2.3. The characteristic curve of the power detector was obtained by measuring the output voltage after the RMS power detector. Figure 2.5 shows the dynamic range of the dosimeter, by representing the mean output voltage value, along with the max/min interval, illustrating the stability of the measurements within the whole range. As can be observed, the dynamic range is above 60 dB, and the requirement defined in Table 2.1 is thus fulfilled. The results were obtained for the 900 MHz band, but this behavior was verified for the other bands as well.

2.1.3 LEXNET Exposure Index

As mentioned earlier, the LEXNET EI was conceived to assess the population global exposure, considering both personal devices and network access elements. The design of this metric had to face potentially contradicting goals. On the one hand, it should be a reasonably simple parameter, so that it could be easily incorporated into network management procedures. On the other hand, it had to combine a large number of information elements of different nature. In order to tackle these objectives, the EI was defined to estimate population exposure during a time interval, in a given area, considering both downlink and uplink communications. Furthermore, since the influence of
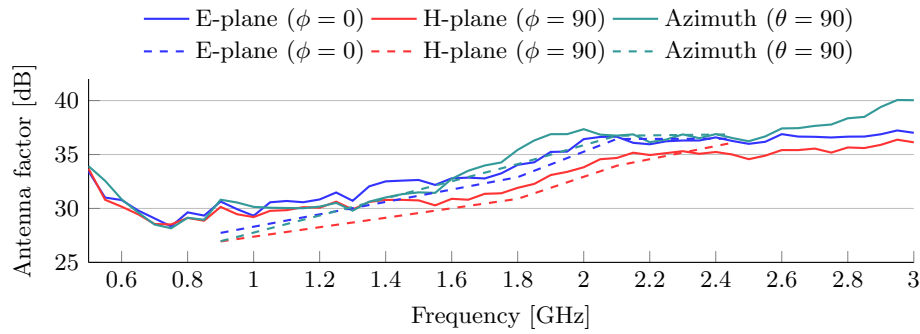
Figure 2.4: Antenna factor as a function of the frequency. Solid lines represent the measured values while dashed ones are obtained by simulation

---

1   Satimo Industries, Brest, France `www.satimo.com`
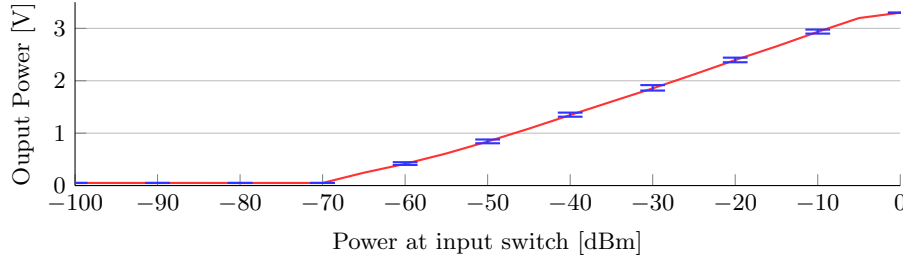
Figure 2.5: Dynamic range of the low-complexity dosimeter for the band of 900 MHz. The output power of the RMS detector is represented as a function of the input power at the corresponding switch. Maximum-minimum interval is also shown

electromagnetic fields strongly depends on other factors (such as age, gender, posture, etc.), an averaging procedure was adopted to derive a value that adequately considers the whole population and the different casuistry. It is worth mentioning that the EI only accounts for the exposure induced by cellular wireless communication systems, as well as WiFi, leaving aside other technologies, such as radio and television broadcasting services, which are less flexible in terms of their management.

More specifically, the EI is based on the dose [Var15a], or SAR, multiplied by the exposure duration, and defined as a summation of different components, which, in turn, correspond to a particular configuration. Each configuration is determined for a population segment under particular circumstances (posture, type of mobile device, etc.) and using a technology. The following parameters have been considered for the definition of exposure configurations:

- N_T is the number of time periods within the reference time frame, which is assumed to be 1 day.

- N_P corresponds to the number of different population categories in the area.

- N_E accounts for the number of environments, such as urban, sub-urban or rural.

- N_R indicates the Radio Access Technologies (RATs) used by the end-users.

- N_C is the number of cell types in the area.

- N_L corresponds to the number of user traffic profiles.

- N_{po} is the number of different postures of the user.

Hence, the total number of configurations, or EI components, N_{EC}, can be calculated as the product of the previously defined parameters: $N_{EC} = N_T \times N_P \times N_E \times N_R \times N_C \times N_L \times N_{po}$. Considering the previous definitions, the overall EI is defined as follows:

$$EI\left[\frac{W}{Kg}\right] = \frac{1}{T}\sum_{k}^{N_{EC}} fr_k \cdot \left[\sum_{u}^{N_u}(d_{k,u}^{UL} \cdot \overline{P}_{tx_k}) + d_k^{DL}\overline{S}_{rx_k}\right] \tag{2.1}$$

where k indicates the current configuration, N_u represents the number of different devices used by the target population. $\overline{P}_{tx}$ and $\overline{S}_{rx}$ are the average transmitted power (in W) and incident power density (W/m^2), respectively. The parameter fr represents the fraction of population corresponding to each configuration. The coefficients d^{UL} and d^{DL} are associated to the exposure induced by either uplink or downlink communications. They are defined as normalized raw dose values multiplied by the time spent in each configuration.

In particular d^{UL} is the absorbed dose, normalized to a reference transmission power P_{tx}^{ref}, defined as follows:

$$d_{k,u}^{UL}\left[\frac{s}{Kg}\right] = \frac{T_{k,u}^{UL}[s] \cdot SAR_{k,u}^{UL}[W/Kg]}{P_{tx}^{ref}[W]} \tag{2.2}$$

where $T_{k,u}^{UL}$ is the time duration of the uplink traffic in a particular configuration k, and device usage u; and $SAR_{k,u}^{UL}$ is the SAR value (either whole body, organ-specific or tissue-specific) calculated for an emitted power of P_{tx}^{ref}. As for the d_k^{DL} coefficient, it is defined as the absorbed dose normalized to a reference power density, as shown in Eq.2.3.

$$d_k^{DL}\left[\frac{sm^2}{Kg}\right] = \frac{T_k^{DL}[s] \cdot SAR_k^{DL}[W/Kg]}{S_{rx}^{ref}[W/m^2]} \tag{2.3}$$

Similar to the definition of uplink coefficient in Eq. 2.2, T_k^{DL} represents the time duration at a particular configuration k, and SAR_k^{DL} is the SAR value induced by the access element corresponding to such configuration for a received power density S_{rx}^{ref}.

As can be observed, the EI combines a relatively large number of cases. Hence, depending on the stakeholder using it, different aspects might be given a stronger relevance. Network operators willing to exploit it for management operations could consider different possibilities. For instance, techniques that are able to reduce the transmission power on a dynamic way would have an impact on the uplink exposure, and just for those periods of time when the user is actually having a ongoing communication. On the other hand, planning strategies are more focused on the downlink exposure, induced by the base stations. Precisely, the platform presented in this chapter mainly focuses on monitoring the downlink contribution of the EI, and the results might be therefore exploited to broaden current network planning procedures. In this sense, authorities (e.g. town councils), which have no control over neither the network or devices, might monitor downlink component of the EI, by measuring the power density, to ensure that the values do not exceed certain limits.

Considering the large number of parameters comprised in the EI, there may be circumstances in which not all of them are available, thus hindering its calculation. In order to provide a common baseline for the EI evaluation, a reference scenario with statistical values of the different involved parameters was defined in [Cor15]. This reference scenario is focused on urban environments and it was defined from the data gathered in [Var15b].

The information related to the population and network users comes from a statistical analysis in Lyon downtown. According to this, the population is segmented in four

categories: *children*, *young people*, *adults* and *seniors*. Besides, each category encompasses outdoor and indoor users [Var15b].

The network data has been estimated from recent measurements in the French Orange network, giving rise to different usage profiles (*light*, *medium* and *heavy*) for the different services (i.e. voice and data) and technologies. It is worth highlighting that it is assumed that neither *VoIP* or *VoLTE* are deployed. Finally, the reference SAR values were obtained assuming people being outdoors are standing, while people being indoors are seated.

This reference scenario will be exploited to calculate the EI from the values provided by the EMF sensing tool in Section 2.3.

2.2 Scenario characterization

Before deploying of the LCDs, a set of measurement campaigns were carried out with a two-fold objective. First, to characterize the target scenario, namely Santander downtown, to optimize the deployment of the LCDs and better exploit the E-field values provided by the dosimeters in different ways:

- Characterization of the E-field distribution over the target area to establish a sensible number of dosimeters, to ensure the accuracy of the results.

- Assessment of the technologies deployed by operators in the target area, since the LCD is not able to distinguish among technologies using the same frequency band.

- Analysis of the transmission power of the users' devices, and its correlation to the E-field measured by the sensing tool.

Furthermore, it is also important to take into account the limitations imposed by the sensing tool. For this, a number of uncertainty sources (i.e. simplified design of the dosimeter or limited number of installation points) have been identified, defining correction factors for each of them.

It is worth noting that, due to logistic limitations, the measurement campaigns were limited in time, and in some cases also in space, so the results presented in this chapter could be broadened with more extensive experiments. Nevertheless, the main contribution of this section is the procedure to complement the EMF measurements and to mitigate inaccuracies of the samples obtained by the sensing tool, yielding a more reliable EI estimation.

2.2.1 Scenario characterization

Deployment dimensioning

In order to obtain meaningful information from the EMF sensing tool, it was necessary to obtain a good approximation of the statistical distribution of the E-field strength within the target area. To do so, a thorough simulation-based study was carried out, exploiting a precise 3D model (with a resolution of 5 m) of the area, by using the Volcano [Vol17] coverage-based simulation platform [Cor09]. This characterization was then used to estimate the dosimeter density.

The simulation strategy consisted in comparing the RMSE of different dosimeter densities. First, the downlink E-field strength was obtained over the whole map (i.e. in every pixel).

Table 2.3: Error statistics from different dosimeter deployments based on simulations. Values are given in dB

	10 Probes	25 Probes	50 Probes
Mean error	1.71	1.4	1.38
RMSE	3.87	2.63	2.52

This was considered as a realistic EMF realization, which was afterwards used as a reference. Then, the dosimeters were emulated by picking sample values in the reference map, at locations selected according to the deployment strategies. Finally, the E-field samples from the emulated dosimeters were compared with the continuous map. This process was carried out in each cellular frequency band (i.e. 900, 1800 and 2100 MHz), considering all base stations in Santander downtown. The reader may refer to [Die15f] for further details of the simulation strategy.

Using this methodology, three densities were evaluated, with 10, 25 and 50 dosimeters deployed overall. In all the cases the locations of dosimeters were uniformly chosen over the target area. Table 2.3 shows the average values of the RMSE as well as the mean error. As can be observed, there is a remarkable gain ($\approx 40\%$) when increasing the number of probes from 10 to 25, but this is much less noticeable when moving from 25 to 50. The main conclusion drawn by this analysis is that a number of probes between 25 and 50 would provide reasonably accurate statistical values of the E-field distribution. Based on this study, it was decided to deploy 40 low-complexity dosimeters, ensuring reliable statistical E-field distribution.

Network usage

Although, at the time of designing the dosimeter, no specific bands had been assigned to deploy 4G, the quick adoption of LTE by the operators made it necessary to analyze the actual usage of the spectrum in the area. The low-complexity dosimeter, although it is able to measure the frequency bands used by the operators, cannot distinguish the contribution to EMF coming from different technologies using the same band.

The characterization of the spectrum usage was carried out with a spectrum analyzer, to identify how the frequency bands were shared between the available operators and how they were using it. In this sense, Figure 2.6 depicts one example of the results obtained from the 1800 MHz band. By analyzing the power values, we could infer that mobile operators were deploying LTE technology in this band.

This underscores the relevance of appropriately characterizing the actual spectrum usage, so that the information provided by the low complexity dosimeters can be further processed according to such usage.

Additional analysis, performed with a Scanner device [Cor15] able to provide Received Signal Strength Indicator (RSSI) values from GSM and LTE in the band of 1800 MHz, yielded that the E-field induced by 2G technologies is around 4.5 times higher than the one caused by LTE. This ratio will be applied to the measurements obtained by the LCD
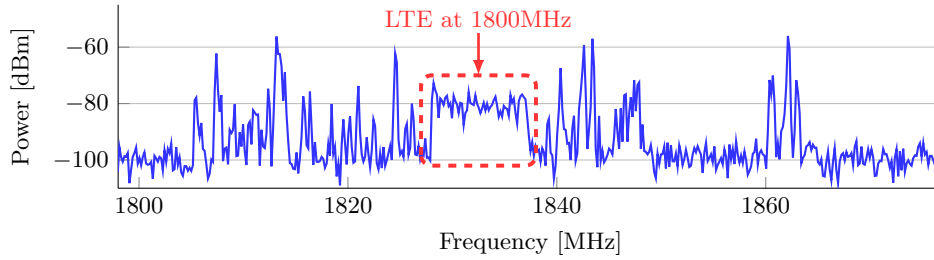
Figure 2.6: Example of a spectrum analyzer capture at the band of 1800 MHz. The figure also highlights the presence of LTE

to better estimate the EI.

Uplink-downlink correlation

As mentioned before, the low-complexity dosimeter was designed to measure far-field values due to downlink communications. On the other hand, one of the most relevant aspects of the EI is that it jointly considers exposure from uplink and downlink communications. In order to overcome the limitations of the measurements obtained by the sensing tool, it is necessary to complement them with additional information, so that the overall EI can be estimated.

Hence, a measurement campaign was performed to compare the received and transmitted power during active service sessions. In particular, we made use of a drive-test device (JDSU RanAdvisor[1]) able to simultaneously monitor transmitted and received powers during voice and data session, using 2G and 3G technologies respectively. In the case of voice communication, the procedure embraced two user terminals, which were having a voice call, while a File Transfer Protocol (FTP) session was used for data communications.

In both cases, the samples obtained by the drive-test were processed to establish a one-to-one relation between received E-field and transmitted power. To do so, the transmitted power samples corresponding to a single received power level were averaged. Then the received power was converted into electric field, assuming far field conditions and free space:

$$P_{rx} = A_e \cdot S_{rx} \tag{2.4}$$

where A_e is the equivalent area of the antenna and S_{rx} is the power density, with units W/m^2. The power density can be defined as a function of the electric field (E), so that $S = \frac{E^2}{Z_0}$, where Z_0 is the characteristic impedance of the vacuum ($Z_0 = 120\pi$). Furthermore the equivalent area is defined in terms of the receiving antenna gain, considering all losses

---

1 https://www.viavisolutions.com

of the receiving system, and the working frequency, as follows:

$$A_e = G \cdot \frac{(c_0/f)^2}{4\pi} \tag{2.5}$$

where f is the working frequency, G holds for the gain, and c_0 is the speed of the light. Altogether, the electric field can be estimated as indicated in Eq.2.6. For simplicity, the correlation of the uplink and downlink will consider unitary receiving gain.

$$E = \frac{\pi f}{c_0}\sqrt{\frac{480 \cdot P_{rx}}{G}} \tag{2.6}$$

Figure 2.7a shows the temporal evolution of both transmitted and received power, during a voice call. As can be seen, there exists a clear trend, so that the higher the transmitted power, the lower the received one. Besides, Figure 2.7b illustrates the mapping between transmitted and received powers. A rather linear trend can be observed, corroborating the duality observed in [Gat10]. Finally, Figure 2.7c depicts the transmitted power as a function of the received E-field. Furthermore, a continuous curve was obtained from those values, applying polynomial fitting, so that we can obtain a value of transmission power for any sample given by the LCD.

Likewise, Figure 2.8 depicts the results obtained with the drive-test for data transfer sessions making use of 3G technology. Similar to the previous case, a linear relation between received and transmitted power can be observed, reinforcing the observations made in [Gat10], but with a different slope. As for the relationship between downlink E-field and transmitted power, a polynomial fitting was also used to obtain a continuous curve that allows mapping for any value provided by the sensing tool.

Although the values obtained during this measurement campaign are limited and could not be applied to other scenarios, this methodology can be helpful to complement the values obtained by the sensing tool, so as to estimate the transmission power of users' devices.

2.2.2 Calibration of the measurements

As has been said before, the dosimeter deployed to build the EMF sensing tool presents a simplified design to reduce its cost, allowing a large scale deployment. As a counterpart, the simplifications, in particular regarding its single-polarized probe, make the samples obtained by the dosimeter less accurate than those that might be obtained with more capable devices.

Furthermore, as the dosimeter are installed in the urban furniture (e.g. lampposts, walls, roofs), the particular surrounding environment may also impact the results. Finally, taking into account deployment constraints, availability of potential locations or permissions, most of the devices had to be installed outdoors, so that indoor exposure is not directly measured.

In order to improve the quality of the EI estimation, a set of correction factors were
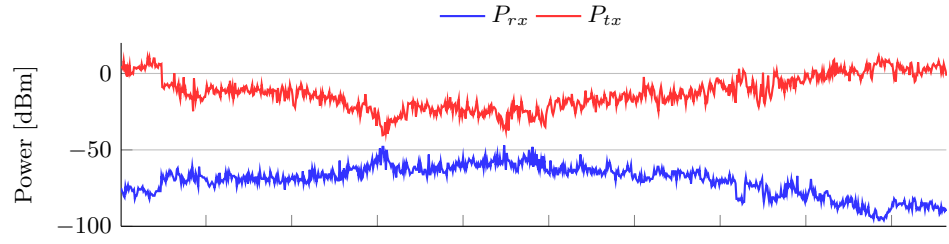
(a) Evolution of transmitted and received powers during a voice call
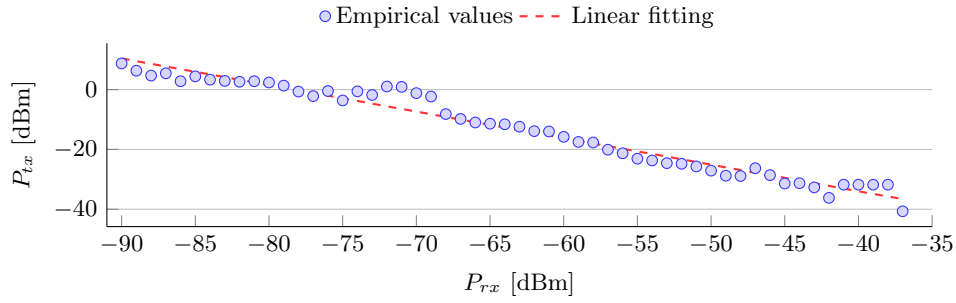
(b) Transmitted power as a function of the received power
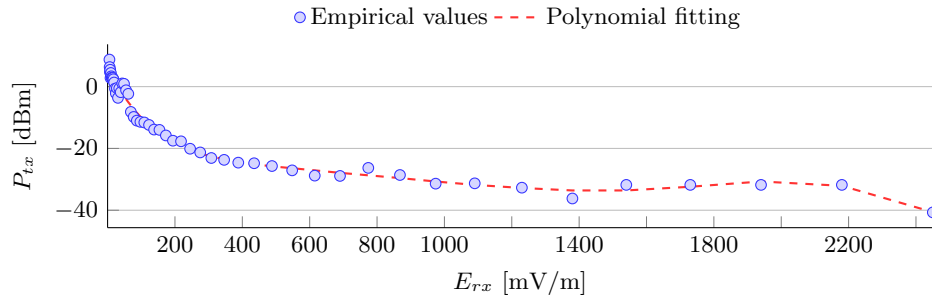
(c) Transmitted power as a function of the received electric field

Figure 2.7: Correlation between transmitted and received power during a voice call using 2G technology

established for the above mentioned cases. The ultimate objective of these factors is to convert the samples obtained by the sensing tool to E-field levels equivalent to those induced to the population. Bearing in mind that the EI is averaged during one day, these parameters were defined as average scale-factors to be applied to the samples obtained by the sensing tool for every frequency band. To this end, samples gathered by the low-complexity dosimeter have been compared with those measured by a more capable device, so that the results provided by the sensing tool can be fine-tuned, considering the dosimeter design and scenario characteristics.
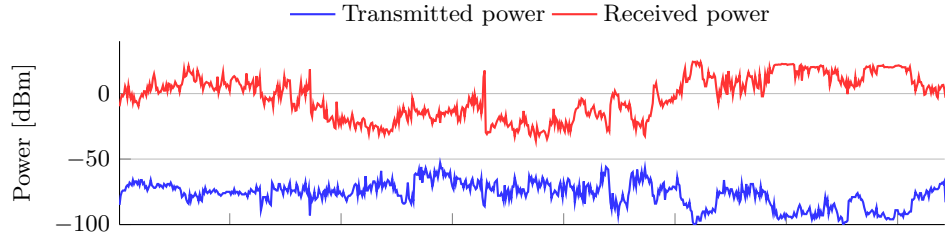
(a) Evolution of transmitted and received powers during data transfer session
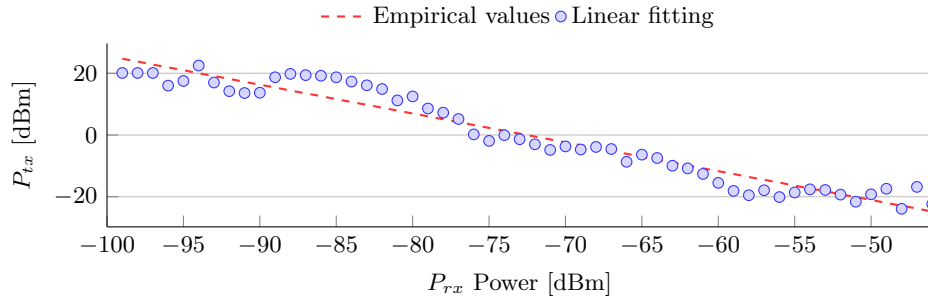
(b) Transmitted power as a function of the received power
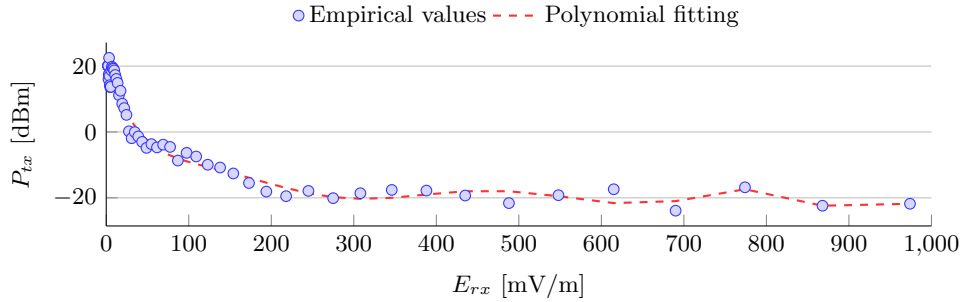
(c) Transmitted power as a function of the received electric field

Figure 2.8: Correlation between transmitted and received power during a data transfer session with 3G technology

In particular, we used the professional wearable dosimeter *EME Spy 200*[1], which is able to provide accurate measurement of the field strength, thanks to the integrated and fully-characterized three-axis antenna, and dedicated in-lab calibration. The measurements were repeated at several time periods in a day, in order to consider the daily variation of the exposure.

In the following, we present the definition of different the correction factors, as well as their values, according to the data gathered during the measurement campaign with the wearable device.

---

1   Satimo Industries, Brest, France `www.satimo.com`

Mono-axial to isotropic

Since the LCD has been designed with a single-polarized probe, it was necessary to define a correction factor, so that the isotropic E-field can be estimated from the samples it provides. In this sense, an exhaustive one week measurement campaign was carried out in Santander, in the month of April, 2015. In particular, the campaign focused on the relationship between each polarization and the overall E-field.

First, a single-axis extrapolation factor ν was defined as follows:

$$\nu_i = E/E_i \qquad\qquad \forall i \in x,y,z \qquad\qquad (2.7)$$

where ν_i represents the extrapolation factor of axis i, E_i is an E-field component (x, y or z axis), and E holds for the total E-field, defined as $E = \sqrt{E_x^2 + E_y^2 + E_z^2}$. It is worth indicating that, following conventional notation, z is assumed to be vertical polarization and $x - y$ represents the horizontal plane.

Then, the uncertainty for the three axis was defined as the ratio between average value and standard deviation:

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N}\nu_{ij} \qquad\qquad \forall i \in x,y,z \qquad\qquad (2.8)$$

$$U_i = \frac{\mu_i}{\sqrt{\frac{1}{N}\sum_{j=1}^{N}(\nu_{ij}-\mu_i)^2}} \qquad\qquad \forall i \in x,y,z \qquad\qquad (2.9)$$

where, μ_i in Eq. 2.8 is the mean value of the extrapolation factors, N is the total number of samples and Un_i holds for the uncertainty in the different axis.

Table 2.4 shows the resulting parameters obtained for the different frequency bands and polarizations. It can be observed that the results at 900 MHz show very large uncertainty values for E_x and E_y (i.e. horizontal components). The reason is the internal

Table 2.4: Mono-axial to isotropic correction factor

Frequency band	E_x			E_z			E_y		
	μ_x	σ_x	$Un_x\%$	μ_y	σ_y	$Un_y\%$	μ_z	σ_z	$Un_z\%$
900 MHz	838.55	1270.97	154.51	1.33	0.74	44.38	1235.81	2059.30	167.44
1800 MHz	2.38	1.26	52.74	1.70	0.66	33.60	2.08	0.83	39.26
2100 MHz	2.36	0.85	35.72	1.39	0.24	12.97	2.45	0.89	35.82
2400 MHz	2.63	1.38	52.44	1.27	0.23	15.18	3.58	1.54	42.35
Average [1]	2.46	1.16	46.97	1.46	0.38	20.58	2.70	1.08	39.14

[1] The band of 900 MHz is not included for the reasons commented in the text.

post-processing of the dosimeter, which is applied to frequencies lower than 1 GHz.

We can also observe that vertical polarization (z axis) yields the lowest uncertainty, 20.58%, which is almost half of the one yielded by the other polarizations. All in all, it can be concluded that the vertical polarization (z-axis) requires the lowest correction factor at all bands. In this sense, the dosimeters will be vertically installed and the z-axis average correction factors (i.e. ν_i) will be applied to the samples so as to obtain average equivalent isotropic measurements.

Location extrapolation factor

Another aspect that has be to taken into account is the impact that the device location might have over the measurements. In this sense, there are several factors that may impact the EMF samples, such as small scale fading, buildings, furniture or the difference between installation height (around 3 m) and the user level (1.5 m). These effects might be better corrected by characterizing each single installation point with a more suitable device. Nevertheless, in cases where the number of installation points is large, this solution is not feasible. Besides, since the EI is averaged over one day, an average correction, while inducing punctual errors, may be appropriate.

The location correction factor has been calculated by comparing the results obtained with the *EME Spy 200* dosimeter with those provided by the LCD, at the same time and at different points of the city; further details of the measurement campaign can be found in [Cor15]. Specifically, the correction factor was defined, for each frequency band, as the ratio of the mean isotropic E-field value from the wearable dosimeter and the mean extrapolated isotropic E-field value from the fixed dosimeter. That is, after applying the mono-axial to isotropic correction factor previously described.

Table 2.5 shows the average values obtained with both mobile and fixed dosimeters. As can be observed, the low-complexity dosimeter underestimates the EMF level, being specially remarkable in the band of 900 MHz. On the other hand, for the 2100 MHz band, the E-field levels provided by the LCD are rather similar to those obtained with the wearable device. As for the standard deviation, we can observe a similar behavior. The band of 900 MHz has the worst behavior while the sparsity of samples in 2100 MHz is rather similar for both devices.

The correction factors obtained from this analysis will be used during the EI calculation

Table 2.5: Average value of E-field with low-complexity and wearable *EME Spy 200* dosimeters for different frequency bands, and correction factor due to installation compared to worn device

Frequency band	EME Spy 200		Low complexity		Corr. factor
	μ	σ	μ	σ	
900 MHz	200.60	218.00	37.50	16.51	5.35
1800 MHz	208.20	247.75	82.62	68.23	2.52
2100 MHz	138.80	139.25	118.68	117.64	1.17
2400 MHz	37.60	25.25	16.81	99.88	2.24

along with those related to the mono-axial probe of the dosimeter. This will assure that the measurement data represents the downlink E-field exposure at the user level in outdoor environment.

Indoor-Outdoor extrapolation factor

Since most of the low-complexity dosimeters are installed outdoors (e.g. urban furniture), it is also necessary to define appropriate extrapolation values to estimate the potential indoor exposure from the dosimeter samples. These values have been obtained comparing the EMF levels gathered with the wearable dosimeter in, and around, two public market places. The measurement itineraries are thoroughly described in [Cor15].

The average and standard deviation of the measured values are summarized in Table 2.6. Based on those results, the correction factor, for each frequency band, has been defined as the ratio of the average E-field value inside and outside the markets.

Table 2.6: Analysis of outdoor-indoor correction factor. The table shows average and standard deviation of indoor and outdoor values, as well as the correction factor

Frequency band	Indoor		Outdoor		Corr. factor
	μ	σ	μ	σ	
900 MHz	25	10	53	18	0.472
1800 MHz	45	20	92	29	0.489
2100 MHz	60	21	105	39	0.571
2400 MHz	10	5	12	5	0.833

As can be seen, for the bands hosting cellular technologies, the EMF level indoors is roughly half than of the one outdoors. It is worth noting that these results are based on a rather limited number of buildings with specific characteristics. More reliable conclusions may be obtained with additional data sets, obtained at different scenarios.

2.3 Deployment and EMF evaluation

This section details the integration of the LCD as an IoT device within *SmartSantander*, thus leading to the EMF sensing tool. Furthermore, based on the results gathered by the tool during a whole year, the E-field over the target area is thoroughly analyzed, which is also used to calculate the EI.

2.3.1 Dosimeter Integration

As mentioned before, the LCDs are deployed as IoT nodes connected to regular repeaters within *SmartSantander* testbed. Before starting the deployment and installation of the new devices the following alternatives were studied:

- Dosimeter inside existing repeater box. This option reuses the existing repeaters. It is likely to be the simplest and cheapest option, considering both installation and deployment costs. However, the main disadvantage is that the dosimeter measurements might be affected by the electromagnetic fields generated by the repeater and power

source. It would be particularly relevant in the 2.4 GHz band, since the repeaters use this frequency for communication purposes. As a possible solution, the controller of the repeater may switch off the radio modules while the dosimeter is measuring, imposing modifications in the current controller. Besides, as was seen before, the dosimeter must be vertically oriented to provide accurate measurements, and the current boxes were not long enough in the vertical axis.

- Dosimeter outside existing repeater box. This alternative would bring about a greater degree of flexibility, specially regarding the orientation of the dosimeter. Besides, this would reduce the influence of the electromagnetic fields generated by the repeater. On the other hand, it would require that the dosimeter is installed within an additional water-proof box, incrementing the total cost of the deployment. Besides, this would also require modifications in the current controller of the repeater to obtain the measurements of the dosimeter.

- New dedicated repeater. This option implies the deployment of new repeaters to control the dosimeters. Since this alternative would not modify the current operation of the existing network, the repeater configuration may be tailored to our needs. For instance, it can be configured as an end-point, ensuring that the radio-module is in idle state until it needs to send the dosimeter information, avoiding any potential interference from the repeater. As a counterpart, this option requires the development of a software controller to gather EMF measurements and to perform communications with the existing infrastructure.

After considering the different alternatives, it was decided to deploy new dedicated repeaters with the dosimeter in a separated box. This decision was made due to the additional costs of modifying the already deployed repeater boxes to embed the new device. Besides, it simplifies the integration with the rest of the platform and potentially avoids additional sources of inaccuracies in the measurements. After considering different options within the LEXNET project, the repeaters TSmoTe provided by the company TST[1] were chosen.

The interface between the dosimeter and the repeater is made of 8 signals, whose layout is described in Table 2.7. These signals are managed by the software controller, located within the dosimeter, according to the following steps:

- Power supply: by controlling the signal V_cBS1 the dosimeter supply (V_{in}) is enabled, and later regulated from 3.3 to 5 V. Some experiments showed that it is sensible waiting, at least, 1 ms after V_cBS1 is enabled, before starting the measurement.

- Frequency band selection: this step consists in controlling the 2 RF switches (i.e. signals $V_{cf\{1,2,3,4\}}$) to select the desired band.

- Sampling: once the frequency band is selected, the sampling of the V_{out} signal is performed. According to the specifications of the dosimeter, after a thorough

---

1 Tecnologías, Servicios telemáticos y Sistemas S.A., Santander, Spain `http://www.tst-sistemas.es/`

Table 2.7: Summary of the functionalities of the low-complexity dosimeter pin-out

Signal	Functionality
V_{out}	Output voltage (0 to 3.3 V) that is sampled to estimate the E-field level measured by the dosimeter
V_{in}	Input voltage (3.3 V) to supply the dosimeter
V_cBS1	Input control to enable or disable the dosimeter supply
V_{cf1}	Switch input to enable measurements at 2400 MHz
V_{cf2}	Switch input to enable measurements at 2100 MHz
V_{cf3}	Switch input to enable measurements at 900 MHz
V_{cf4}	Switch input to enable measurements at 1800 MHz

Table 2.8: Sampling time and processing of the samples obtained from the dosimeter for each frequency band

Frequency band	Sampling time	Processing
900 MHz 1800 MHz	6.616 ms	Median
2100 MHz	6 ms	
2400 MHz	100 ms	Maximum

calibration process, the sampling time was tailored for each frequency band, as detailed in Table 2.8.

- Samples treatment: once the set of values of each band are gathered, these samples are processed to obtain a single voltage value. Depending on the frequency band, either the maximum or median are taken, as specified in Table 2.8.

- Field calculation: finally, the voltage value are converted to E-field level. This conversion is done by using lookup tables loaded in the controller, and that have been created individually for every dosimeter based on the AF and RF chain parameters. The lookup tables convert every possible value provided by the Analog to Digital Converter (ADC) ($[0, 3300]$ mV) to EMF values within the range $[0, 5001]$ mV/m, where 5001 indicates that the measured field is higher than the dosimeter limits (5 V/m).

Before starting the final deployment, the integration and correct operation of the low-complexity dosimeter connected to the repeater was tested, by comparing its results with those obtained with the wearable dosimeter *EME Spy 200*. It is worth noting that, at the time these measurements were performed, correction factors were not yet known. In this sense, the objective of this measurement campaign was to assess the integration and

appropriate operation of the dosimeter, as well as to roughly study the potential of the samples obtained by the LCDs.

Two areas with different characteristics were considered: downtown of Santander city, and the University of Cantabria campus. The former scenario has a relatively large density of mid-high buildings, and validates the measurements that would be taken by the LCD in a real deployment. On the other hand, the later scenario covered a long wide area, and it was intended to assess how a single measurement point can be used to characterize the EMF within a wider area, by means of extrapolation techniques.

The LCD was connected to a PCB repeater, to control the dosimeter and was attached to a metallic mast of 3 m height, which is kept at fixed position, somehow mimicking the future deployment. On the other hand, the wearable *EME Spy 200* device was hand-held at around 1.5 m height, thus getting more accurate measurements of the induced exposure.

For the first scenario, the measurement points were chosen to cover both an open area (wide street) and other locations between mid-high buildings in narrow streets. At each point 100 samples, every 5 seconds each, were gathered with the LCD in a static position. In the case of the wearable dosimeter, the measurements were obtained every 3 seconds with the dosimeter moving around the mast. Figure 2.9 shows the average value of the measurements collected by both devices, low-complexity and wearable dosimeters. It can be observed that the LCD slightly underestimates the E-field as compared with the values offered by the wearable device. The worst case (largest difference) happens at location 3 for the 2100 MHz band, with a negative gap of 479 mV/m (−3.3 dB).

The second scenario covered a wide avenue along University of Cantabria campus. In this case, we compare the measurements provided by the fixed dosimeters with the ones offered by the portable device, while it was moving along the street. The area is approximately 1 Km at full length, and the LCD locations are separated 100 m from each other. Alike in the previous scenario, at each point 100 samples were gathered by the LCD every 5 seconds.

In addition, the wearable device measurements were processed with a moving average

(a) 900 MHz   **(b)** 2100 MHz
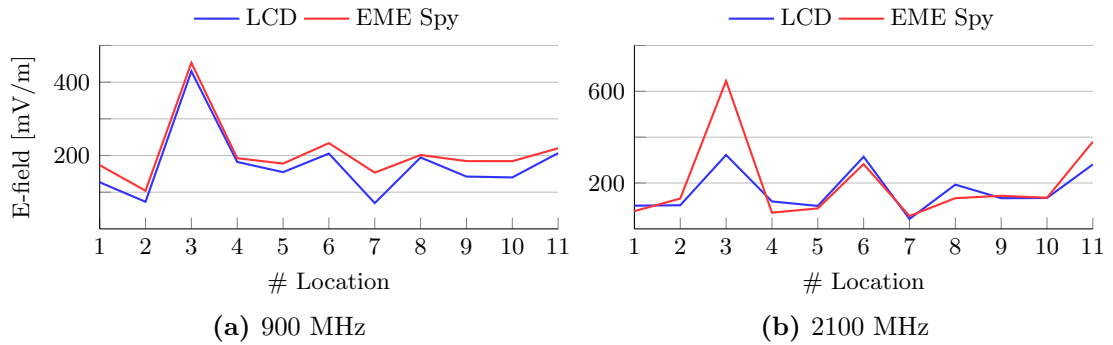
Figure 2.9: Comparison of measurements between the low-complexity dosimeter and the *EME Spy 200* wearable dosimeter. Each point represent the average E-field value at different locations along a path of 1 Km length

(a) 900 MHz
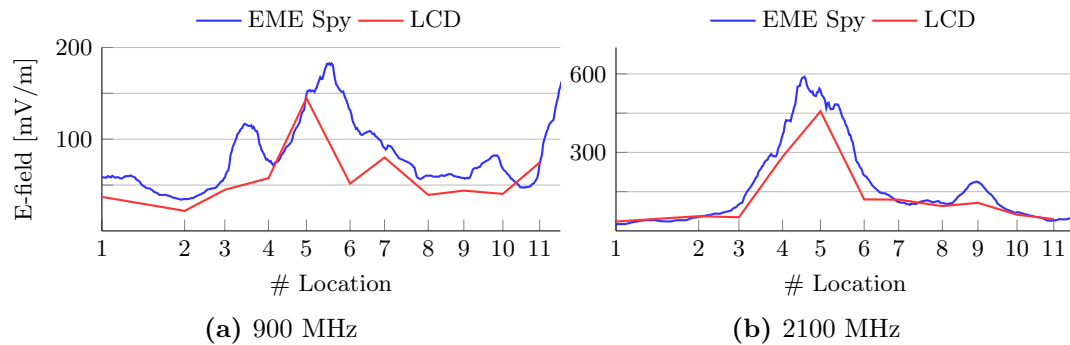
(b) 2100 MHz

Figure 2.10: Comparison of measurements obtained with the low-complexity dosimeter and the *EME Spy 200* wearable dosimeter. Figures 2.10a and 2.10b show the mentioned values for the band of 900 and 2100 MHz respectively

comprising 25 measurement points, to obtain the tendency of the E-field for the different bands. That is to say, the value that we plot for the wearable dosimeter is calculated with the following expression: $E[i] = \frac{1}{M_w} \sum_{j=-\frac{M_w-1}{2}}^{\frac{M_w-1}{2}} e[i+j]$, where $e[k]$ is the E-field measurement provided by the wearable device at the k_{th} point and M_w the number of points used to calculate the moving average. Figure 2.10 shows the results for the bands of 900 and 2100 MHz. As can be seen, the static values provided by the LCD follows the trend of the E-field drawn by the wearable device.

Altogether, this set of results served to validate the integration of the LCD and its operation before carrying out measurement campaigns to obtain the correction factors.

2.3.2 Exposure Index Computation

As was mentioned earlier, the EI definition considers not only electro-magnetic field values, but also a set of data elements that are related to life-segmentation, usage patterns of communication devices, and SAR values. Since these parameters might not be always available, the LEXNET project developed reference scenarios, with values for the different EI components. The sensing tool presented in this chapter is only able to provide E-field values, so that life-segmentation and other ICT usage information has to be given to estimate the EI. The reference scenario for urban environment takes into account all the parameters considered in the EI definition, with the values shown in Table 2.9.

The population is assumed to be segmented in four categories, each of them with four traffic profiles: none, light, moderate and heavy. It is worth noting that users within the none-traffic profile are considered, since they would also be impacted by the E-field exposure (passive receivers). The traffic profiles are different for every population segment and RAT. The reader may refer to [Var15b] for a succinct discussion of these values, and the assumptions that were made to obtain them.

According to the EI definition that was described in Section 2.1.3, the corresponding contribution from each configuration k can be defined as a scaled value of either the mean transmission power ($\overline{P}_{tx_k}$) or the mean received power density ($\overline{S}_{rx_k}$) for uplink

Table 2.9: Reference urban scenario with life segmentation data [Var15b]

Parameter	Values
Time periods	Day, Night
Population segments	Child ($<$15), Young(15-29), Adult(30-59), Senior ($>$59)
Traffic load	None, Light, Moderate, Heavy
Environment	Outdoor, Indoor
RAT Type	2G, 3G, 4G, WiFi
Cell Type	Macro
Services	Voice, Data
Device	Mobile phone, tablet, laptop (dongle)
Posture	Standing, Sitting

and downlink exposure, respectively. In order to simplify the computation of the EI, a simplified version of it was developed, which can be directly obtained by scaling power and E-field values. In this sense, the original representation of the EI can be simplified as shown in Eq. 2.10.

$$EI\left[\frac{W}{Kg}\right] = \frac{1}{T}\sum_{k}^{N_{EC}} fr_k \cdot \left[\sum_{u}^{N_u}(d_{k,u}^{UL} \cdot \overline{P}_{tx_k}) + d_k^{DL}\overline{S}_{rx_k}\right]$$

$$= \sum_{k}^{N_{EC}} \left(\delta_{UL}^k \cdot \overline{P}_{tx} + \cdot\delta_{DL}^k \cdot \overline{S}_{rx}\right) \tag{2.10}$$

$$\text{where} \quad \delta_{DL}^k = \frac{fr_k}{T} \cdot d_k^{DL} \quad \text{and} \quad \delta_{UL}^k = \frac{fr_k}{T} \cdot \sum_{u}^{N_u} d_{k,u}^{UL}$$

$$\tag{2.11}$$

where δ^k is a constant exposure coefficient for a specific configuration k, and is calculated by scaling a reference Whole Body SAR (WBSAR) to account for ICT usage and technology. The value of the WBSAR is typically used as a basic constraint for human exposure to radio frequency [Hir10].

The reference values of the WBSAR were estimated from studies conducted in different European countries, assuming a transmission power and a power density of 1 W and 1 W/m, for uplink and downlink, respectively. Besides, WBSAR reference values were calculated for different services (voice, data, etc.) and for both outdoor and indoor conditions. Finally, the coefficients to calculate the EI were obtained by modulating the WBSAR reference values with ICT usage statistical information.

For instance, we will assume that configuration k accounts for indoor voice calls, within a day using 3G technology. This way, the EI uplink contribution of population under

senior category for a time interval of one day, would be defined as:

$$\delta_{UL}^{k} = \frac{r_{3G} \cdot r_{voice} \cdot r_{senior}}{24 \cdot 3600} \cdot \left[\overline{T}_k \cdot WBSAR_k^{UL} \right] \tag{2.12}$$

where the factor $WBSAR_k$ represents the normalized reference value for this particular configuration, and $\overline{T}_k$ holds for the average time of indoor voice calls during a day. Besides, the exposure is scaled by the ratio of people under this configuration, so that r_{senior} represents the ratio of *senior* population, r_{3G} is the ratio of $3G$ communications and r_{voice} indicates the proportion of voice calls.

The values that have been used to calculate the EI, in the considered urban scenario, are detailed in Table 2.10. As can be observed, from the statistical studies, 4G is not used for voice services and WiFi networks are mostly used indoors. Besides, the values of the coefficients are averaged for some of the configurations, such as device or posture.

Altogether, Figure 2.11 represents the overall methodology to compute the different components of the EI from the data gathered by the sensing tool. In particular, it considers downlink exposure components for both indoor and outdoor scenarios, as well as the uplink outdoor contribution. The uplink component for indoor scenario has not been considered, since it cannot be directly extrapolated from the values gathered by the dosimeters. As can be seen, the methodology makes use of the scenario characterization and correction factors presented in Section 2.2 and it consists of 4 main steps:

1. E-field values are gathered by the sensing tool and averaged according to the reference time interval (one day).

2. An estimation of the E-field induced to users is obtained by applying the corresponding correction factors, which consider both the effect of the installation of the LCDs and the loss of accuracy due to the dosimeter design.

3. Then, the corrected values are used to derive the different components of the EI. As can be seen in Figure 2.11, these values permit us to extrapolate indoor E-field and users' devices transmit power. It is worth highlighting that, for downlink components, E-field has to be converted to power density, according to the EI definition.

4. Finally, the extrapolated values are multiplied by the corresponding coefficients (see Table 2.10) to obtain the actual EI components.

2.3.3 E-field analysis

This section presents an extensive evaluation and characterization of the EMF within the target area (Santander downtown) during a long time period, to study the variation of the E-field, both in temporal and spatial terms. The analysis has been performed from the data gathered during a whole year (2016), focusing on outdoors E-field and leaving aside of 2.4 GHz band, since the obtained values were much lower than those from the cellular technologies, thus hindering its representation. It is worth noting that the results presented here have been obtained by applying the correction factors and network usage extrapolation procedure previously described.

Table 2.10: Coefficients δ of the simplified expression of the Exposure Index for the reference urban scenario [Var15b]

2G				Child	Young	Adult	Senior
UL	Indoor	Day	Voice	0.1427	0.2797	0.2768	0.2039
			Data	0.0303	0.0645	0.0180	0.0019
		Night	Voice	0.0849	0.1662	0.1645	0.1212
			Data	0.033	0.0704	0.0197	0.0021
	Outdoor	Day	Voice	0.0612	0.1199	0.1186	0.0874
			Data	0.0118	0.0242	0.0068	7.1×10^{-4}
		Night	Voice	0.0364	0.0712	0.0705	0.0520
			Data	0.0129	0.0264	0.0074	7.8×10^{-4}
DL	Indoor	Day		243.54	168	166.32	152.88
		Night		405.9	265.44	263.76	263.76
	Outdoor	Day		52.92	31.2	32.76	45.24
		Night		7.56	15.6	17.16	17.16
3G				Child	Young	Adult	Senior
UL	Indoor	Day	Voice	0.0666	0.1970	0.0935	0.0689
			Data	0.79	1.97	0.9	0.4565
		Night	Voice	0.0507	0.1566	0.0712	0.0525
			Data	1.5649	2.8367	1.5048	0.6251
	Outdoor	Day	Voice	0.0260	0.0844	0.0401	0.0295
			Data	0.1672	0.2505	0.0352	0.0054
		Night	Voice	0.0198	0.0671	0.0305	0.0225
			Data	0.1944	0.2913	0.034	0.0058
DL	Indoor	Day		210.87	159	157.41	144.69
		Night		351.45	251.22	249.63	249.63
	Outdoor	Day		48.51	28.2	29.61	40.89
		Night		6.93	14.1	15.51	15.51
4G				Child	Young	Adult	Senior
UL	Indoor	Day	Voice	0	0	0	0
			Data	0.4932	0.9144	0.7315	0.3531
		Night	Voice	0	0	0	0
			Data	1.3438	1.5	1.2194	0.4849
	Outdoor	Day	Voice	0	0	0	0
			Data	0.044	0.1087	0.0548	0.0066
		Night	Voice	0	0	0	0
			Data	0.0458	0.1266	0.0620	0.0061
DL	Indoor	Day		196.02	147	145.53	133.77
		Night		326.7	232.26	230.79	230.79
	Outdoor	Day		45.36	25.2	26.46	36.54
		Night		6.48	12.6	13.86	13.86
WIFI				Child	Young	Adult	Senior
UL	Indoor	Day	Voice	0	0	0	0
			Data	6.8569	16.1083	13.4286	6.4898
		Night	Voice	0	0	0	0
			Data	8.7125	17.2282	14.4512	6.7819
	Outdoor	Day	Voice	0	0	0	0
			Data	0	0	0	0
		Night	Voice	0	0	0	0
			Data	0	0	0	0
DL	Indoor	Day		6.8310	4.2	4.1580	3.8220
		Night		45.36	25.2	26.46	36.54
	Outdoor	Day		11.3850	6.6360	6.5940	6.5940
		Night		6.48	12.6	13.86	13.86
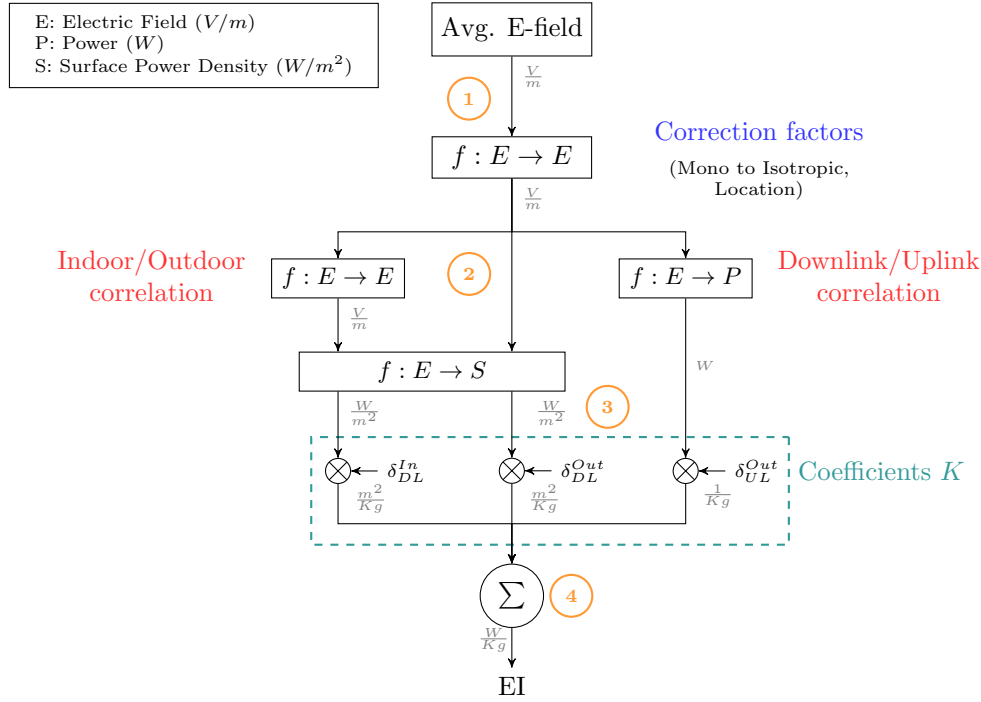
Figure 2.11: Methodology to calculate the Exposure Index By using the IoT sensing tool. The surface power density is calculated assuming impedance in the vacuum: $S = \frac{E^2}{Z_0}$; where $Z_0 = 120\pi$. Besides, $f : A \to B$ represents a mapping function that obtains values of B from values of A

For the first set of results, we select a single dosimeter and we study the temporal variation of the E-field values. This analysis aims to validate the 24 hour usage, as the reference time period for the calculation of the EI. Figure 2.12 illustrates such temporal evolution for three different days of the week (Wednesday, Friday and Sunday).

For a single day, we have data sets of 288 samples, since the E-field is measured every 5 minutes. Then, we take the 52 data sets during the year (one per week), and the average values, as well as the 95% confidence intervals, are plotted. Each of the 9 figures is thus obtained by processing ≈ 15000 samples. We can see that there is a non-negligible variance of the results, although the trend is quite stable.

The E-field values for the 1800 MHz band are the lowest ones, due to the fact that this is the least used band. We can also see that the 3G band (2100 MHz) shows a more remarkable variation of the E-field, which is likely due to the stronger impact of traffic patterns over Code Division Multiple Access (CDMA)-based technologies. Note that this is the only band without GSM technology. Another interesting conclusion can be derived by comparing the evolution of different days. In this sense, the E-field on Saturday nights (left part of the lower figures) is slightly higher than in the other days. This reflects the influence of social activity on the E-field, since the downtown has several restaurants and pubs.
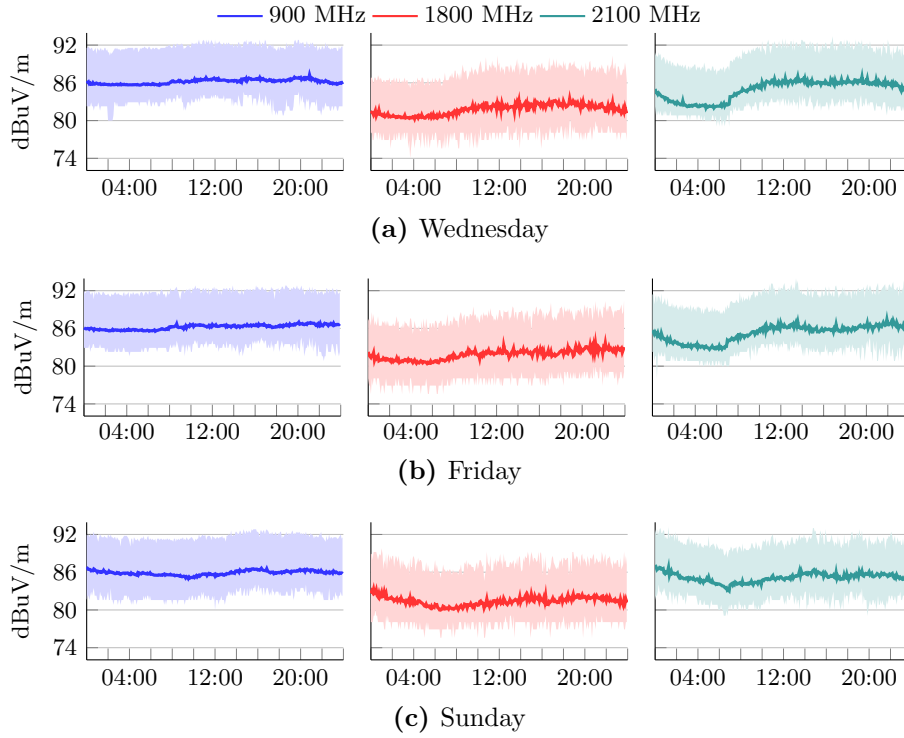
Figure 2.12: E-field variation for different days of the week. Each point is obtained every 5 minutes. Upper and lower values represent the confidence interval

We have afterwards studied the statistical behavior of the E-field during different day periods. In particular, we have defined 5 time intervals, according to activity levels and habits in Spain:

- Working Time Period (WTP): comprises the time periods in the ranges $9:00 - 13:00$ and $16:00 - 19:00$.

- Rest Time Period (RP): is defined to be in the ranges $14:00 - 16:00$ and $19:00 - 22:00$.

- Night: comprises the time interval when communications are less likely, comprising from $23:00$ to $7:00$.

Based on such classification, Figure 2.13 depicts the statistical distribution of the E-field during different day periods for one particular dosimeter, using a box-plot diagram. For each period of the day and frequency band, the boxes indicate the 25 ($Q1$), 50 ($Q2$ or median) and 75 ($Q3$) percentiles. Furthermore, based on $Q3$ and $Q1$, we establish the interquartile range (IQR), as $IQR = Q3 - Q1$, so that the bottom/top lines in the graphs are defined as $Q1 - 1.5 \cdot IQR$ and $Q3 + 1.5 \cdot IQR$, respectively.

Alike the previous results, this metric yields similar variability for all frequency bands. Furthermore, there is not a clear impact of the day time over the E-field values, but for

a slight increase during the evening rest period. Besides, the results also show a more noticeable E-field reduction during night hours.

Once we have analyzed how the E-field evolves during a day, Figure 2.14 shows the distribution of values taken by a single dosimeter, for the seven days of a week. The statistical distribution is represented by using box-plot diagrams, whose values have the same meaning as in Figure 2.13.

We are also considering the measurements taken during a whole year. We can again see that there is not a clear difference between the daily E-field observed. In general, we can observe that the E-field is rather constant for the different days of the week. Besides, a slight upturn of the E-field can be observed on Saturday, with an increase of the median value and a decay around the middle of the week.

All in all, we can conclude that the values at a particular point are rather independent of the time of the day, validating the initial assumption of using a 24 hour interval as the reference period for EI calculation. This way, we ensure not overlooking EMF peaks and valleys during the day, for the different frequency bands. It is worth pointing out that the same behavior was indeed observed in the remaining dosimeters, and the conclusion could be therefore generalized to the whole set of deployment locations.

Once we have seen how the E-field varies at a single measurement point, Figure 2.15 shows the correlation between the values provided by nearby dosimeters. We sort them according to their geographical longitude, from the most Western dosimeter towards the most Eastern one. Then, we plot, for every device, the 10, 50 (median) and 90 percentiles of the samples.

These results yield interesting insights on how the electric field varies within the scenario under study. On the one hand, the measurements taken by a dosimeter do not show a high variability, yielding a tight distribution, where the highest and lowest percentile values are quite close to each other, thus confirming what was previously seen. On the other hand, although there is a clear correlation between some of the nearby devices, there are some cases where the measurements exhibit larger differences. This somehow corroborates the initial assumption that the density of probes needs to be high enough to yield an accurate
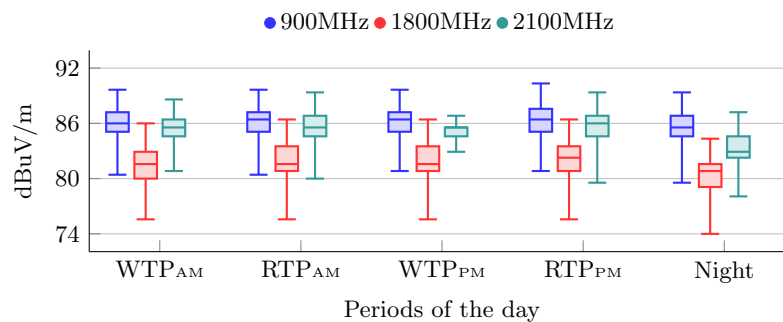
Figure 2.13: Distribution of the E-field for different frequency bands and during different periods of time. The results are grouped for the different times of the day (standard working time and resting periods), and the values for the different frequency bands are indicated in each group

● Tuesday ● Thursday ● Saturday
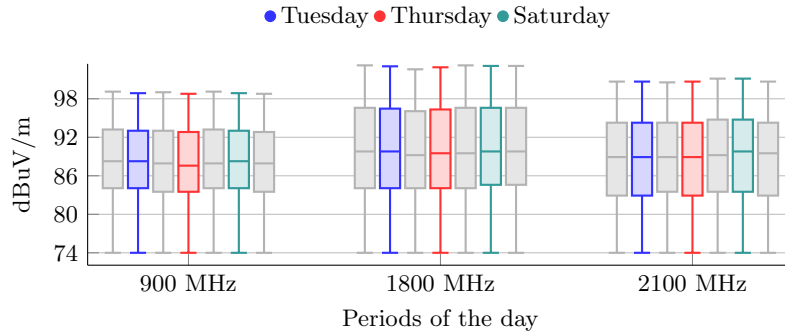
Figure 2.14: Distribution of the E-field for different frequency band and days of the week. The results are grouped for the frequency band, and the days of the week are highlighted in each group

characterization. Both aspects are observed for the three frequency bands. This is indeed confirming the simulation-based analysis presented in Section 2.2.1.

2.3.4 EI evaluation

After analyzing the E-field distribution in the target area, the methodology described in Figure 2.11 has been used to calculate different components of the EI. In this regard, both the correction and correlation factors obtained in Section 2.2 have been used. Besides,
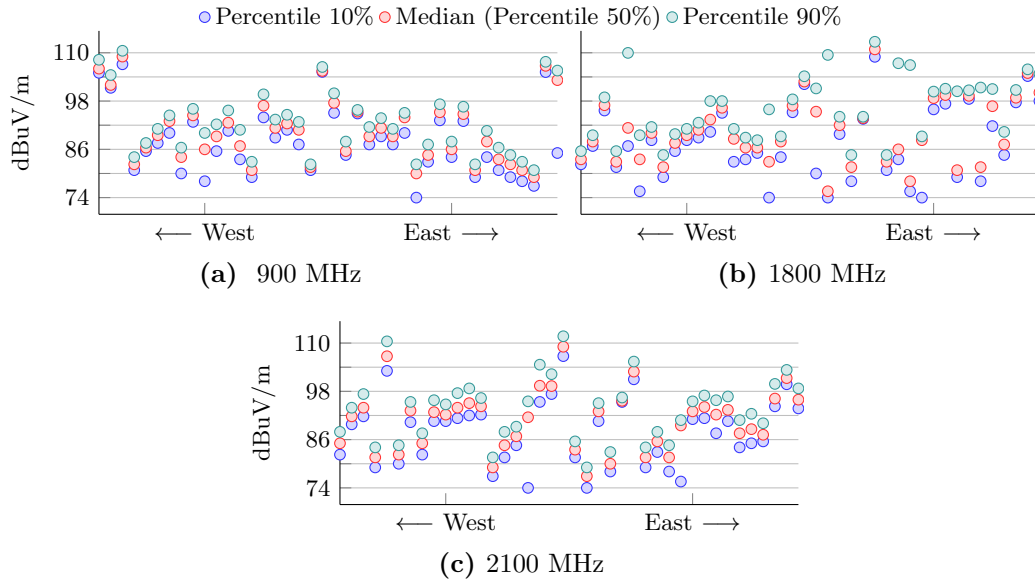
Figure 2.15: Spatial diversity of measurement among the dosimeters for each frequency band. Each point represents a percentile of the values obtained by one dosimeter during the year 2016. X-axis indicates the geographical location of the dosimeter, so that the lower x-value correspond to easter locations
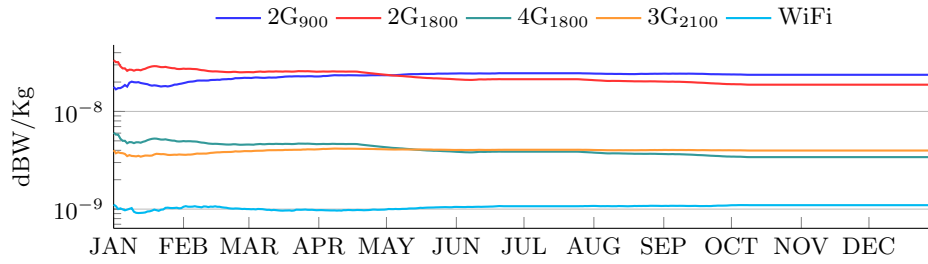
Figure 2.16: Temporal evolution of the downlink outdoor component of the EI in Santander downtown

it is worth recalling that, due to spectrum re-farming techniques exploited by operators, some bands are shared between different technologies. In particular we observed that the 1800 MHz band is used to deploy both 2G and 4G technologies, and based on field measurements [Cor15], it is assumed that the E-field induced by 2G is 4.5 times higher than that caused by 4G networks. Furthermore, the values for the 2.4 GHz band have been included so as to illustrate the different exposure levels coming from WiFi technology.

Figure 2.16 shows the temporal evolution of the downlink EI, considering the contribution of the different technologies. It can be clearly seen that the strongest contribution comes from 2G, which is almost 10 dB higher than the one observed for both 3G and 4G technologies. Besides, we can also observe, for the two 2G bands, a smooth decrease of the EI, being slightly more relevant for the 1800 MHz case. This might be caused by the fact that during the last months of 2016, cellular operators deployed 4G networks using white-spaces left by TV broadcasting services (800 MHz). In this sense, it is sensible to assume that the E-field values can be affected by adjacent bands, due to the sensitivity of the dosimeters. On the other hand, the values observed from the 2100 MHz band, uniquely devoted to 3G, shows a rather constant level during the whole year.

In general, EI variation is rather slow, as could have been expected taking into account that this metric averages values for a whole day. Hence, it would be required studying the evolution of the metric during larger time periods, making the EMF sensing tool appropriate for such goal.

Finally, Table 2.11 illustrates the average values of the different components of the

Table 2.11: Average values of the EI components during 2016 in Santander downtown. Figures are given in dBm/Kg, **n.a.** indicates that either the metric is not defined or there were not measurements to calculate it

Component	2G$_{900MHz}$	2G$_{1800MHz}$	4G$_{1800MHz}$	3G$_{2100MHz}$	WiFi
Downlink$_{Outdoor}$	-46.3429	-46.5330	-53.9501	-54.0074	-59.7898
Downlink$_{Indoor}$	-49.6035	-49.6399	-57.0570	-56.4410	-60.5833
Uplink$_{Outdoor}$	-79.3726	n.a.	n.a.	-59.6966	n.a.

EI during the year 2016. As can be observed, according to the data obtained during the measurement campaign, indoor component of the EI in downlink communications is around 3 dB lower than the one induced outdoors. Regarding the uplink exposure, it has been calculated only for the 900 and 2100 MHz bands, assuming voice and data traffic, respectively. As for 900 MHz, we can observe that the uplink EI component is rather low, compared to the downlink one. This is due to high and quite stable measured values that, after applying the correction factors, are around 200 mV/m. In this sense, high E-field values in the downlink imply good network coverage (proximity), so the transmission power for uplink communications can be reduced.

On the other hand, the E-field values obtained at 2100 MHz band are relatively low, slightly above 60 mV/m after applying the correction factors, which yields quite higher uplink exposure, but still below the one considering downlink communications. It can be due to the effect of the δ factors used to calculate the EI, see Table 2.10, which are much higher for downlink, according to life-segmentation data.

2.3.5 Monitoring and open data

Once the deployment was completed, the sensing information has been made publicly available by means of a web portal[1], to facilitate data acquisition and visualization. Figure 2.17 depicts the main view of such portal, where users can check the deployment of the dosimeters (i.e. their position), including whether they have been installed indoors or outdoors.

Besides the dosimeter location, the web interface also shows the position of base stations within the area, see Figure 2.17a. These access elements can be filtered based on their

(a) LEXNET deployment   (b) E-field variation
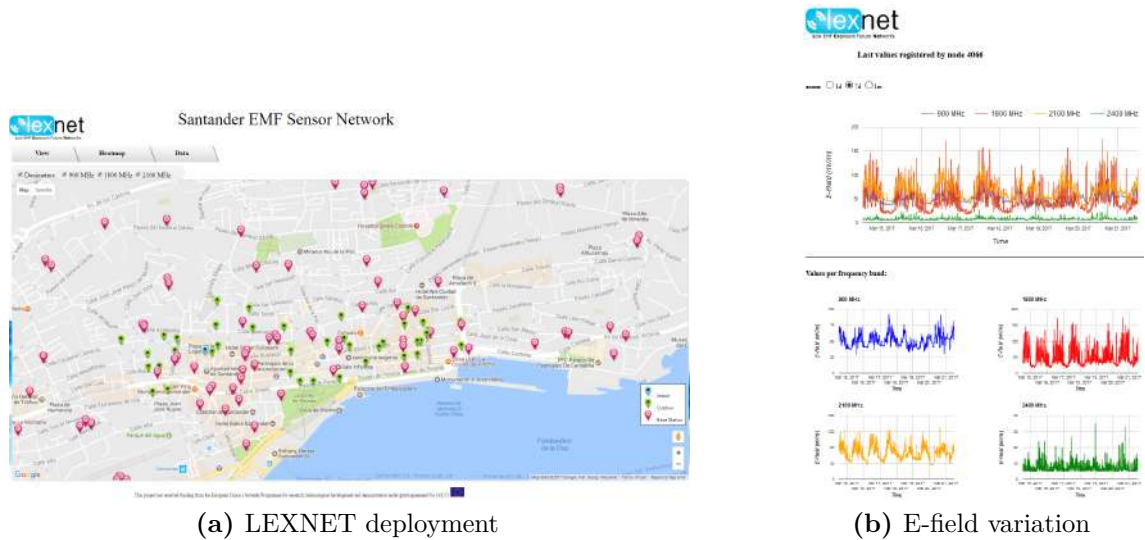
Figure 2.17: Web Portal: deployment map and E-field variation at a single dosimeter

---

1   http://emf-lexnet.services.smartsantander.eu/lexnet/

frequency band, and a link is provided for each of them, offering additional pieces of information, in particular E-field measurements in its vicinity when they were installed. This context information can be afterwards used to better understand the measurements gathered by the dosimeters. Furthermore, quick data visualization can be also obtained from each dosimeter, as shown in Figure 2.17b. This preliminary visualization can be used to observe how the E-field varies, as well as to roughly compare the values provided by different dosimeters. In this sense, it could be also used as a means to assess the validity of the whole measurement methodology, or to identify failing probes.

As previously mentioned, one of the main objectives of the measurement tool is the assessment of the EMF exposure over large areas. While this can be performed by analyzing the data provided by the dosimeters, the web portal also gives an overall representation of exposure throughout the target area, by means of a dynamic heatmap. The heatmap illustrates the evolution of the EMF measurements during the last 24 hours, filtered by the corresponding frequency band or considering the accumulated exposure. While the information provided by the map is clearly not enough for a thorough EMF analysis, it allows to quickly identify potential hot-spots in the city, as well as the temporal evolution of the E-field.

Finally, the web portal also offers the possibility to retrieve data from the different devices, by simply clicking on the,, as illustrated in Figure 2.18. We can use different time-frames (day or week) and filter the information in terms of the frequency bands. In addition, more detailed filters can be applied, by selecting the day periods reflecting the usage patterns in Spain (working hours, breaks, rest, etc.). The information can be also downloaded, as text files, which might be afterwards used to study the E-field samples provided by the LCDs.

Figure 2.18: Dosimeter selection for data retrieval

2.4 Conclusion

There exist a large number of dosimetry devices able to accurately measure the instantaneous EMF levels. However, the growing concern about wireless communications adverse effects makes it necessary to develop new solutions to assess the overall population exposure. In turn, it requires the deployment of sensing devices covering large areas and capable of collecting data during prolonged periods of time, leading to complex sensing facilities.

In the last years, a number of solutions fostering the deployment of a large number of EMF sensors have been presented in different countries. However, they need dedicated infrastructure able to gather and process the data, which may hinder its adoption by authorities.

In order to overcome this potential limitation, an EMF sensing tool leveraging a smart-city testbed has been deployed as a proof of concept. Opposed to other massive sensing solutions, the one presented in this chapter exploits the system and data management functionalities provided by a smart-city facility, so that EMF sensing can be adopted as a new city service.

The sensing testbed has employed low-complexity dosimeters, that were integrated in the *SmartSantander* facility. Along this chapter, the different integration steps have been described, as well as the data treatment applied to the dosimeter samples. In this sense, the main difficulty lies in the conversion from the samples gathered by the dosimeters to representative E-field values. To this end, a number of extrapolation factors have been defined, considering different aspects (installation, single polarized antenna, etc.).

Furthermore, the information gathered during a whole year has been analyzed both to validate the overall integration and to study the EMF evolution. We have seen that there are not remarkable variations on the values measured at a single location. In addition, although the results showed some correlation between nearby dosimeters, we also observed that there might be relevant differences. The high density of probes is precisely addressing this issue. Last, we have exploited additional pieces of information (usage statistics and detailed measurements) to estimate the exposure induced by the different wireless communication technologies. We have seen that the impact of GSM is quite relevant, as compared with 3G and 4G networks.

The deployed testbed is still active, at the time of writing, and can be publicly accessed through a web portal. In this sense, its data is fully available to the scientific community, which might therefore use real E-field values induced by cellular networks in exploitation.

CHAPTER 3

Generic Wireless Network System Modeler

Although 4G technologies are still being deployed worldwide, both industry and scientific community are already working on the evolution of this technology towards the so called 5G. This is believed to seamlessly integrate multiple access technologies ranging from well known cellular ones (GSM, or LTE), Wide Area Network (WAN) such as WiFi, and others devoted to novel communication paradigms, such as IoT and M2M, where technologies like *LoRa*[1] or *SigFox*[2] are rising as promising candidates. Furthermore, an increase of the number of network elements is also foreseen, leading to so called small-cells densification [Bhu14], which might have a positive impact on the population exposure due to the proximity between terminals and access elements. In order to efficiently utilize the network resources, together with the small-cells densification, novel cooperation techniques, such as Coordinated Multipoint (CoMP) or Downlink Uplink Decoupling (DUDe) [Boc14] are believed to be used.

On top of access network modifications, there is another key component of the forthcoming wireless communications landscape, which is the utilization of virtualization techniques and Software Defined Networks (SDN) [Agi16], leading to the Network Function Virtualization (NFV) paradigm. This new approach will allow centralizing network functions when required, thus enabling cooperation techniques among different access elements.

Besides, a remarkable increase of the data traffic is also expected in the next years [Cis17]. One of the main reasons of this growth is the consolidation of new services such as video-streaming, gaming applications, etc., which demand a high capacity.

Hence, it becomes necessary studying the tool that allow an accurate analysis at different abstraction levels. In particular, we have established the following requirements:

- Deployment of large networks, with a high number of access elements and users.

- Evaluation of the impact that new techniques have over the QoS.

- Possibility to consider architectural changes. Although it will eventually require modifications, it would be appealing that new architectural solutions (e.g. NFV) can be modeled and considered in the analysis.

- Ability to assess the performance of new techniques, specially focused on the access part, that may impact the modeling of network entities.

---

1   https://www.lora-alliance.org/
2   https://www.sigfox.com/en

On the one hand, analysis based on real testbeds probably provides the most accurate results. In this sense, the open source implementation of radio interfaces and the appearance of affordable Software Defined Radio (SDR) devices allows the analysis of several technologies that, in the past, required the use of expensive vendor specific devices, often with limited access to its configuration. However, testbeds are usually limited to small deployments and the integration of new solutions is complex, if not impossible.

Simulation offers cheap solutions to assess the performance of complex systems in lab environments, at the expense of some accuracy. According to the precision goal and the way the network is modeled, we can differentiate three main possibilities: network simulators, link-level simulators, and system-level simulators.

The first type simulates the network elements and provides an implementation of the communications protocols. This way, network simulation is able to mimic the transmission of all bits between network entities, at the expense of high computational cost. Among this group, the choice that has gathered more attention is the *ns3* platform [Ns317], and its LTE-EPC Network Simulator (LENA) extension [Pir11], which is in constant evolution. Furthermore, due to their complex implementation, network simulation may interact with real devices[1]. However, when working with large systems, the complexity of network simulators may lead to prohibitive simulation time. For this reason, simulation of cellular networks is usually performed by means of link-level or system level alternatives.

Simulation at link level focuses on the behavior of the radio link between a single transmitter and receiver, considering aspects of both physical and Medium Access Control (MAC) layers, and radio propagation phenomena at small temporal scales, for instance at the symbol duration. In the physical layer, link-level simulators tackle functions such as modulation and interleaving. On the MAC layer, it considers procedures such as the Hybrid Automatic Repeat Request (HARQ), which can significantly affect the system performance. This type of simulation has been extensively studied in cellular systems, and accurate models and guidance to build them are even provided by the standards. Vienna LTE simulator [Meh11] appears as one of the most widely accepted solution.

The third group, system-level simulators, studies the performance of the network considering aspects such as handovers or radio resource management procedures. In this case, the objective is not to analyze the performance of a single radio link during small temporal periods, but to evaluate the performance and behavior of the entire network over longer periods of time. Link-level and system level simulators are usually independently executed due to the different time scales and computational costs. However, in order to achieve an accurate and reliable network evaluation, they may still interact through static interfaces such as lookup tables or analytical models. These interfaces are usually generated from link-level simulators, and reflect link performance results in a simplified manner. For example, radio link quality in terms of Signal to Interference plus Noise Ratio (SINR) can be mapped onto a radio link performance indicator, such as Block Error Rate (BLER). The link level interfaces are used by the system level simulators to assess radio link performance, resulting in a simple yet efficient approach. It allows including the effects of both MAC

---

1 https://www.nsnam.org/wiki/HOWTO_make_ns-3_interact_with_the_real_world

and physical layers, as well as the fluctuations of the radio channel, which occur in a small time scale, into system-level simulations, which consider longer time scales.

Concerning the available solutions for system-level simulation, most are based on proprietary developments (in this case the use of Matlab is quite popular), while others use some of the few tools that have been made available, outstanding again the Vienna LTE Simulator [Meh11]. As a consequence of using proprietary and ad-hoc solutions, the results obtained are not easily reproducible and it is usually complex to integrate them into other frameworks.

Table 3.1 provides a comparative analysis of the different solutions considering three main aspects:

- The spacial and temporal scale of the simulation solution, and its accuracy.
- Adaptability of the solutions to technology modifications. It includes new services, modifications of current solutions and new network models.
- Issues related to the usability and capacity of integration with external tools.

Among the different network techniques that are to be analyzed by the simulation tools, the ones that are likely to have a highest impact on the EMF population exposure are those targeted to access selection and resource allocation, since they are closely related to the usage and performance of wireless connections. First, access selection selects the wireless technology used by the services, which, as has been seen in Chapter 2, has a relevant impact on the overall population exposure. Besides, different management solutions may improve the efficiency of radio resources, thus reducing either the number of transmitted packets or the connection time, depending on the type of service.

A common aspect from most of the works promoting new access selection techniques, is that they compare their performance with that shown by legacy alternatives, but they do not assess the gap of their approach with the best possible solution [Fal10; Tos12]. Other works make use of optimization techniques [Luc12; Xue12], but they do not consider the service performance, leaving aside relevant aspects, such as the handover. In particular, [Luc12] focuses on load balancing and efficient radio resource management mechanisms, and [Xue12] provides greater relevance to the physical parameters of wireless technologies, and it is therefore more constrained in terms of scenario size and number of network elements. Besides, different mathematical techniques have been exploited, such as Game Theory [Niy12] and Multi-Attribute Decision Making [Gu09]. In this latter case, the works do not usually obtain the optimum solution.

As a response to the new requirements to evaluate network solutions, a generic simulation framework, Generic Wireless Network System Modeler (GWNSyM), has been developed. This chapter discusses the functionalities provided by this framework, as long as its design principles. Afterwards, GWNSyM operation is assessed by evaluating access selection policies on a complex heterogeneous scenario [Cho16]. This analysis provides the overall optimum and sets the basis to further and more accurate evaluation, which will be presented later.

Table 3.1: Coverage of different parameters by analysis alternatives for wireless networks (LTE). A subjective ranking is established per parameter, filled circles mean good rank while empty circles represent poor rank

	Parameter	Link Level Simulation	System Level Simulation	Network Simulation
Scenario characteristics	Scenario complexity: # of users and base stations	◔ Due to the great level of detail within this type of tools, the number of elements is usually low; typically one access element and a number of users [Sch12]	◕ Some simplifications are usually considered, so that the number of elements that can be included in the experiments is usually larger	◑ Simulation time required to analyze large network deployments is usually unacceptable [Fuj03], without the use of parallelization techniques [Pel11]
	Time dimension: time that can be simulated, and possibility to keep track of services evolution	◔ Due to the computational load [Cab11], simulated time is usually rather short, no need to keep track of services history	◔ The use of heavy development environments (Matlab) usually prevents from using long simulated times	◕ Service evolution is usually considered; however, rather large computational time is required for long experiments
	Accuracy: degree of precision of the models that are used in the different experiments	● The detailed modeling of the lower layer mechanisms is the main goal of link level simulators, which ensures a rather accurate outcome	◕ Some simplifications are usually taken, although existing works follow the recommendations of the 3GPP	◕ Although some simplifications are assumed, precise implementation of the protocols is used
Technological landscape	Architecture shift: possibility to add support to new networking paradigms: SDN and NFV	◔ As link-level solutions, they do not usually consider architectural issues, focusing on lower layers mechanisms and techniques	◕ Some of the possibilities brought by novel networking functionalities (tighter cooperation schemes) can be usually modeled	● Although it would require a large implementation effort, the integration of novel architectural approaches is usually possible
	Support for different technologies and solutions/techniques	◑ They are rather limited to the functionalities that were included from the beginning. The integration of different technologies is quite complex	◕ They usually have the flexibility to incorporate novel techniques and solutions due to some simplifications in the modeling of the physical layer	◕ Broad network simulation platforms, such as `ns-3`, are rather flexible, and allow the integration of different technologies and novel techniques and solutions
	Services modeling. Saturation conditions and/or constant load is assumed	◔ Very little attention is paid to the services modeling; the solutions usually focus on how the packets arrive at the lower layers	◕ Basic services characteristics can be included, although it is more frequent to see constant load and/or full-buffer [Tar15]	● In network simulators real applications and services can be even used to generate the corresponding traffic
Additional aspects	Specific Vs. Generic purpose and learning curve	◕ Since their scope is usually very focused, link-level frameworks require a shorter learning curve	◑ They are much more focused than network simulators, but not all its components might always be relevant	◔ They are usually rather heavy frameworks, and it takes a long time before being able to analyze the required scenarios
	Use of complementary methodologies. Use of optimization techniques?	◔ The usual goal is to analyze the performance of particular techniques, and the do not usually seek optimum performances	◑ Although this is not within the objectives of this type of tools, optimization techniques could be integrated	◑ The simulator architecture can be exploited to get an overall vision from which global optimization strategies might be executed

3.1 Simulation tool

In order to cope with the aforementioned requirements, GWNSyM has been designed as a simulation environment that allows flexible modeling and analysis of rather large and complex communications systems. The framework is specially focused on the evaluation of service performance upon different network techniques.

According to the features of the different simulation solutions presented at the beginning of this chapter, system-level simulation seems to be the best suited solution to evaluate the network techniques related to population exposure, since they cover large scale scenarios both in time and space. For instance, they are used to analyze the benefits and drawbacks of network densification, or to evaluate the impact of self-organization techniques on the overall network performance.

Although there exist general-purpose system-level simulators, different tools tailored to specific targets are typically used by vendors, operators and academy to investigate possible performance enhancements. System-level simulations are usually divided into static and dynamic approaches [Chu13] attending to their ability to consider the temporal axis.

In static simulation, the aim is to study the average performance of the networks over large areas and for long periods of time. In this case, simulations are typically based on *Monte Carlo* methodologies, where multiple and independent snapshots are used to evaluate network performance in a static manner. *Monte Carlo* simulations have been widely used in network planning and optimization. In each snapshot, the time domain is neglected and a random realization of the network is generated, following appropriate distributions (e.g. deployment of users and base stations). The performance of the overall network is then analyzed, so as to obtain KPIs of a particular configuration, for instance in terms of throughput and outage probability. Since the time domain is not considered, *Monte Carlo* experiments can be easily implemented and its simulation time can be significantly reduced using parallelization techniques. In order to get statistically representative results, system-level static simulations require a high number of snapshots. Fast simulation algorithms, able to support hundreds of snapshots within a limited period of time, are thus necessary to assess the overall performance of the network. On the downside, this kind of simulation is not suitable to study dynamic parameters such as handovers, or the temporal evolution of services.

In contrast, dynamic system-level simulation aims to model the system operation, paying special attention to the network dynamics. In this case, the evolution of the network over time is taken into account, and the simulation allows keeping the state when moving between snapshots, thus considering the temporal axis. In order to account for the network dynamics, this kind of simulation implements features of both users (such as mobility models) and service traffic (traffic models), as well as the radio channel fluctuations over frequency and time (e.g. correlated shadowing, fast fading). In this way, the behavior of different performance parameters, such as handovers, power control, and scheduling can be analyzed in terms of convergence time, or delay. Accordingly, the running time of a dynamic system-level simulation significantly increases with the network size, and the performance is typically assessed in terms of QoS (throughput, drop rates, etc.).

All in all, GWNSyM adopts the main concepts of system-level simulation, allowing both

static and dynamic approaches, so as to cope with large networks and to be able to assess service performance.

Furthermore, by using some abstractions, it allows to configure the relationship between the different network entities (both logical and physical) in a modular way, so as to ease the addition of new solutions and architectural modifications. The simulation framework has been implemented as a set of *C++* libraries, where new models can be easily added.

Usually, proprietary solutions require simulation models to be adapted to such framework (e.g. inheritance, dependencies of libraries, etc.). In contrast, as a primary requirement, GWNSyM has been implemented to perform weak binding between the models implementation and the framework itself, so as to foster code portability and addition of external implementations. As will be seen afterwards, simulation models can be implemented as plain *C++* classes or *structs*.

GWNSyM defines two types of generic abstractions: *NetTypes* and *NetActions*. *NetTypes* model any network entity with a defined configuration, ranging from User Equipments (UEs) to operators, and store the state of such entity during the simulation. The last are stateless logic that implement the interactions that take place between different network entities at a snapshot, ranging from physical models (e.g. propagation), to networking solutions or policies.

In the following, the main aspects of the GWNSyM simulation framework are presented, highlighting the abstractions to model the system, as well as the overall simulation flow fostered by the tool. Afterwards, the overall concepts of the library model are explained so as to illustrate how the simulation tool is used. In addition, Annex A depicts a practical example to illustrate the flexibility and simplicity of scenario definition within the GWNSyM framework. Furthermore, the simulation tool has been made publicly available[1] along with working examples and scenarios.

3.1.1 GWNSyM Simulation principles

Network elements

In order to facilitate modifications of the system architecture, GWNSyM features a configuration system that specifies both the static configuration of the network entities, as well as the ownership relationships between them.

The goal is to allow us to change ownership dependencies between network elements by simply modifying its configuration, instead of requiring hard-code refactoring. Since we consider network entities in a generic way, this applies to both physical and virtual entities. In this sense, architectural changes that imply shifting functionalities between entities can be accomplished by simply modifying their ownership relation.

Within GWNSyM, network entities are defined in terms of the element they implement and its configuration, so that all entities with the same model and configuration define a *NetType*. In turn, the element model specifies the actual implementation of the network entity, while the configuration indicates both static parameters shared by the *NetType* elements and entities owned by each of them.

---

1 `https://github.com/ldiez/GWNSyM`
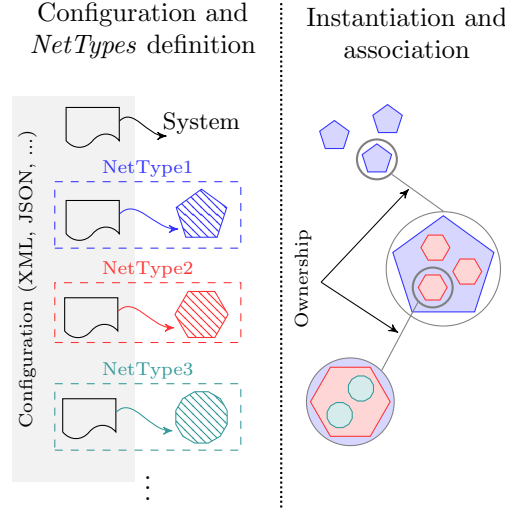
Figure 3.1: Abstract representation of the *NetType* instantiation within GWNSyM. Dashed colored shapes represent models and solid ones indicate instances of *NetType*, including its model and configuration

Figure 3.1 depicts an abstract representation of the instantiation system in GWNSyM. First, models, represented by dashed colored shapes, are attached to a configuration, conforming the *NetType*. Then, the configuration indicates the ownership relationship of different *NetTypes* instances. As can be seen, on the right hand side of the picture, once instances are created the dependencies indicate how networks entities are aggregated.

For instance, a *NetType* can be defined for LTE base stations. The model would store the state of the base station during the experiment, such as the number of connected users, while the configuration would establish static parameters, such as the operating frequency. Furthermore, the configuration might also provide the number of cells that belong to the base station, so that different configurations would be defined for macro or small base stations.
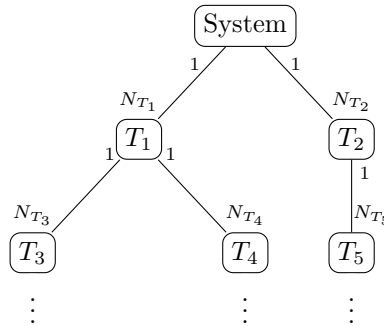
Figure 3.2: Example of instantiation tree GWNSyM. Boxes indicate *NetTypes* and lines connecting them show the ownership association with the corresponding multiplicity

As a result of the configuration, GWNSyM instantiates the elements of the system in a recursive way, leading to a tree of dependencies. Figure 3.2 shows a simple example of such tree, where T_i represents a *NetType* and N_i indicates multiplicity values. The base of the tree is the scenario whose configuration establishes the first level of network elements, directly depending on the scenario.

Afterwards, for each element of the sets T_1 and T_2 the tool instantiates the sets of elements belonging to them. In the example depicted in Figure 3.2, the scenario has N_{T_1} elements of type T_1, each of them containing N_{T_3} entities of type T_3. It is worth noting that the instantiation and connection between entities is transparent to the library users, and different structures can be defined by a simple modification of the configuration.

Once *NetTypes* are defined and the network entities are instantiated, GWNSyM provides the means to address sets of entities by traversing the tree, employing the so-called *set resolution operator*. Using as example the tree dependencies depicted in Figure 3.2, we can see that each element T_1 has a set of elements of T_3. In this sense, the expression $T_1 :: 1 :: T_3$ would indicate the set of elements of *NetType* T_3 belonging to first element of T_1. Likewise, $T_1 :: * :: T_3$ would refer to the set comprising the elements T_3 that belong to all T_1 elements. As will be seen afterwards, this feature is very useful to indicate the set or sets of network elements that will be passed around to the *NetActions* during the simulation.

Simulation workflow

Along with the *NetTypes*, which define network entities, the simulation framework provides another abstraction to model events that take place during the simulation, the so-called *NetActions*. In short, *NetActions* model the system behavior in a wide sense, ranging from physical phenomena (e.g. propagation models) to specific policies (e.g. access selection).

Each *NetAction* takes as input parameters one or more sets of entities, which are addressed using the *set resolution operator* described above. Then, during each snapshot of the simulation the actions are applied in a sequential way, updating the system state. While this approach is generic, there might be situations where actions should be only applied either at the beginning or the end of the whole simulation process. The deployment of access elements, such as base stations, or the creation of trace files are the most straightforward examples. These situations are considered by defining two special types of actions: *InitNetAction* and *EndNetAction*. Those from the former type are executed at the beginning of the simulation in a specified order. Likewise, the actions defined as *EndNetAction* are all performed at the end.

Altogether, the general simulation process is represented in Figure 3.3. In the first step both the *NetTypes* and *NetActions* are defined. Then, based on the configuration, the network elements are instantiated and the dependency tree is generated. After the instantiation stage, actions are performed. Firstly, the set *InitNetActions* are executed, setting the network elements in the default state. Afterwards, at every snapshot, the set of *NetActions* are applied, in a defined order, over the set or sets of elements indicated by means of the *set resolution operator*. Finally, the simulation finishes by executing the set of *EndActions*.

It is worth noting, that after applying a *NetAction*, the state of the involved network entities may be modified. In this sense, both static and dynamic simulations can be
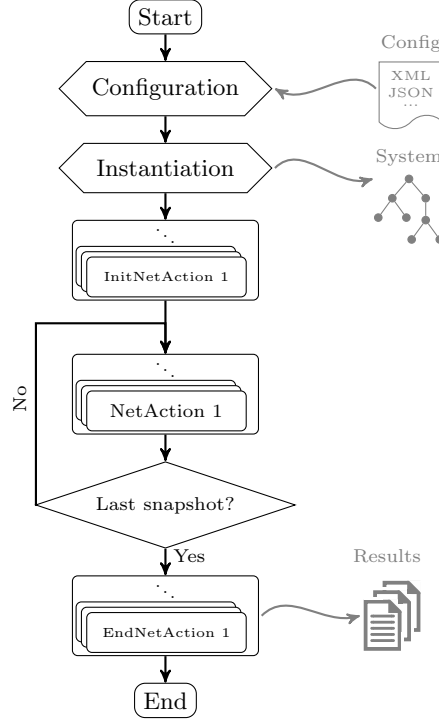
Figure 3.3: GWNSyM general simulation process

performed, being the only difference the state at the end of the snapshot and whether or not that state is used in the next snapshot.

3.1.2 Library model

The implementation of the GWNSyM has been carried out with the goal of code re-usability in two ways: (1) that code generated within GWNSyM can be used in other environments, and (2) that legacy code can be easily integrated within the framework. To this end, the simulator has been implemented so that it requires low binding between the framework and the actual implementation of the different models, both *NetTypes* and actions, including *NetActions*, *InitNetActions* and *EndNetActions*.

As for actions, the only requirement imposed by the simulation framework is that they can be called. In this sense, actions need to be implemented either as simple functions or callable classes[1]. Additionally, the actions have to accept as input parameters sets of *NetTypes*.

Furthermore, actions can be defined to accept individual elements or groups of them. For instance, calculating the pathloss between a base station and a UE can be implemented by only taking one element of each type. On the other hand, resource allocation would require to consider all the user devices at the same time, so as to ensure that the capacity

---

1   In *C++*, callable classes are easily implemented by defining the function-call operator (i.e. *()*)

of the base station is not exceeded.

In order to avoid additional constraints over actions, GWNSyM has been equipped with an inference system that analyzes the definition of actions to call them in the correct way. To explain the inference system, it is convenient to differentiate between *NetTypes* (an abstraction), and instances of them. In the following we will use the term reference to refer to instances of elements (e.g. an instance of a *NetType*). Likewise, sets of instances will be referred to as vectors[1]. The goal of the inference system is to adapt the way the actions are called according to its particular implementation, provided they follow one of the predefined patterns. In particular, the GWNSyM has been designed to accept actions with one-to-one, many-to-many, and one-to-many semantics.

For instance, if an action is applied to two sets of *NetTypes*, GWNSyM accepts calling operators taking as parameters:

- One reference of each type. In this case, during each snapshot, and without user interaction, the framework iterates over both sets, passing to the action all combination of elements.

- A vector containing the elements of one type and a reference of the other one. Similar to the previous case, the system iterates over the set whose elements are passed individually, and it calls the action passing the vector and each element of the other set.

- One vector for each type. This is the easiest case, the system passes the vectors and the action gets called only once in each snapshot.

The overall inference process is depicted in Figure 3.4. Before starting the simulation, the system iterates over all the defined actions and, for each of them, it obtains the sets of involved *NetTypes*, which are defined by using the *set resolution operator*. The reader may refer to Annex A for a practical example of the attachment between sets of *NetTypes* and *NetActions*. Then, the action is analyzed and stored in one of the so-called *ActionWrappers*, along with the vectors of involved elements. The different types of *ActionWrappers* implement the calling patterns that are accepted by the system, and expose function call operators without parameters.

For instance, as illustrated in Figure 3.4, an action taking as parameters a vector and an reference (i.e. one-to-many pattern) would be stored in an *ActionWrapper* and executed following a double-dispatch model (i.e. the system calls the *ActionWrapper* that afterwards calls the actual action). The wrapper iterates over the elements of the second vector, and for each of them, it calls the action passing such element and the first vector as argument.

This technique provides a high degree of flexibility to implement actions, and simplifies the integration of legacy models into GWNSyM. Besides, the double-dispatch approach enables mapping network elements and actions, as well as call adaptation. Hence, it ensures

---

1   From a implementation perspective, references are used in a wide sense, so that they include proper *C++* references and pointers with or without qualifiers (e.g. *constness*). Similarly, vectors can be replaced by other containers that are also accepted by the tool.
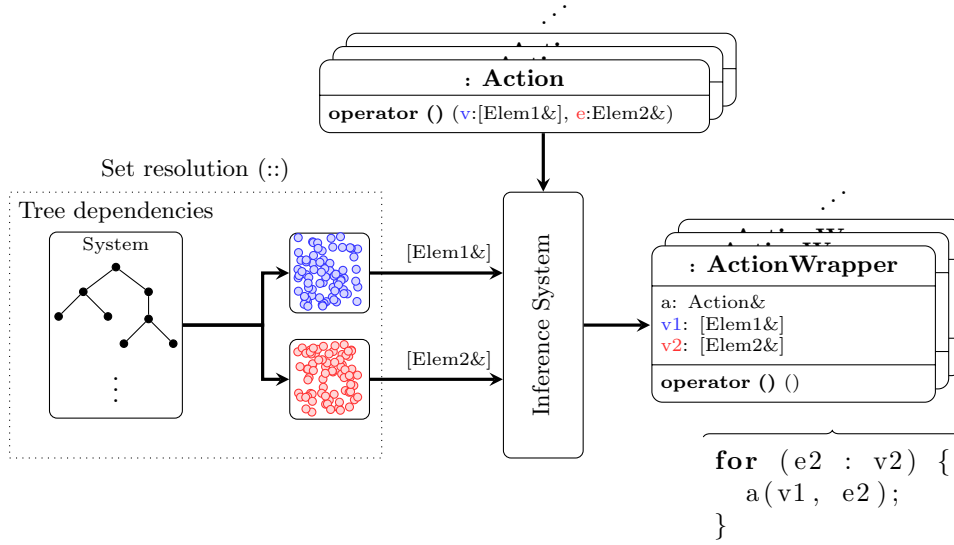
Figure 3.4: Call to inference system that maps elements to actions and adapts actions call to actual implementation. In the figure, the symbol *[]* indicates a set or container of network elements, while & is used for both references or pointers
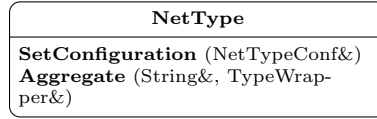
Figure 3.5: Optional interface of *NetTypes* to perform aggregation and locally store configuration.

that no run time overhead is added when the simulation starts, besides the negligible effect of the double dispatch. Furthermore, it is worth noting that no copy is performed during the inference system, and that GWNSyM takes over the memory management of the whole simulation.

As has been mentioned before, the definition of network entities within GWNSyM does not require tailored implementation (e.g. inheritance), but plain *C++* is accepted by the framework. In order to provide this degree of flexibility while keeping the dependency among entities and configuration, separate structures (e.g. dependency tree) are internally created by GWNSyM to allow the simulation.

However, there exist cases where it is necessary to perform the actual aggregation of entities or to have access to the configuration of a particular network element. To this end, the framework defines an optional interface, as depicted in Figure 3.5, which is detected and called during the instantiation of entities. The first function of the interface (*SetConfiguration*) allows network entities to store a reference to its configuration. Besides, a second function can be implemented to get access to the entities belonging to the current one (one level below in the dependencies tree), enabling actual aggregation. When this

function is called, the corresponding sets of entities are passed along with the name of its *NetType*, so that they can be unequivocally identified and aggregated as desired.

As can be observed, the implementation of GWNSyM allows flexible configuration of the simulation setup, so that the level of binding can be tailored according to the simulation needs. Altogether, GWNSyM is responsible of managing the simulation flow, while the user can decide the level of complexity of the system model, since the framework is able to adapt to user requirements.

3.2 Optimum access selection in heterogeneous networks

As was mentioned at the beginning of this chapter, access selection and resource management play a crucial role on EMF population exposure. In this section, we present a generic model to obtain the optimum network performance according to different operator policies, which will be used as the basis for more detailed studies. This analysis will allow us to validate the methodology fostered by GWNSyM, as well as to identify and build some common simulation models, suited to different studies. In particular, we present herewith a service model that will be used in following chapters.

Since we focus on the service modeling and overall evaluation, some simplifications are assumed, following the same approach as [Luc12; Nas04]. In particular, we simplify the differences between technologies, and the impact that link conditions might have over the capacity of radio resources.

On the other hand, service evolution is considered, and we thus take into account the willingness of the users to have an active connection, depending on whether or not a service session is ongoing (i.e. not all users want to be always connected), as well as according to the current state of the service (for instance, dropped services).

The access selection policies are modeled as optimization problems, to find the optimum solution according to the specific configuration, so that a problem instance is executed at each snapshot where access selection is performed accordingly.

In what follows, the main components of the system will be described, including both service models and the optimization problem for access selection. Afterwards, and based on such models, different access selection policies will be analyzed.

3.2.1 Service modeling

One of the most relevant features of dynamic system-level simulation is its capacity to evaluate the impact that network techniques have over the QoS. As a common element for the different analysis, we have implemented a service model that considers the temporal evolution and the effect that access selection decisions may have over the service. The model is implemented by means of a state machine, where each service can have, at each snapshot, one of the following states:

- **Idle:** The service is not active, and it does not, therefore, require any connection.

- **Active:** The service has started a new session, and it has been accepted by some of the available access alternatives.

- **Rejected:** The service started a new session, but it was not accepted by the network.
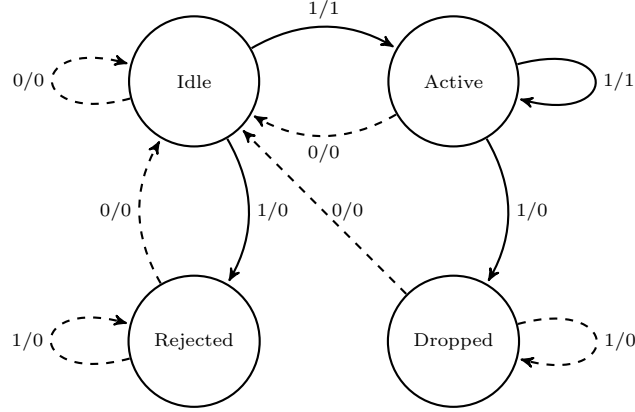
Figure 3.6: State machine for a specific user/service pair. Solid lines indicate transition path to the final state

- **Dropped:** The service had an ongoing session accepted by the network, but it eventually stopped before its correct finalization, due to user mobility, fading, etc.

With the aforementioned states, we can establish the evolution of the state for a particular service as shown in Figure 3.6. The arrows represent state transitions, which depend on two aspects: the service activity and the outcome of the access selection procedure in the previous snapshot. These are represented as a pair of numbers (*activity/selection*) on top of each arrow.

It is straightforward to see that, depending on the previous state and the current activity, there are some cases in which the corresponding service is not considered in the following access selection decision, represented in the figure by dashed arrows. No matter the previous state is, when a service has not an active session, it goes to the *Idle* state. In addition, whenever a service has been rejected or dropped, it is not considered during the access selection, so that a connection that has been lost cannot be recovered until it goes back to the *Idle* state (i.e. until a new session starts).

It is worth noting that the service model is independent of both the access selection policy and the implementation to model network behavior. In this way, the same service pattern can be used in scenarios with different models and configurations.

3.2.2 Generic access selection problem

The access selection problem aims at establishing the optimum association of the current active services among the available access alternatives. The scenario comprises $\mathcal{A}$ available access elements of different RATs, thus having different capabilities in terms of coverage and capacity. It is also assumed that there are $\mathcal{U}$ users, who can start $\mathcal{S}$ different services, equipped with a terminal able to establish a connection with any of the involved technologies, but not simultaneously.

In the scope of this work, resources are modeled using a generic and discrete capacity unit coined Traffic Unit (TU), no matter it refers to time slots, like in Time-Division Multiple Access (TDMA) systems, codes (CDMA) or sub-carriers, like in Orthogonal

Frequency-Division Multiple Access (OFDMA) based systems. Any service would require a number of TUs to be handled by the network, so that in the case a service cannot be granted the required TUs, it is rejected or dropped, depending on the circumstances. Besides, each access element has a pool of resources that can be assigned to users, according to its technology.

Over the aforementioned scenario, the access selection decision is implemented as a Binary Linear Programming (BLP) problem, in which there are $|\mathcal{U}| \times |\mathcal{A}| \times |\mathcal{S}|$ basic variables, which are defined as follows:

$$x_{ijk} = \begin{cases} 1 \text{ if user } i \in \mathcal{U} \text{ uses network } j \in \mathcal{A} \text{ for service } k \in \mathcal{S} \\ 0 \text{ otherwise} \end{cases} \tag{3.1}$$

The access selection problem, defined in Problem 1, aims to maximize the summation of the utility functions, U_{ijk}, assigned to each connection. In turn, the utility function can be tailored according to different criteria. It is worth noting that this utility remains constant for each problem instance, thus ensuring the linearity of the proposed model (i.e. U_{ijk} does not depend on x_{ijk}).

Problem 1 (Generic access selection)**.**

$$\boldsymbol{max.} \qquad \sum_{i,j,k} U_{ijk} \cdot x_{ijk} \tag{3.2}$$

$$\boldsymbol{s.t.} \qquad x_{ijk} \in \{0,1\} \qquad\qquad \forall i \in \mathcal{U}, j \in \mathcal{A}, k \in \mathcal{S} \tag{3.3}$$

$$\sum_{j} x_{ijk} \leq 1 \qquad\qquad \forall i \in \mathcal{U}, k \in \mathcal{S} \tag{3.4}$$

$$\sum_{i,k} x_{ijk} D_k \leq C_j \qquad\qquad \forall j \in \mathcal{A} \tag{3.5}$$

$$\tag{3.6}$$

In addition, we pose some constraints to ensure a sensible behavior of the system. First, Eq. 3.3 ensures that the optimization variables are binary, so that it is not possible to assign fractions of services (i.e. multi-path/multi-flow is not considered). Then, Eq. 3.4 forces that a service can be only connected to one access element. Finally, Eq. 3.5 ensures that the capacity of network elements is not exceeded, according to the total number of TUs that the element can assign (C_j) and the demand by each service (D_k).

It is worth mentioning that for each instance of the access selection problem, some optimization variables can be discarded due to lack of physical connectivity or traffic demand. The former is calculated based on the distance between users and access elements, while the later follows the service model described above.

3.2.3 Access Selection Strategies

Although a broad range of possibilities exist to define the utility function, the validation presented in this chapter will focus on two particular aspects: pricing and RAT affinity

or preference. In addition, we will also study the interaction of the mentioned metrics with some widely assumed QoS parameters. Firstly, we will consider the cost of changing the access element when a service is ongoing (i.e. handover), which can lead to service interruption. Besides, we will prioritize ongoing services over new ones, since it is widely assumed that dropping an already established call is, in general, worse than rejecting a new one.

In this way, the utility function is built as a weighted summation of different parameters, that are defined below. In the following the sub-indexes keep the same meaning as before, so that i, j, k hold for user, access element and service, respectively.

Connectivity

The connectivity criterion, $U_{C_{ijk}}$, assigns utility to ongoing sessions. This criterion is defined for each service as follows:

$$
\forall j \quad U_{C_{ijk}} = \begin{cases} 1 \cdot \frac{D_k}{\max_{\forall k} D_k} & \text{if user/service } i/k \text{ was connected} \\ \xi_c \cdot \frac{D_k}{\max_{\forall k} D_k} & \text{otherwise} \end{cases}
\tag{3.7}
$$

where D_k is the capacity required for service k and ξ_c is a design parameter selected to ensure that ongoing services are always given higher priority than new ones for all the considered capacities: $\xi_c < \frac{\min_{\forall k} D_k}{\max_{\forall k} D_k}$.

Handover criterion

We also consider the influence that the cost of change (handover) has on the service performance. For that, we model the $U_{H_{ijk}}$ criterion as follows:

$$
U_{H_{ijk}} = \begin{cases} 1 & \text{if user/service } i/k \text{ was previously connected to } j \\ 1 - \xi_h & \text{otherwise} \end{cases}
\tag{3.8}
$$

where $\xi_h < 1$ is a design parameter which depends on the particular configuration, to foster keeping connected to the same access element.

Price criterion

This criterion models the monetary preference of end-users, who would opt for more economical connections. The utility perceived by a user increases as far as the base stations offer a lower price. Thus, the utility is a decreasing function of the price.

Since we want to enable relative comparisons (i.e. based on discount percentages) between offered prices, we propose using a logarithm function as depicted in Figure 3.7a. According to this function, the price criterion is defined as follows:

$$
\forall k \quad U_{P_{ijk}} = \begin{cases} -\log(Pr_j) & Pr_j \in [Pr_{min}, Pr_{max}] \\ 1 & \text{otherwise} \end{cases}
\tag{3.9}
$$

(a) Price utility

(b) Base station price as a function of current free load
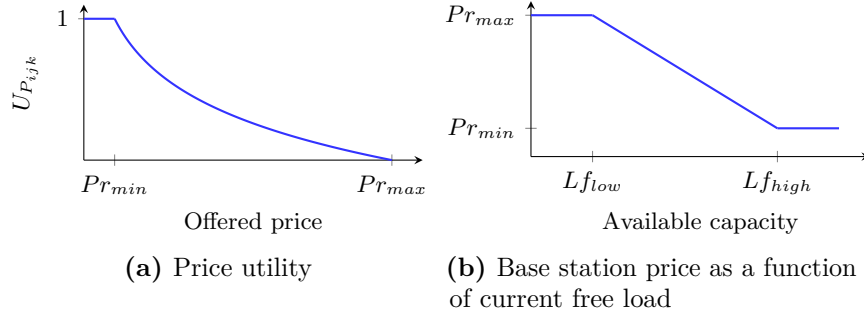
Figure 3.7: Price-related utility functions

where Pr_j holds for the price offered by the base station j given as monetary units per time and capacity unit. Besides, Pr_{min} and Pr_{max} correspond to the highest and lowest price offered by the access element j to use its resources, respectively. As can be observed, the first one is the fee below which end-users would not perceive any utility gain. We also assume that those base stations that offer a price higher than the maximum an end-user would be willing to pay, would be discarded. Relative price units are used, and therefore the maximum price offered by any of the base stations would be 1.0.

Moreover, we also consider that base stations are using their offered price so as to encourage or deter users to connect to them. In particular, we implement an scheme in which base stations are able to dynamically modulate the offered price according to the free load, as depicted in Figure 3.7b.

As can be seen, when the base station is highly loaded (available capacity lower than Lf_{low}), the offered price is the maximum one. On the other hand, if the base station load is low, the offered price decreases to a minimum configured level. For the sake of simplicity, a linear decreasing trend has been used between these two points

RAT Affinity criterion

This criterion is designed to favor that services are attached to the most suitable technology, to perceive better QoS. For instance, we could use WiFi accesses for data transfer services, since they benefit from higher bandwidths, but do not require a strict delay, while we would establish a preference of voice services towards cellular base stations, better suited for them. The corresponding component of the utility function ($U_{R_{ijk}}$) is therefore defined as follows:

$$\forall i \quad U_{R_{ijk}} = \begin{cases} 1 & \text{if service } k \text{ has affinity towards technology of } j \\ \xi_r & \text{otherwise} \end{cases} \tag{3.10}$$

where $\xi_r < 1$ is a design parameter, which modulates the relevance given to this particular aspect of the utility function.

Utility Function

By linearly combining the different criteria, we can define an overall utility function, which establishes the corresponding access selection strategy, as follows:

$$U_{ijk} = \rho_c \cdot U_{C_{ijk}} + \rho_h \cdot U_{H_{ijk}} + \rho_p \cdot U_{P_{ijk}} + \rho_r \cdot U_{R_{ijk}} \qquad (3.11)$$

where the parameters $\rho_{\{c,h,p,r\}}$ will be used as binary parameters that enable or disable the different access selection criteria. Furthermore, in order to favor one of the active criteria over the others, the design parameters ξ_c, ξ_r have been fixed to 0.4 and 0.8 respectively, while the value for ξ_h is discussed below.

First, we assume that RAT affinity criterion is disabled (i.e. $\rho_r = 0$). We then compare the utilities for two different access alternatives for an ongoing service. In this case, the value of $U_{C_{ijk}}$ is not relevant, since it will be alike for the two accesses. The end-user would have two connectivity options, the one to which it is currently connected to (1) and another one (2), which is offering a cheaper fee than (1). In particular, we define the price of the second alternative as function of the first one, so that $Pr_2 = (1 - \Delta_{Pr}) Pr_1$, with $0 \le \Delta_{Pr} \le 1$. Hence, we can write:

$$U_1 = U_2 \quad \rightarrow \quad 1 - \log(Pr_1) = (1 - \xi_h) - \log(Pr_1(1 - \Delta_{Pr})) \qquad (3.12)$$
$$\rightarrow \quad \xi_h = -\log(Pr_1(1 - \Delta_{Pr})) + \log(Pr_1) \rightarrow \qquad (3.13)$$
$$\rightarrow \quad \xi_h = -\log(1 - \delta_P r) \qquad (3.14)$$

Therefore, if we consider that a discount of $\Delta_{Pr} = 0.2$ should be enough for an end-user to change its current access, the value of ξ_h to be used is ≈ 0.1.

All in all, Table 3.2 describes the six different strategies that will be analyzed, where each of them corresponds to a different instance of the utility function.

3.2.4 Scenarios analysis

This section presents a thorough analysis of the different access selection strategies using GWNSyM with a twofold objective. First, this study aims to validate the feasibility of the methodology fostered by the simulation framework to evaluate the potential impact of

Table 3.2: Access selection strategies

	Price		RAT affinity		Combined	
	w-HO	w/o-HO	w-HO	w/o-HO	w-HO	w/o-HO
Connectivity (ρ_c)	1	1	1	1	1	1
Handover (ρ_h)	1	0	1	0	1	0
Price (ρ_p)	1	1	0	0	1	1
RAT Affinity (ρ_r)	0	0	1	1	1	1

access selection strategies over large networks. Besides, this analysis provides a overarching view of the potential impacts of the defined strategies over the services.

We study a 200×200 m^2 scenario comprising access elements of two different types, as depicted in Figure 3.8. The first type corresponds to a traditional cellular technology, with a coverage of 150 m, effectively covering the whole scenario, and a capacity of 16 TUs, while the second one mimics a WiFi access point, with a range of 50 m and a capacity of 8 TUs.

Over such scenario we deploy different densities of users which, ranging from 20 to 200, which move according to a *Random Waypoint* model, whose parameters are given in Table 5.3. Each user generates flows belonging to two different service types, following *ON-OFF* models, having some preference towards a particular technology. All the base stations use the pricing policy that was depicted in Figure 3.7.

For each strategy defined in Table 3.2, we run 10 independent simulations of 1 hour, and we represent the average values. Each snapshot of the simulation is obtained applying the aforementioned models (users movement and traffic pattern) every 10 seconds, so that 360 snapshots are evaluated for every simulation. Besides, within a simulation, the output of one snapshot is used to feed the following one, so that the optimization problem corresponding to each access selection strategy can be properly posed according to the service model. The corresponding instances of Problem 1 have been solved using the optimization toolkit GNU Linear Programming Kit (GLPK) [Gnu17].

We assess the performance of the different strategies, and for the different services, based on the following metrics:

- Success Rate (**SR**): probability that the service is successfully finished. That is, it is neither rejected nor dropped.

- Handovers (**HO**): average number of handovers which are carried out during the service lifetime.
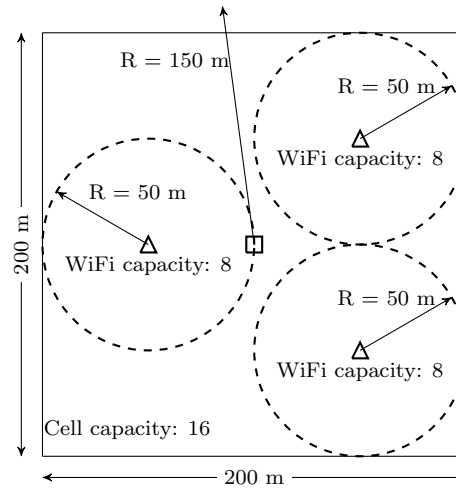
Figure 3.8: Network deployment

Table 3.3: Scenario configuration parameters.

Base stations	
Cellular	
Coverage (m)	150
Capacity (TU)	16
Pricing policy	$Pr \in [0.1,1.0]$
	$Lf_{low} = 0.2,\ Lf_{high} = 0.8$
WiFi	
Coverage (m)	50
Capacity (TU)	8
Pricing policy	$Pr \in [0.1,1.0]$,
	$Lf_{low} = 0.2,\ Lf_{high} = 0.8$
Mobility model	
Random Waypoint	
Speed (m/s)	Uniform distribution within [1,3]
Movement time (s)	Uniform distribution within [800,1000]
Pause time (s)	Uniform distribution [80,100]
Service model	
Service 0: Voice	
Model	On-Off
Idle (*off*) time (s)	420
Service (*on*) time (s)	180
Capacity (TU)	1
Affinity	Cellular
Service 1: Data transfer	
Model	On-Off
Idle (*off*) time (s)	480
Service (*on*) time (s)	120
Capacity (TU)	2
Affinity	WiFi
General parameters	
Simulation time (s)	3600
# of snapshots	360
# of runs	10

- Price per service (**PS**): average service price paid per time and traffic unit.
- RAT affinity (**RA**): we study the time percentage that the service was using the technology towards which it has some preference.

Before discussing the performance metrics that were obtained, we study how the setup described in Table 5.3 affects the system load, so as to fully understand the obtained results. In this sense, Figure 3.9 shows the traffic demand that was generated by all the users, relative to the overall network capacity.

Note that the traffic demand cannot be directly calculated from the scenario parameters, but it is modulated by the outcome of the different optimization problems and the service model commented above. In case of performing static simulations (without considering the temporal evolution), the offered traffic would show a linear increase against the number of users with a slope of 0.7. However, whenever a service is rejected or dropped, it is not longer considered in the future problems until it is restarted. This would effectively reduce the traffic demand, explaining the saturation behavior that is observed in Figure 3.9a.

We also represent in Figure 3.9b the relative load per RAT type. It can be observed that the cellular base station gets almost fully loaded for the three strategies, since it covers the whole scenario. Conversely, although the load of the WiFi access points also increases with the number of users, they do not get saturated. Indeed, despite the traffic demand surpasses the network capacity, there are still some available resources in these access points, since they do not fully cover the whole area.

As for the performance metrics, Figure 3.10 shows the probability that a service finishes successfully. As can be observed, when there is enough available capacity (i.e. 20 or 40 users), the usage of the connectivity parameter (U_C) in the utility function leads to having similar results for the two types of services.

On the other side, when the number of users increases, the success rate for *Service 0* is higher, since it requires fewer resources. We can also observe that the configuration which

(a) Traffic demand

(b) Carried load per RAT type
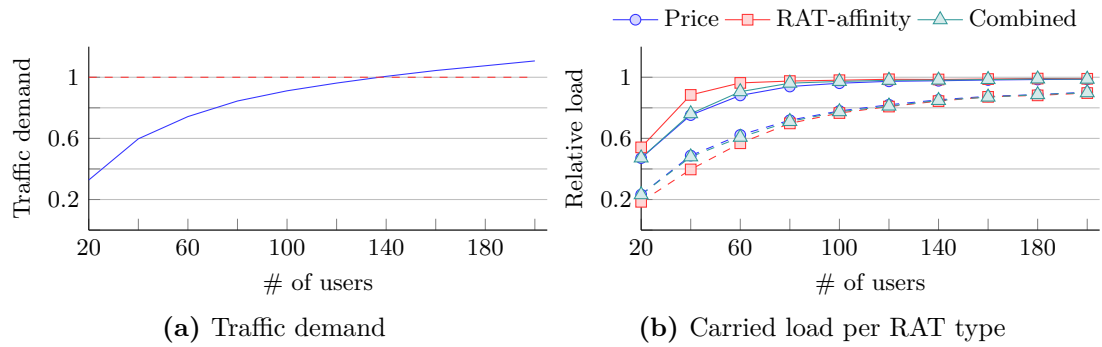
Figure 3.9: Traffic demand and carried load per RAT type. The results in Figure 3.9a are obtained from the strategy that only considers RAT affinity with handovers; for other combinations the results are mostly the same. Similarly, Figure 3.9b shows the load for the strategies with handovers; when handovers are not considered in the utility, the results are quite similar

considers *Price w/o-HO* does not follow this trend. In this case, cheaper connections will be favored (without looking at the possible handover side-effects) and therefore *Service 0* will be moved to WiFi accesses as soon as they become available, since those accesses are likely to be cheaper because the corresponding access points are less loaded, as can be seen on Figure 3.9b. This behavior is not observed for the other strategies, where the RAT affinity criterion fosters *Service 0* to be handled by the cellular base station. In addition, due to user mobility, previous connections to the WiFi Access Points (APs) might not be longer available, and the service would probably need to be dropped, since the cellular base station might be fully loaded, as depicted in Figure 3.9b.

In addition, as far as the system load gets higher, there is as well an increase of the number of rejected services, since it is rather unlikely that a new call from *Service 0* would cause a call from *Service 1* to be dropped. The additional utility of the price criterion is low, since the cell load and offered price are high. On the other hand, in those strategies fostering the RAT affinity criterion, a new call from *Service 0* would quite likely cause an ongoing *Service 1* to be dropped if it is using the cellular base station, due to the stronger weight given to this particular parameter. This, together with the fact that *Service 0* calls are longer than those corresponding to *Service 1*, explain the aforementioned situation.

Furthermore, when the handover criterion is considered in the utility function, we can observe that the price effect is mitigated. Figure 3.10b shows that when handovers are penalized, same services have similar behavior for the different configurations. In fact, the differences between services are simply due to the different capacity demand, since *Service 1* requires twice the capacity of *Service 0*. This way, when the system starts getting saturated, *Service 0* is more likely to be allocated in the scarce remaining resources.

Following with the evaluation, Figure 3.11 shows the impact of considering the handover criterion. As expected, and aligned to the previous results, we can observe a decrease of the number of handovers when the handover criterion is considered in the utility function. In general, the impact is more relevant for *Service 0*, due to its longer duration, especially when the system gets highly loaded.

Furthermore, when the cost of change is not considered ($\rho_h = 0$), we can also see, again for *Service 1*, that the RAT affinity strategy leads to a lower number of handovers, while

(a) Without handover criterion   **(b)** With handover criterion
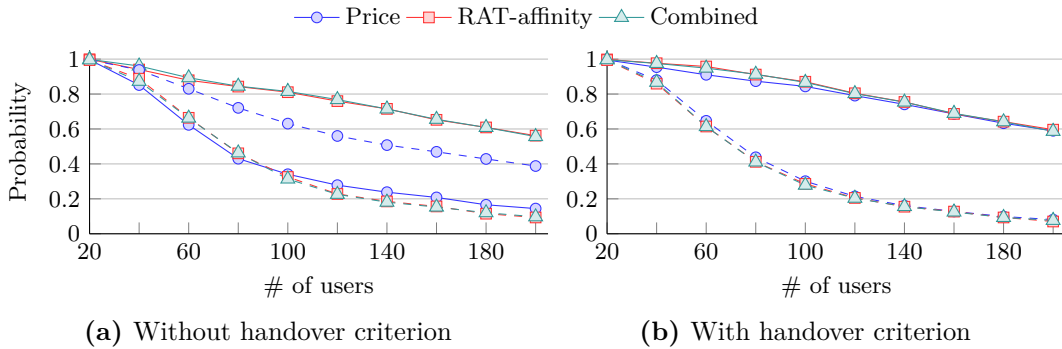
Figure 3.10: Average success rate per service Vs. the number of users. Continuous lines are for *Service 0* and dashed lines for *Service 1*

(a) Without handover criterion   **(b)** With handover criterion
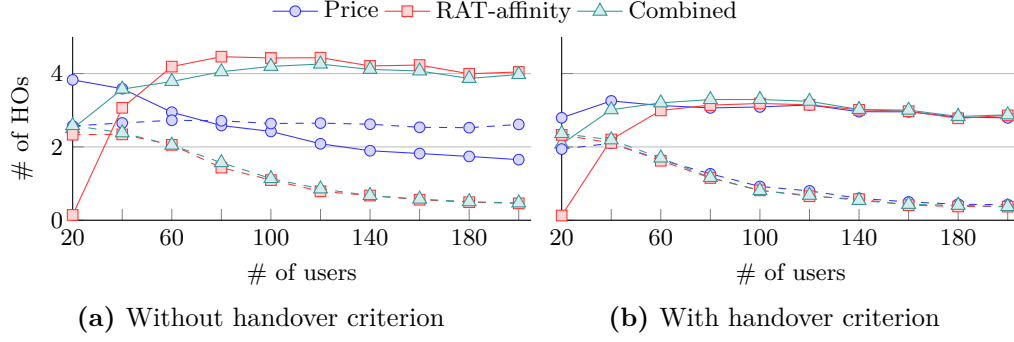
Figure 3.11: Average number of handovers per service Vs. the number of users. Continuous lines are for *Service 0* and dashed lines for *Service 1*

the behavior for *Service 0* is different. This is in part due to the *ping-pong* effect which might appear in this particular strategy, since a double change of access will not have any impact over the overall utility function. Indeed, this effect is more relevant for *Service 0* flows, since they are longer.

Regarding the price that a user needs to pay (per traffic and time unit), Figure 3.12 yields an interesting result. For *Service 0* we get the expected behavior, and the price strategy leads to lower prices, compared to the RAT affinity one. However, for *Service 1* we can see that the RAT affinity strategy leads to prices that are slightly lower than those obtained for the price-based one. The reason is that the WiFi access points, due to their coverage, are less loaded than the cellular base station and the price they offer (following the previously presented pricing policy) is therefore lower. Since services belonging to type 0 would stick to the cellular base station (no matter the price is) according to the RAT affinity criterion, the prices offered by WiFi access points are lower, as the fee paid per service of type 1.

On the other hand, this result also reflects the fact that the price-based strategy just seeks a global price reduction, without distinguishing between service types. In this case it

(a) Without handover criterion   **(b)** With handover criterion
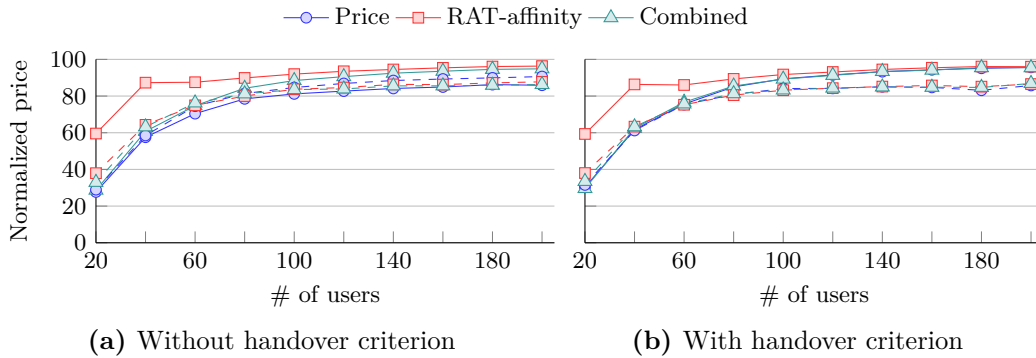
Figure 3.12: Average price per service and TU Vs. the number of users. Continuous lines are for *Service 0* and dashed lines for *Service 1*

is also worth highlighting that there is not a clear dependency on the handover criterion, since the results do not change between the utility functions that did not take it into account and those which did so. Besides, the figure also yields that the higher the carried load, the higher the price per service. It is also interesting to note that, for *Service 0*, the RAT affinity strategy always leads to a higher price, since it aims at handling all the corresponding calls by the cellular base station, which gets fully loaded, even when the number of users is still small, as illustrated in Figure 3.9b.

Figure 3.13 shows the probability for a service to be connected to the technology it has affinity to. We can see that the use of the appropriate strategy does really have a relevant effect. The values obtained for the RAT affinity strategy are much higher than those observed for the price one. We can also observe the little influence of the handover criterion. Last, the obtained results also yield that *Service 1* gets better results than *Service 0* regarding the RAT affinity criterion for all the considered strategies. This is due to the fact that *Service 1* prefers using WiFi access points and, as we have already discussed, these access alternative is less loaded than the cellular base station. In addition, it is also worth bearing in mind that those services are shorter, and it is therefore less likely that they would need to connect to the cellular base station due to user mobility after being connected to a WiFi access router.

In order to provide a global view of the overall performance yielded by the different strategies, Figure 3.14 uses a radar chart in which we represent the four figures of merit for the two service types. The edges of the different axes represent the best potential performance, while the chart center can be considered as the worst result (for the handover parameter we used the empirically observed values to establish the performance bounds). The results were obtained for the scenario with 100 users. As was already discussed, we can see the improvement brought about by the integration of the handover criterion on the utility function, since the rest of parameters do not get clearly affected, while there is an enhancement in the number of handovers that are required per service. In the case of the *Price* strategy, we can also see that the use of the handover criterion results in an improvement on the success rate for *Service 0*, although it is compensated by the decrease

(a) Without handover criterion   **(b)** With handover criterion
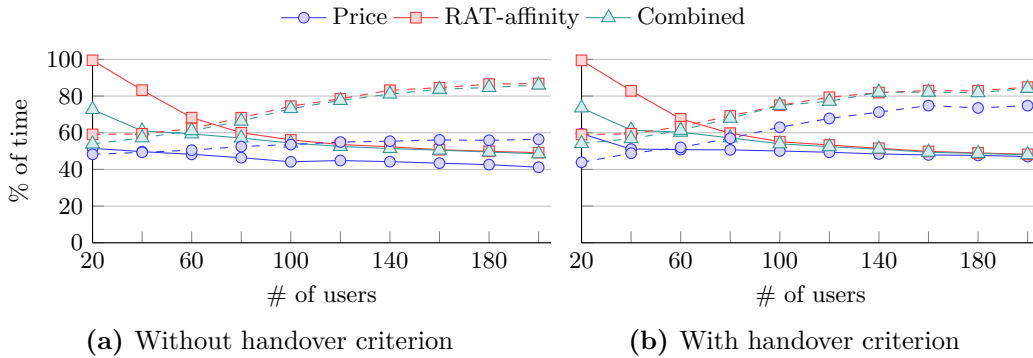
Figure 3.13: Average RAT affinity per service Vs. the number of users. Continuous lines are for *Service 0* and dashed lines for *Service 1*

(a) Without HO parameter ($\rho_2 = 0$)   **(b)** With HO parameter ($\rho_2 = 1$)
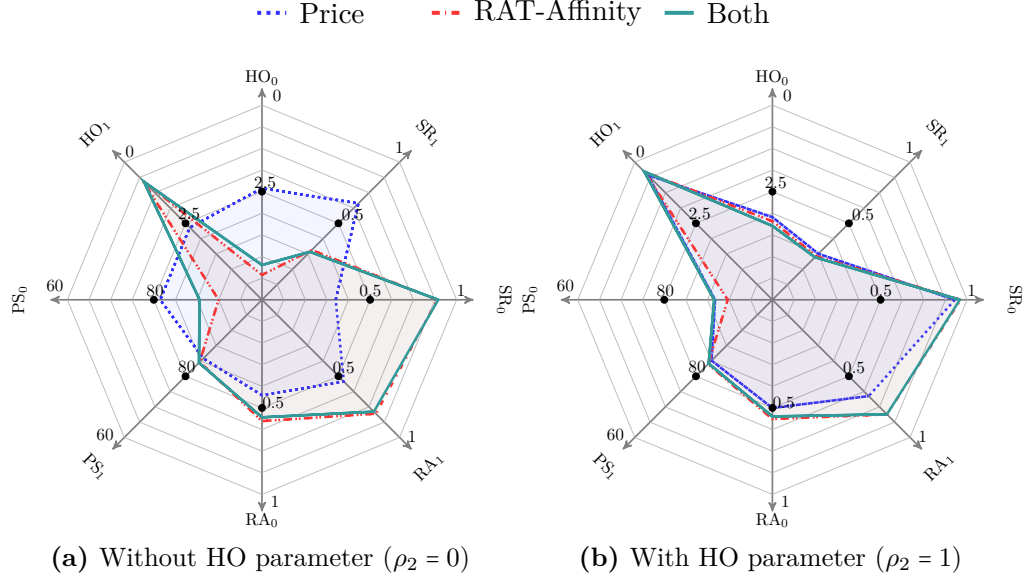
Figure 3.14: Performance for the different utility functions. # of users was fixed to 100

on the corresponding value for *Service 1*.

In general, all figures yield that the strategies considering the two goals (i.e. price and RAT-affinity) lead to performances that are rather similar to those obtained for the one fostering just RAT-affinity (without jeopardizing the price per service too much), while for the price-based strategy, the performance in terms of RAT-affinity are remarkably worse.

3.3 Conclusion

Despite the wide range of analytical and validation tools, the advent of the 5G is broadening the research landscape of cellular networks. In this regard, the research community does not only need to validate novel network techniques or solutions, but new concepts involving novel architectural paradigms have been added to the already complex playground that has to be assessed. As a consequence, it is necessary to rethink the accepted validation methodologies to fulfill these new requirements.

Although there exist solutions that allow the actual deployment of technology to be studied, such as testbeds or SDR developments, when it comes to large and complex systems, simulation is still the preferred approach. In order to overcome the limitations of simulation solutions to tackle novel technological challenges, we have presented the design and implementation principles of Generic Wireless Network System Modeler framework. It is a novel simulation framework whose design, built on top of widely accepted simulation principles, simplifies the integration of new network solutions and architectural changes, as well as the integration of existing network models. This framework adopts system-level simulation methodology and fosters systematic simulation while leaving the implementation open to the researcher.

The simulation flow provided by GWNSyM has been validated by studying the potential

performance of different access selection policies over a heterogeneous scenario. In particular, we have proposed a generic model, leveraging binary linear programing, to find the optimum expected behavior according the different policies.

Operator policies considering service price, service RAT preference, and handovers have been studied. The results highlight the impact of the utility function definition over the whole system performance. In particular we saw that, for the considered scenario, penalizing the handovers in the utility function does not only reduce the average number of handovers, but it does not jeopardize the rest of figures of merit.

The selection of this validation scenario responds to the impact that access selection and resource management may have over the population EMF exposure. In the following chapter GWNSyM will be exploited to validate more accurate radio models, advanced access selection techniques and resource management solutions over different scenarios. Besides, while the studied access selection policies do not consider EMF exposure, the optimization model could be exploited to obtain optimum bound to compare with the performance of other strategies.

CHAPTER 4

Access selection and resource management for OFDMA-based networks

During the last years we have witnessed a large increase of the mobile data traffic, and this trend is expected to remain in the near future [Cis17]. In order to support the forthcoming traffic demand and overcome the limitations of current technologies (i.e. WiFi, 3G or 4G), 5G systems are expected to support a thousandfold capacity increase [Qua13]. Among the different techniques that are being studied in the framework of 5G related solutions, network densification stands out as one of the most attractive alternatives to appropriately deal with the expected traffic boost [And14; Bhu14].

Network densification can be achieved by both spatial and spectrum densification, aggregating potentially non-contiguous radio frequency chunks, using the same or different technologies [Bhu14]. A cost-efficient way to accomplish spatial densification is the deployment of low-power access elements, such as small-cells or WiFi access points, which brings about a number of benefits. First and foremost, distance between users and access elements is shortened, thus increasing the quality of wireless links or reducing the required transmission power. In fact, although the LEXNET exposure index is a rather complex metric embracing many parameters, it is defined as a scaled metric of the transmit power, for both downlink and uplink components. In this sense, reducing the overall transmit power by means of densification, would be helpful to diminish the exposure received by the population.

Regarding spectrum densification, multi-connectivity with multiple RATs is attracting a lot of interest to 5G Radio Access Network (RAN) architects [Ban14], who are in the hunt for larger chunks of spectrum [Rav16]. By exploiting multi-connectivity, the amount of spectrum resources can be extended by using non-contiguous bands, such as sub-6GHz, Industrial Scientific and Medical (ISM) bands, Millimeter Wave (mmW) or TV white spaces, unified under a common OFDMA based air interface [Qua16]. Although multi-connectivity can be implemented at different layers of the stack [Rav16], MAC layer aggregation allows much finer granularity and it is a natural evolution of legacy Carrier Aggregation (CA) introduced by 3rd Generation Partnership Project (3GPP) in LTE Release 10 specification [She12].

On the downside, the load that each individual RAT needs to manage becomes highly volatile and unpredictable, as evidenced in the literature [Hua13; Wan15; Wu15]. Although macro-cells are able to compensate this volatility by aggregating multiple flows, small-cells handle fewer flows [And14], requiring the development of flexible and adaptive control techniques for radio resources. Besides, the usage of unlicensed spectrum chunks may also cause the radio resource scheduling to become substantially more complex. For instance,

unlicensed LTE, either LTE in unlicensed spectrum (LTE-U) or licensed-assisted access (LAA), operating in the 5 GHz ISM band might suffer from uncontrolled interference from external users [Che17], adding uncertainty to the amount of available radio resources.

On top of it, the definition of new architectures, such as C-RAN [Pen15], will enable the cooperation among access network elements, making the access selection decision even more complex. Several studies have already analyzed the performance of novel access techniques for both downlink, such as Cooperative Multi-Point (COMP), and uplink [Dam11] [Boc14]. Altogether, the high number of access elements and technologies will actually cause EMF exposure to be more relevant in the access selection and resource management procedures.

As we have seen in the previous chapter, system level simulation can play an important role to analyze the overall performance of access selection policies. However, when it comes to EMF exposure, the usage of abstract models does not provide the required level of realism, but approaches that accurately mimic the response of the underlaying access technologies are required.

As a response to this requirement, Sections 4.1 and 4.2 will be devoted to the definition and validation of advanced simulation models for OFDMA-based radio resources, considering both downlink and uplink connections. Typically, legacy models for system level simulation assume static behavior of the radio conditions during one snapshot, so that the effect of access selection decision over radio resources is ignored. Differently, the models proposed in this chapter consider such effect, leading to optimization problems, endowing the simulation with more realism.

Based on those models, different access selection policies will be evaluated in dense and heterogeneous scenarios. In particular, in Section 4.3 we will analyze, in terms of downlink EMF exposure, the impact of traffic offloading, involving LTE and WiFi technologies. Section 4.4 will be devoted to evaluating the effect that access selection policies have over the uplink EMF exposure. Furthermore, taking into account the aforementioned need for flexible and adaptive scheduling mechanisms, in Section 4.5 we will propose a novel scheduling framework able to fulfill the challenges brought about by 5G networks, including multi-connectivity, unpredictable traffic patterns and uncertain availability of radio resources.

4.1 Advanced modeling of downlink radio resources

The downlink performance in cellular networks mainly depends on two factors: the network geometry, and the network load. The former can be considered by applying propagation models, or, in a simplified way, by means of the interference factor, *f-factor* [Kar09]. Most of the available works that can be found in the literature put considerable effort on using appropriate propagation models [Cas07; Nas16], which accurately model the system interference and its impact over access network processes performance.

However, little attention is paid to the network load, which is usually assumed to be constant. This assumption is typically used in studies focused on the worst-case scenario (e.g. rush hour), thus providing a performance bound. On the other side, constant load, or full-buffer, does not allow considering the load fluctuation and its impact on the overall system performance, which may be more relevant in dense and heterogeneous scenarios.

In this sense, some studies, such as [Dav15], have pointed out the impact of network load

on the access selection procedures performance of LTE systems, and how it may impact the system throughput. In general, in OFDMA-based systems, the load of a cell modulates the interference induced by such cell, according to the scheduling scheme and its cooperation with other cells. In turn, the interference is used by access selection processes to decide the attachment of the UEs, which would eventually lead to a variation of the network load. In real systems, this cyclic dependence happens over time in a continuous manner, so that the feedback reported by UEs is used to consider the interference fluctuations caused by load variations. However, in snapshot-based analysis, this behavior has to be modeled by jointly considering access selection and resource management, leading to complex optimization problems that, in general, cannot be solved in a closed form [Boc11]. Nevertheless, under some assumptions the resulting problems can be simplified and solved with optimality guarantees.

In this section two different approaches to mimic the network behavior for system-level simulations are presented, considering both network geometry and load variations. The first one makes use of well known heuristics, in particular a Simulated Annealing (SA) algorithm, which solves the problem in a iterative way. The second approach assumes some simplifications, so that the underlying problem can be casted to a generalized geometric program, which can be solved as a standard convex optimization problem.

The resulting transformation can be then efficiently solved with optimality guarantees. Although the same approach has been previously used in the past for resource allocation problems [Kra11], their application scope was different and so was the corresponding problem formulation.

4.1.1 System model

We consider a single carrier LTE scenario in which there is a set of users $\mathcal{U}$ and LTE cells $\mathcal{B}$. In this scenario, cells share the same system resources, so that all base stations, but the serving one, cause interference. System resources are modeled as Physical Resource Blocks (PRBs), which are allocated to users according to their traffic demand and the particular link conditions (network geometry). We also assume that access selection has been already completed, so that users are attached to the cells from which they received the highest Reference Signal Received Power (RSRP), and we are thus aware of the corresponding user/cell associations.

We define the function β to identify the base station a user i is connected to, so that $\beta(i) \in \mathcal{B}$. Furthermore, we define the set of users connected to a particular cell k as $\mathcal{U}(k) \subseteq \mathcal{U}$. Besides, for each user-cell pair (i,k), we denote $\gamma_{i,k}$ as the power user i receives from base station k, while Γ_i is the one corresponding to the serving cell of user i, so that $\Gamma_i = \gamma_{i,\beta(i)}$. Furthermore, each user has a particular traffic demand, D_i, and its serving cell, $\beta(i)$, needs to allocate enough resources x_i to satisfy such demand.

In order to calculate the amount of resources x_i to be allocated to each user, we first need to estimate the spectral efficiency of each connection, which depends on the corresponding interference. The spectral efficiency, η, can be calculated from the SINR perceived by a user, following the Shannon-Hartley's channel capacity formula [Mog07], $\eta = \log_2(1 + \text{SINR})$.

In order to keep the formulation simple, we consider that cells do not cooperate, and use a random scheduler, so that the amount of interference power coming from a cell

is proportional to the load of such cell. It is worth noting that other cooperation and scheduling solutions would lead to different relationships between interference and load. In this sense, the applicability of this model would depend on whether or not those relationships remain linear.

As was mentioned earlier, a widespread approach, in system level evaluations, is to assume that traffic is evenly distributed within the network, which yields a Constant Load (CL) for all cells. In this case, the spectral efficiency η^{cl} for a particular user i can be calculated as:

$$\eta_i^{\mathrm{cl}} = \log_2\left(1 + \frac{\Gamma_i}{\sigma^2 + \sum\limits_{k \in \mathcal{B}/\beta(i)} \frac{L_k \gamma_{i,k}}{C}}\right) \tag{4.1}$$

where L_k holds for the load of an interfering cell $k \neq \beta(i)$, σ^2 corresponds to the system white additive noise power and C is the system capacity.

In the previous approach the resources to be assigned can be straightforwardly calculated, since all the parameters are known. On the other hand, if we strictly calculate the cell load that characterizes the particular access selection, we would need to consider the Mutual Interference (MI) among various associations. In this case, we can define the spectral efficiency η^{mi} perceived by user i as:

$$\eta_i^{\mathrm{mi}} = \log_2\left(1 + \frac{\Gamma_i}{\sigma^2 + \sum\limits_{k \in \mathcal{B}/\beta(i)}\left(\frac{\gamma_{i,k}}{C}\sum\limits_{j \in \mathcal{U}(k)} x_j\right)}\right) \tag{4.2}$$

In this case, the solution for the resource assignment problem is more complex, since the spectral efficiency perceived by user i, and so the resources to be allocated, depends on the amount of resources granted to other users, $x_j \ \forall j \neq i$.

In order to compare the two approaches, we model the corresponding resource assignment as the following optimization problem, which seeks minimizing the required number of resources, while satisfying the traffic demand for all users:

Problem 2 (Downlink resource allocation)**.**

$$\boldsymbol{min} \quad \sum_{i \in \mathcal{U}} x_i \tag{4.3}$$

$$\boldsymbol{s.t.} \quad \sum_{i \in \mathcal{U}(j)} x_i \leq C_j \qquad\qquad \forall j \in \mathcal{B} \tag{4.4}$$

$$\qquad x_i \cdot BW_{PRB} \cdot \eta_i \geq D_i \qquad\qquad \forall i \in \mathcal{U} \tag{4.5}$$

where constraint 4.4 prevents from exceeding the cell capacity, while 4.5 ensures that the traffic demand is satisfied for every user. In this last constraint the bandwidth of the physical resource, BW_{PRB}, has been added so that the demand can be defined in traffic units.

It is worth noting that, in the case of constant load, the problem can be directly solved, by calculating the resources that satisfy the user demand. However, when using the MI model, Problem 2 becomes far more complex. In the following, we present the two approaches that have been followed to solve it. The first one makes use of heuristics, in particular Simulated Annealing, to obtain the optimum solution. Although this approach is able to reach a global minimum, it is rather time consuming, and needs to be manually tweaked if the problem configuration is modified (e.g. different densities). Afterwards, we propose a modified formulation that yields a geometric programming model, which can be casted to a convex optimization problem. In this case, the problem can be efficiently solved with optimality guarantees, exploiting well known convex solvers. Afterwards, we will study the system performance when using both interference models, CL and MI, and the results obtained with the solvers for the MI one will be compared, to assess the validity of the geometric programing simplification.

4.1.2 Solvers for the downlink mutual interference problem

Simulated annealing

Simulated Annealing is a robust technique used in a wide variety of realms to solve single and multiple objective problems. It iteratively performs a random search, worsening the current solution in some cases to avoid local minima. The optimization process is based on thermodynamics and the way metal cools during the annealing process. If the metal cools slowly its atoms form a pure crystal, corresponding to the state of minimal energy of the metal (optimum), at the expense of time. On the other hand, if the metal cools quickly the resulting anneal is still at a high energy state, which may lead to getting stuck at a local minima. The reader may refer to [Sum06] for a thorough description of SA methodologies and its applications.

The particular implementation of the SA algorithm follows the generic approach depicted in [Hen03], whose implementation details are described in Algorithm 4. In brief, it contains two nested loops: the first one corresponds to the *cooling* process, which defines a new temperature, while the inner one implements the thermal equilibrium (*atoms accommodation*) for each temperature. As can be seen, the algorithm starts by selecting an initial solution x_0 with a cost $\omega(x_0)$. In general, the initial value selected when using heuristics is very important, since it can significantly reduce the convergence time. In our implementation we use the solution that would be obtained assuming constant load. In particular, the load value has been calculated by averaging the results obtained from a number of independent scenario instances where the interference was neglected. As expected, the resulting load is quite dependent on the particular scenario configuration (i.e. density).

For the cooling procedure, we use an exponential scheme, since it showed an appropriate performance for our scenarios. We start from an initial temperature $Te_0 = 500$, and $N_{Te} = 300$ iterations were used. Furthermore, the algorithm carries out $N_{it} = 400$ iterations

to look for new neighboring solutions, $Nei(x)$, within a temperature level. During this process, potential neighboring solutions $x' \in Nei(x)$ are obtained. Then, if the cost evaluated at $x', \omega(x')$, improves the previous one (i.e. $\Delta_{\omega(x),\omega(x')} < 0$), this new solution is approved. Otherwise, the neighboring solution is approved with a certain probability, to avoid local minima. The acceptance probability of worse solutions decreases with the temperature.

Algorithm 1: Simulated Annealing Algorithm

1: **procedure** SA(x_0, Te_0)
2:　　Set initial solution x_0, temperature Te_0 and iterations N_{it}
3:　　Set number of temperature levels N_{Te} and cooling counter $k = 0$
4:　　Set cooling schedule: $Te_k = Te_0 \cdot e^{-\alpha \cdot k^{\frac{1}{N_{Te}}}}$　　▷ α parameter defines the cooling pace
5:　　**while** Stopping criteria is not met **AND** $k < N_{Te}$ **do**
6:　　　　Set counter $m = 0$
7:　　　　**while** $m < N_{it}$ **do**
8:　　　　　　Generate new solution $x' = Pe_k(x) \in Nei(x)$
9:　　　　　　Calculate $\Delta_{\omega(x),\omega(x')}$
10:　　　　　　**if** $\Delta_{\omega(x),\omega(x')} \leq 0$ **then**
11:　　　　　　　　$x \leftarrow x'$
12:　　　　　　**else if** $\Delta_{\omega(x),\omega(x')} > 0$ **then**
13:　　　　　　　　$x \leftarrow x'$ with probability $e^{-\frac{\Delta_{x,x'}}{Te_k}}$
14:　　　　　　**end if**
15:　　　　**end while**
16:　　**end while**
17: **end procedure**

While the overall process could be used for different problems, the procedure used to seek new neighboring solutions heavily depends on the particular characteristics of the current problem. We have implemented it as a perturbation of the current solution, $Pe(x)$, defined as:

$$x' = x - g(x; 0, 1) \cdot \frac{\Delta_{\text{traff}}}{\eta} \tag{4.6}$$

where $g(x; 0, 1) = \frac{1}{\pi(1+x^2)}$ is the standard Cauchy distribution, η holds for the spectral efficiency of the current association, and Δ_{traff} is the difference between the traffic demand and the value carried by the current assignment. As can be seen, every new solution randomly modulates the current assignment, considering the corresponding capacity demand. This algorithm is able to solve the MI problem within a sensible time frame, but, as it can be observed, it needs a tailored configuration for each particular scenario to keep fast convergence time while avoiding local-minimum solutions.

Generalized Geometric Programming

In order to overcome the limitations of the heuristic approach, we propose a transformation of the mutual interference problem, so that it can be efficiently solved. In particular, we transform the problem to obtain the general expression of a Geometric Programing (GP) [Boy07], which is defined as follows:

Problem 3 (Standard Geometric Program)**.**

$$\textbf{\textit{min}} \quad f_0(x) \tag{4.7}$$

$$\textbf{\textit{s.t.}} \quad f_i(x) \le 1 \qquad\qquad i = 1, \dots, m \tag{4.8}$$

$$\qquad\quad g_j(x) = 1 \qquad\qquad i = 1, \dots, p \tag{4.9}$$

where f_i and g_i are generalized posynomials and monomials, respectively. Then, by applying a logarithmic change of variable and a logarithmic transformation of the objective and function constraints, the GP problem is casted to a convex one, see [Boy07] for further information regarding basic concepts of GP and its transformations.

In Problem 2, both the objective function and the capacity constraint, 4.3 and 4.4 respectively, are generalized posynomial functions [Boy04]. On the other hand, the demand constraint 4.5 is not a posynomial function when $\eta_i = \eta_i^{mi}$.

However, if we assume that $SINR \gg 1$, Eq. 4.5 can be simplified, and the original problem can be transformed into a GP one. This assumption would be sensible for highly dense networks, where the distance between users and access elements is notably reduced, so that the link connection has a high SINR.

Under such assumptions we need to apply a set of transformations so that the initial constraint 4.5 is converted to a posynomial one. First, we take the inverse of the function and apply logarithms at both sides:

$$\frac{\sigma^2 + \sum\limits_{k \in \mathcal{B}/\beta(i)} \frac{\gamma_{ik}}{C_k} \sum\limits_{j \in \mathcal{U}(k)} x_j}{\Gamma_i} \le e^{-\frac{\log 2 \cdot D_i}{x_i}} \tag{4.10}$$

As can be observed, the left-hand side of the equation has posynomial form, while the right-hand side is exponential. Then we add a slack variable t_i to remove the exponential, so that $t_i \ge e^{\frac{\log 2 \cdot D_i}{x_i}}$. This way the original constraint is replaced by two constraints:

$$\frac{\sigma^2 + \sum\limits_{k \in \mathcal{B}/\beta(i)} \frac{\gamma_{ik}}{C_k} \sum\limits_{j \in \mathcal{U}(k)} x_j}{\Gamma_i} \le \frac{1}{t_i} \tag{4.11}$$

$$t_i \ge e^{\frac{\log 2 \cdot D_i}{x_i}} \tag{4.12}$$

Finally, applying the logarithmic change of variable to both x and t (i.e. $y = \log x$ and $s = \log t$) and logarithmic transformations to the objective and constraints functions, we

obtain a convex formulation of the initial MI problem stated above, which is shown in Problem 4.

Problem 4 (Geometric form of downlink mutual interference model)**.**

$$\boldsymbol{min} \quad \log \sum_{i \in \mathcal{U}} e^{y_i} \tag{4.13}$$

$$\boldsymbol{s.t.} \quad \log \left(\frac{\sigma^2 + \sum\limits_{k \neq \beta(i)} \frac{\gamma_{ik}}{C_k} \sum\limits_{j \in \mathcal{U}(k)} e^{y_j}}{\Gamma_i} \right) + s_i \leq 0 \qquad \forall i \in \mathcal{U} \tag{4.14}$$

$$\log(2) \cdot \frac{D_i}{BW_{PRB}} \cdot e^{-y_i} - s_i \leq 0 \qquad \forall i \in \mathcal{U} \tag{4.15}$$

$$\log \sum_{i \in \mathcal{U}(j)} e^{y_i/C} \leq 0 \qquad \forall j \in \mathcal{B} \tag{4.16}$$

4.1.3 Model evaluation

In this section we discuss the evaluation of the two described models with a twofold objective: to compare the performance of CL and MI approaches; and to assess the validity of the convex approximation of the MI problem.

We have carried out such analysis, exploiting GWNSyM, under a two-tier LTE urban dense scenario. The first tier corresponds to macro enhanced NodeBs (eNodeBs) deployed following an hexagonal pattern. The second one is made of small-cells randomly located within the coverage area of the macro base-stations. For such deployment, we define clusters of cells consisting of one eNodeB and a number of small-cells. Besides, we assume a single carrier scenario where all the cells share 20 MHz (i.e. 100 PRBs) without any cooperation among them.

The analysis is carried out in the central cluster (embracing the macro and the corresponding small-cells), while the other clusters are seen as interfering cells with a constant load. Differently, the load in the cells of the central cluster, both in macro base station and small-cells, follows either CL or MI models. The propagation models are summarized in Table 4.1[1]. Over this network, we deploy 100 active users, each of them with a downlink traffic demand of 500 Kbps. Since this evaluation focuses on the performance of the different interference models, leaving aside service behavior, the simulation has been configured as static, so that each snapshot is a random independent realization of the scenario. In particular, the results presented in this section are obtained from 100 independent runs of each configuration.

As for the access selection strategy, the users are attached to the cell from which they receive higher RSRP, after applying a certain bias to the small-cell, exploiting the so called

---

[1] The scenario described in Table 4.1 will serve as reference scenario in the following sections. Variations will be indicated when necessary.

Cell Range Extension (CRE) technique[1]. Since the transmission power of the eNodeB is remarkable higher than the one from small cells (sometimes also called low power cells), only measuring the RSRP would prevent users from selecting small cells. In order to avoid that effect, users apply a bias to the RSRP measurements from the small-cells, fostering traffic offloading to the small-cells tier.

In all cases, the load of the interfering cells, as well as the cells in the central clusters when CL model is used, is set to 40%. Besides, we have carried out the analysis for CRE bias values of 3 and 9 dBs.

In order to assess the impact of considering constant load, Figure 4.1 depicts the Complementary Cumulative Distribution Function (CCDF) of the relative load in the central cluster differentiating between macro and small cells. This result is the effective load after the access selection when using both CL and MI models. In this case, we have used the SA solver for the mutual-interference model, so as to obtain a fair comparison, avoiding the assumptions made by the convex formulation.

First, we can clearly observe that the load is rather variable, far from the constant value that is assumed in the CL model, being this variability particularly higher for the small-cells. Furthermore, the effect of using different access selection strategies, different CRE bias values, is also clearly observed. According to the results, the load of the small-cells is remarkable higher when the CRE is 9 dB. Finally, although the curves obtained for the two models follow a similar trend, there is a clear difference (note that the x-axis is in logarithmic scale). We can thus conclude that the widespread constant-load model could yield inaccurate load values, and that in some cases, especially when the traffic is not fairly distributed or in highly heterogeneous scenarios, the use of the mutual interference approach shall be preferred.

After corroborating the advantages of the MI model, we have assessed the validity of the convex solution by comparing it with the one obtained by the SA solver. To solve the
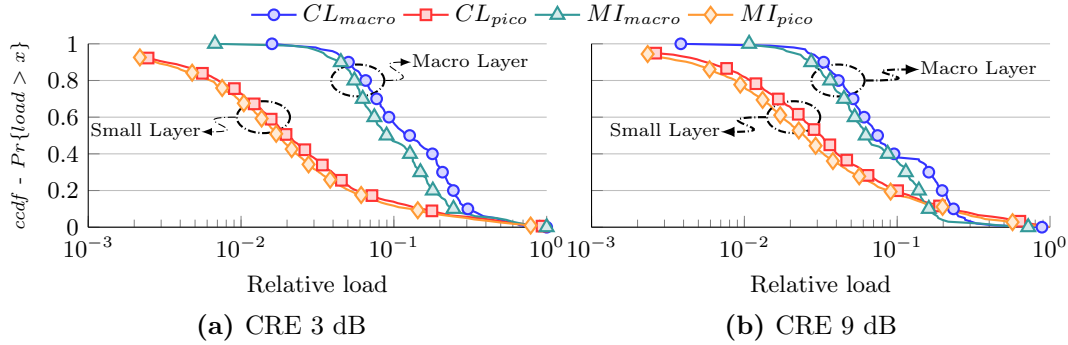
Figure 4.1: Comparison between CL and MI models. CCDF of the relative load for the two cell types

---

1 In this section, CRE is used to highlight the impact of different access selection policies on the model. Further analysis of CRE in heterogeneous scenarios will be performed in subsequent sections.

Table 4.1: Basic LTE heterogeneous scenario. Propagation models are obtained from 3GPP scenario description for system-level simulation [TR36814]

	LTE layout 20 MHz $BW_{PRB} = 180$ KHz @2.1 GHz
Macro layer	ISD[2] 500 m, 7 tri-sector sites
	Max. Tx. power 46 dBm
	Antenna gain 16 dBi, 15° down-tilt
	Random Location, 12 cells per cluster
Pico layer	Max. Tx. power 37 dBm
	Omni-antenna, 5.0 dBi
UE	DL NF 7 dB, Rx. Gain 0 dB
LTE layer	Pathloss (L in dB) as a function of the distance (d in m)[3]
Macro$_\text{NLOS}$	$139.1033 + 39.0864 \cdot (\log_{10} d - 3)$
Macro$_\text{LOS}$	$36.2995 + 22 \cdot \log_{10} d$ if $d < 328.42$
	$40 \cdot \log_{10} d - 10.7953$ if $d > 328.42$
Small$_\text{NLOS}$	$145.48 + 37.5 \cdot (\log_{10} d - 3)$
Small$_\text{LOS}$	$103.8 + 20.9 \cdot (\log_{10} d - 3)$
	LOS probability as a function of the distance d (m)
Macro	$P_\text{LOS} = \min\left(\frac{18}{d}, 1\right) \cdot \left(1 - e^{\frac{-d}{36}}\right) + e^{\frac{-d}{36}}$
Small	$P_\text{LOS} = 0.5 - \min\left(0.5, 5 \cdot e^{\frac{-156}{d}}\right) + \min\left(0.5, 5 \cdot e^{\frac{-d}{30}}\right)$

[2] Inter site distance defines the distance between eNodeBs   [3] Pathloss is defined for both line-of-sight (LOS) and non-line-of-sight (NLOS) conditions

convex problem, we have integrated GPKIT library [Bur15] within GWNSyM.

In order to assess whether the assumptions and simplifications that were taken are valid, Figure 4.2 compares the results that were obtained for both the SA and GP solutions for different loads of the interfering cells (i.e. surrounding clusters). We represent the Cumulative Distribution Function (CDF) of the relative difference of the resources allocated for the same scenario instance. Hence, we define $\Delta_\text{load} = \dfrac{\left|\sum_{i \in \mathcal{U}} x_i^\text{sa} - \sum_{i \in \mathcal{U}} x_i^\text{gp}\right|}{\sum_{i \in \mathcal{U}} x_i^\text{sa}}$, where x_i^{sa} and x_i^{gp} are the resource allocated by cell i for the Simulated Annealing and Geometric Programing solutions, respectively. The results yield that the GP solution exhibits an appropriate performance, especially when the interference induced by the surrounding cells is low. This was expected, since higher interferences would lead to smaller SINR, and the simplification that was taken ($SINR \gg 1$) would be less precise. If coordination and cooperation techniques were used to mitigate the interference, GP solutions would be even more accurate.

In any case, for a relatively high load (40%), the difference stays below 10^{-2} for almost
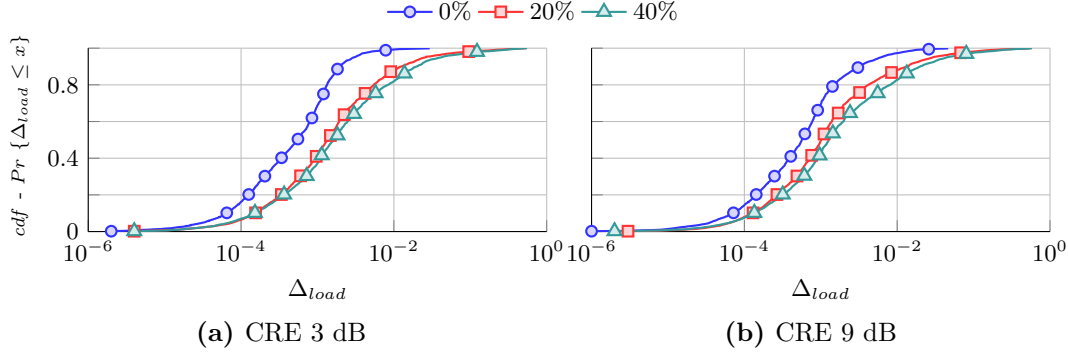
Figure 4.2: CDF of resource assignment difference with both SA and GP for various interfering cells loads

90% of the cases. We can also see that when CRE is higher, there is a smaller difference between the Δ_{load} for the three scenarios. The reason is that the load of small cells would increase, leading to a slightly higher interference.

4.2 Advanced modeling of uplink radio resources

Due to the reduced capabilities of user terminals, as compared to those of the access elements, LTE uplink makes use of Single Carrier Frequency Division Multiple Access (SC-FDMA) for multiplexing, instead of the OFDMA scheme used for downlink. In both cases, however, the resources allocated to users can be considered orthogonal, so that they can be modeled in a similar way.

This orthogonality ensures that, in principle, there is not any interference between users attached to the same cell, but they only experience inter-cell interference [Sim08]. Similarly to the downlink case studied in Section 4.1, the interference level depends on the users position and pathloss (i.e. network geometry) as well as on interfering cells load.

Differently to the downlink, however, the uplink resource allocation depends on the power control scheme implemented in the UEs. In the downlink case, under homogeneous power allocation assumption, users are granted resources until traffic demand is satisfied, so that the SINR is a consequence of the network resource allocation, as was discussed in Section 4.1. Conversely, in the communication from users to base stations, resource allocation is performed considering a target SINR [Sim08], which depends on the network configuration, and constrained to a maximum transmit power.

Taking these differences into account, the amount of resources that need to be allocated to a particular user, can be calculated based on the target SINR. On the other hand, the power assigned to such resources, dictated by the power control policy, must be assigned to reach such target.

In general, the power allocated by a user will affect the interference experienced by others, and, in turn, the power required by the latter to reach their target SINR. In real systems, power is adapted over time using a closed loop power control [Sim08], being adjusted based on the radio conditions experienced by the user. However, in system level evaluations this iterative power control needs to be modeled, so that the power allocated by one user

considers that allocated by interfering ones. Usually, interference is considered as a static factor added to the noise [Liu14] (i.e. interference over thermal noise). Nevertheless, this approach, while simple, does not account for the network geometry nor the effect that particular access selection policies would have over the resulting interference, and in turn, over the transmission power.

As we can see, there exist some parallelism between the requirements to model the uplink power control, and the resource allocation model presented in Section 4.1, due to the existence of mutual interference. Nevertheless, in this case, the underlying problem can be posed as a linear one, and it can be therefore efficiently solved without further modifications.

In the following we will provide a short description of the uplink power control performed by the UEs. Afterwards, we will describe a power control model for LTE uplink communications that can be applied to different access selection policies. Finally, the proposed model will be compared with a static interference solution in order to assess the differences yielded by both approaches.

4.2.1 Uplink power control

In general, we can define the power to be transmitted by a UE as follows [Ber14]:

$$P = \min\{P_{tx_{\max}}, P_0 \cdot N_{RB} \cdot L^\alpha \cdot \delta_{\mathrm{mcs}} \cdot f(\Delta)\} \tag{4.17}$$

where $P_{tx_{\max}}$ holds for the maximum power that the UE is allowed to transmit, N_{RB} is the number of resources allocated to the user, L is the pathloss between the UE and its serving access element, and α is the pathloss compensation factor. The parameter δ_{mcs} indicates a power offset depending on the Modulation and Coding Scheme (MCS) used, and it is specific of the UE. Besides, $f(\Delta)$ is the closed loop correction function, which enables adjusting the transmission power over time.

Parameters P_0 and α are indicated by the network to the users in a broadcast manner. Furthermore, pathloss is computed by the UE based on the RSRP. By using these three variables, open loop parameters, the UE can set the initial transmission power without any indication from the serving cell.

Furthermore, δ_{mcs} and $f(\Delta)$ factors are used to compensate the transmit power over time (i.e. closed loop) so that the actual transmit power oscillates around the value obtained with the open loop parameters. The reader may refer to [Muh10] for a thorough explanation of how the closed loop mechanism is performed.

Assuming that the UE evenly distributes the power between allocated resources, the power transmitted perPRB can be defined as :

$$P = \min\{P_{tx_{\max}}, P_0 \cdot L^\alpha \cdot \chi\} \tag{4.18}$$

where $\chi = \delta_{\mathrm{mcs}} \cdot f(\Delta)$ accounts for the closed loop factors.

4.2.2 System model

We consider a scenario similar to that described in Section 4.1, comprising a set of users $\mathcal{U}$ and a set of LTE cells $\mathcal{B}$. Without loss of generality, we assume that cells share the same spectral resources (i.e. single carrier scenario). Hence, all users, but those attached to the same cell, induce interference. We also assume that resources are modeled as PRBs and the access selection has been already applied, so that the users attachment is known. We will not make any assumption on the specific access selection policy.

We will follow almost the same nomenclature as in the previous section. Function $\beta(i) \in \mathcal{B}$ indicates the serving cell of user i. Likewise, $\mathcal{U}(k) \subseteq \mathcal{U}$ holds for the set of users attached to the cell k. Furthermore, we define γ_{ik} as the pathloss between user i and cell k, while Γ_i is the one corresponding to the serving cell of user i, $\Gamma_i = \gamma_{i,\beta(i)}$. Finally, χ_i accounts for the closed loop power factors of user i.

We then define the interference per PRB experienced by user i, h_i, as follows

$$h_i = \sum_{k \in \mathcal{B}/\beta(i)} \sum_{j \in \mathcal{U}(k)} \frac{N_{RB_j}}{C_k} P_0 \frac{\gamma_{jk}^\alpha}{\gamma_{j\beta(i)}} \chi_j \tag{4.19}$$

where N_{RB_j} is the amount of resources allocated to the interfering user j, and C_k holds for the total capacity of cell k. As can be observed, the interference is defined as the normalized summation of the resources allocated to interfering users ($j \in \mathcal{U}(k)$), weighted by the received interference power at the serving cell, $\beta(i)$.

This way, the power transmitted by user i, P_i, can be defined in terms of the target SINR, St_i, and the interference, so that $St_i \leq \frac{P_i/\Gamma_i}{\sigma^2 + h_i}$, where $P_i = P_0 \Gamma_i^\alpha \chi_i$. Then, if we group all the dependent terms, we can express the SINR of one user as a lineal combination of the power transmitted by all the remaining users:

$$P_i - St_i \sum_{k \in \mathcal{B}/\beta(i)} \sum_{j \in \mathcal{U}(k)} \frac{N_{RB_j} P_j \Gamma_i}{C_k \gamma_{j\beta(i)}} \geq St_i \Gamma_i \sigma^2 \tag{4.20}$$

As we can see, the power that each user needs to transmit linearly depends on the power transmitted by the interfering users. If we assume that UEs transmit the minimum power to ensure that the target SINR is reached, the overall uplink transmission power in the system can be modeled as a linear optimization program, as shown in Problem 5.

Problem 5 (Uplink power allocation)**.**

$$\boldsymbol{min} \quad \sum_{i \in \mathcal{U}} P_i \tag{4.21}$$

$$\boldsymbol{s.t.} \quad P_i - St_i \sum_{k \in \mathcal{B}/\beta(i)} \sum_{j \in \mathcal{U}(k)} \frac{N_{RB_j} P_j \Gamma_i}{C_k \gamma_{j\beta(i)}} \geq St_i \Gamma_i \sigma^2 \qquad \forall i \in \mathcal{U} \tag{4.22}$$

$$P_i \geq 0 \qquad \forall i \in \mathcal{U} \tag{4.23}$$

In this sense, the mutual interference model represented by Problem 5 obtains the overall transmission power, both open and close loop components. It considers a particular instance of the scenario, where the topology is defined by the pathloss coefficients and the traffic by the number of allocated PRBs.

4.2.3 Model evaluation

The mutual interference model for uplink transmission power has been evaluated in the scenario described in Table 4.1. In this case, however, we have only focused on the central cluster, assuming that interference from users in adjacent clusters is negligible.

It is assumed that users select the cell from which they receive the highest RSRP. Furthermore, the target SINR is set to 5 dB (values between 0 and 30 dB are assumed in the literature [Sim08]) and each user has an uplink traffic demand of 1 Mpbs, so that 2.1492 PRBs are allocated per user[1].

Figure 4.3 depicts the CDF of the uplink transmission power when using the mutual interference model defined by Problem 5, and when applying a simple constant interference model with the same average value. The constant load value has been obtained by averaging the output of the mutual interference model, so that in both cases the average load is the same.

Furthermore, the scenario has been analyzed under different user densities and small cells. It is worth highlighting that the transmission power is not low bounded, and in some cases it takes quite low values, $1e - 7$ W, which are typical values of P_0 [Mar16].

This evaluation draws some interesting results. In general, we observe that both models, mutual and constant interference, present rather different behaviors. Indeed, the use of constant interference model leads to incrementing the uplink transmission power by a factor between 5 and 10. This is due to the assumption that the interference evenly affects all users.

Let us consider that a user is close to the serving cell, but having moderate (or even high) interference. In this case, due to the proximity to the cell, a small increment of the transmission power may be enough to reach the target SINR. On the other hand, for users close to the cell edge, small upturns of the interference will cause similar increases of the transmission power. Hence, the assumption of constant interference may have a high impact in those users close to the cell edge. Furthermore, this effect can become more relevant in heterogeneous scenarios, due to the uneven coverage patterns of cells. In this sense, users attached to small cells can experience high interference due to nearby users connected to macro cells, requiring higher transmission power.

Besides the general trend, the results depicted in Figure 4.3 also indicate that the statistical behavior of the mutual interference model is more robust upon the variation of the network load. For any density of small cells, the mutual interference model hardly varies with the system load (only in the case with 60 small cells), represented by the number of users. On the other hand, for the scenario under analysis, the constant model shows a remarkable variation when we increase the number of users.

---

1   This can be easily calculated by using the Shannon-Hartley's channel capacity formula considering that the bandwidth of a PRB is 180 KHz : $N_{RB} = \frac{1}{0.18 \cdot log_2 (1+\text{SINR})}$.
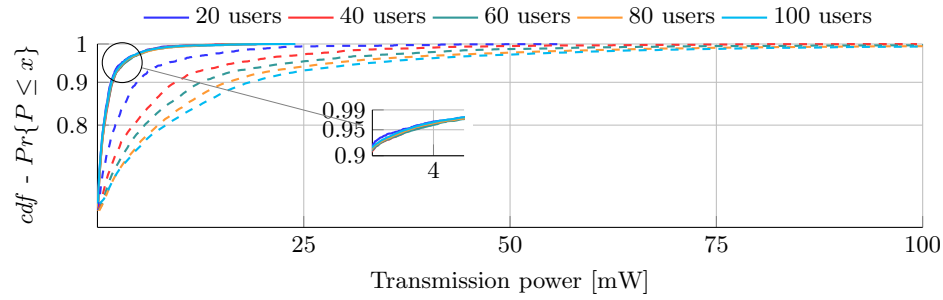
(a) 20 small cells
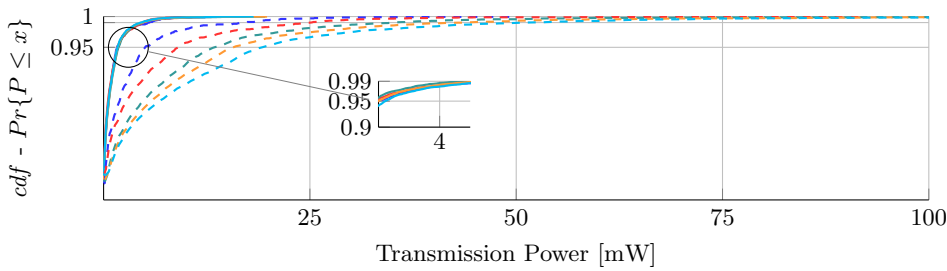
(b) 40 small cells
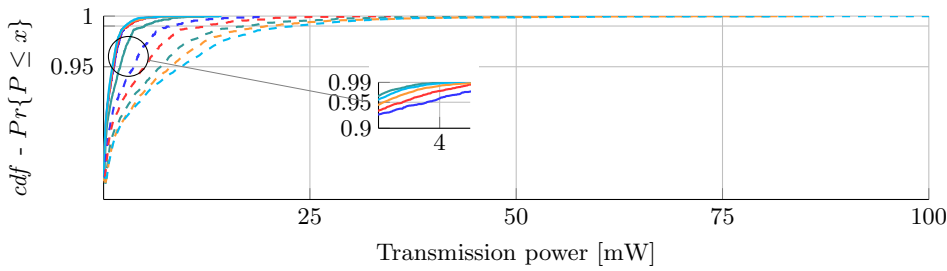
(c) 60 small cells

Figure 4.3: CDF of the uplink transmission power per PRB using both mutual and constant interference models with the same average values and different system loads. Solid and dashed lines indicate the results for mutual and constant interference models respectively

Finally, we can also observe a slight power reduction as the density of small-cells increases. Additionally, if we focus on Figure 4.3c, we can perceive that the mutual interference model has some dependency with the system load, when the density of small cells is high.

4.3 LTE-WiFi offloading scenario

Based on the radio resource model proposed in Section 4.1, we analyze here the use of non-cellular technologies as a means to tackle the traffic increase in wireless networks. In particular, we will study a scenario where WiFi access points are deployed to offload traffic from the LTE cellular deployment. Traffic offloading to WiFi is governed by the Access Network Discovery and Selection Function (ANSDF) block, where operator offloading

policies are defined. In this sense, upon specific configuration, network operators select the set of services that shall be, when possible, transferred to WiFi, so as to optimize the core network efficiency in terms of delay, capacity and cost. This technique is indeed viewed as a very cost-effective solution to support traffic increases by densification, with limited investment in the infrastructure.

The study presented in this section was conducted in a collaborative way, together with other partners within the LEXNET project, where different and complementary evaluation techniques over the same scenario were applied. The overall conclusions were presented in two publications [Bra16; Die15d], where the results obtained with GWNSyM complemented those provided by a radio-planning industrial tool [Vol17].

The evaluation presented here shows the best performance that may be obtained when the access selection policy tries to minimize the EMF exposure. It makes use of the power transmitted by the access elements as an indicator of the population exposure, and it is thus focused on the downlink components of the EI.

In the following, we will first introduce the system models implemented within GWNSyM. Based on those models, an optimization problem to minimize the power transmitted by the network will be posed, embracing both LTE and WiFi layers, so that we can obtain the bound for a EMF aware access policy. Later on, the scenario will be evaluated based on the simulation results.

4.3.1 System model

Both LTE and WiFi technologies have been implemented within GWNSyM, paying special attention to the computation of downlink transmit power as a proxy indicator of the EMF exposure. Nevertheless, differences exist between both technologies concerning the way radio resources are consumed. Accordingly, we have modeled them from different perspectives. In the case of LTE we propose a model that manages the resources on the eNodeBs, while the WiFi model is based on a statistical function of the distance. In both cases we are able to estimate the transmission power of each network element, so that it satisfies the users traffic demand, according to the access selection criteria and policies.

LTE model

The LTE model is similar to the mutual interference one described in Section 4.1. The resources of the eNodeBs are allocated to users according to their traffic demand and the spectral efficiency of the link between the cell and UE.

In order to endow the study with more realism, we have slightly modified the Shannon-Hartley's channel capacity formula, so as to consider specific LTE features, as proposed in [Mog07]. In particular, the SINR and spectral efficiency are modulated by constant parameters obtained from on-the-field measurements over LTE networks, resulting in the following formula:

$$\eta_{ij} = \mathrm{BW_{eff}} \log_2 \left[1 + \frac{\mathrm{SINR}}{\mathrm{SINR_{eff}}} \right] \tag{4.24}$$

where $\mathrm{BW_{eff}}$ and $\mathrm{SINR_{eff}}$ are correction factors that account for the system overhead and

the adaptive modulation and coding, respectively. It is worth noting that the correction factors do not alter the properties of the optimization problem posed for the mutual interference model.

Furthermore, we assume that the transmit power of a LTE cell j, with capacity C_j, is evenly distributed among its PRBs, being P_{RB_j} the power allocated by cell j to each physical resource. All in all, the power transmitted by LTE cell j for its connection with user i is calculated as follows

$$P_{LTE_{ij}} = N_{RB_i} \cdot P_{RB_j} \tag{4.25}$$

where N_{RB_i} is the number of resource blocks allocated to the corresponding connection, and obtained by the mutual interference model introduced for Problem 2, given a specific demand[1].

WiFi model

Unlike LTE, WiFi technology does not simultaneously allocate resources to the different users, but they contend to use the shared wireless channel. In this case, we adopt a statistical model, based on the distance between the user and the WiFi AP. In particular, we estimate the raw data rate, DR, that a user might obtain from an AP, according to the particular distance between them and the probability of having a line-of-sight link. According to the study presented in [Col13], the data rate of 802.11g, in outdoor scenarios, can be mathematically modeled as follows:

$$DR_{NLOS}(p_1, p_2, d) \quad = DR_{\max} \cdot \left[1 - \frac{1}{1 + e^{-p_1(d-p_2)}} \right] \tag{4.26}$$

$$DR_{LOS}(p_1, p_2, d) \quad = \begin{cases} A & d < p_1 \\ k \cdot \left(\frac{1}{d} - \frac{1}{p_2} \right) & p_1 \le d < p_2 \\ 0 & d \ge p_2 \end{cases} \tag{4.27}$$

where $DR_{\max}$ is the maximum data rate supported by the AP (54 Mbps for 802.11g), d is the distance between the user and the AP, and both p_1 and p_2 are parameters defined according to the scenario.

In this case, we define the time that a user is using the shared channel as the basic resource unit. Let DR_{ij} be the maximum data rate at which user i could be connected to AP j, depending on the distance between them, and D_i the traffic demand of such user. Then, the fraction of time that the corresponding connection (when the AP is not available for other users) will be active can be calculated as $T_{ij} = \frac{D_i}{DR_{ij}}$. Finally, the power transmitted by the AP can be estimated as shown in Eq. 4.28, where $Pmax_j$ represents

---

1 The change of variables is done to simplify the access selection policy notation that will be presented below.

the maximum transmit power of the access point (typically 20 dBm).

$$P_{WiFi_{ij}} = P_{\max_j} \frac{D_i}{DR_{ij}} \tag{4.28}$$

4.3.2 Access selection optimization problem

In the following we will use the same notation as in Section 4.1, but including some new parameters. Let $\mathcal{U}$, $\mathcal{B}$, $\mathcal{W}$ be the set of users, LTE base stations and WiFi access points respectively. We also define $\mathcal{A}$ as the set of all access elements in the network, so that $\mathcal{A} = \mathcal{B} \cup \mathcal{W}$.

Furthermore, we define a binary decision variable, $x_{ij} \in \{0, 1\}$, to reflect the connection between user $i \in \mathcal{U}$ and access element $j \in \mathcal{A}$. This variable would equal 1 if user i is connected to access element j and 0 otherwise. Next, the transmission power for an association between a user i and access element j, $P_{tx_{ij}}$, is defined as follows:

$$P_{tx_{ij}} = \begin{cases} P_{\text{LTE}_{ij}} & \text{if } j \in \mathcal{B} \\ P_{\text{WiFi}_{ij}} & \text{if } j \in \mathcal{W} \end{cases} \tag{4.29}$$

All in all, the access selection policy to reduce the overall EMF exposure can be defined as follows:

Problem 6 (Offloading access selection)**.**

$$\boldsymbol{min} \quad \sum_{i \in \mathcal{U}, j \in A} x_{ij} \cdot P_{tx_{ij}} \tag{4.30}$$

$$\boldsymbol{s.t.} \quad \sum_{j \in \mathcal{A}} x_{ij} = 1 \qquad\qquad \forall\ i \in \mathcal{U} \tag{4.31}$$

$$\sum_{i \in \mathcal{U}} x_{ij} \cdot N_{RB_{ij}} \leq C_j \qquad\qquad \forall\ j \in \mathcal{B} \tag{4.32}$$

$$\sum_{i \in \mathcal{U}} x_{ij} \cdot T_{ij} \leq 1 \qquad\qquad \forall\ j \in \mathcal{W} \tag{4.33}$$

$$\sum_{j \in \mathcal{B}} x_{ij} \cdot N_{RBij} \cdot BW_{PRB} \cdot \eta_{ij} \geq D_i \qquad\qquad \forall i \in \mathcal{U} \tag{4.34}$$

$$\sum_{j \in \mathcal{W}} x_{ij} \cdot DR_{ij} \geq D_i \qquad\qquad \forall i \in \mathcal{U} \tag{4.35}$$

$$\tag{4.36}$$

where constraints were added to ensure sensible behavior and that system capacities are not exceeded. First, we ensure that each user can be only connected to one access element in Eq. 4.31. Besides, we also guarantee that the amount of resources does not exceed the capacity of the access elements in Eqs. 4.32 and 4.33 for LTE and WiFi, respectively. Finally, constraints 4.34 and 4.35 are included so as to enforce that users' demands are satisfied.

All in all, the access selection problem is a linear binary optimization problem, which can be solved with optimization toolkits such as GLPK [Gnu17]. Furthermore, the underlaying LTE model requires solving an additional convex problem for each possible solution of the linear one, leading to a rather complex evaluation system.

4.3.3 Scenario evaluation

For the LTE network deployment, we consider a macro hexagonal deployment, similar to the one described in Table 4.1. Nevertheless, the configuration parameters were tuned to make the setup compatible and comparable to the other tools that complemented the results obtained with GWNSyM, as mentioned at the beginning of this section.

Similar to previous evaluations, each eNodeBs has 3 sectors (cells) with bandwidth of 20 MHz, corresponding to 100 physical resource blocks of 180 KHz each. Besides, we consider a maximum transmission power of 43 dBm, a transmission gain of 18 dBi and a noise figure of 6 dB. The antenna pattern and propagation models are the same as in Table 4.1 for macro base stations. Furthermore, we consider a lower bound for the SINR of −10 dB, so that below this value, we assume the signal to be too low, and therefore the cell is deleted from the candidate list. Last, the maximum spectral efficiency has been fixed to 5 bps/Hz. The deployment of the base stations follows an hexagonal pattern where macro base stations are 450 m away from each other, covering a surface of 1350×1350 m^2.

Concerning the WiFi layout, the maximum data rate DR_{ij} is calculated by a step-wise function according to the distance model previously described, and maximum transmission power of 20 dBm is assumed. The complete mapping of the data rate for this analysis is shown in Table 4.2. Furthermore, we scale it by 0.5, accounting for the overhead of the lower layers. The number of the APs deployed in the scenario is modified to analyze the system behavior under different densities.

Over this network topology, we consider different types of users and services, as shown in Table 4.3. Each of the services follows an ON-OFF model, while users move according to a random waypoint model with a speed that is randomly selected within [1,3] m/s. It is worth highlighting that, following the service model described in Section 3.2.1, those services interrupted due to lack of connectivity, are not considered in the following instances of the problem, until they start again.

Additionally, a set of rules were imposed to users according to their profiles. These rules can be included in Problem 6 by simply adding additional constraints over the set of decision variables x_{ij}.

First, we assume that users of type 3 are not allowed to connect to the WiFi network. On the other hand, the rest of users would first attempt to connect to WiFi, and, only if it is not possible, they would connect to a LTE cell. For this scenario, we have analyzed 360 snapshots, generated every 10 seconds, according to the aforementioned models, and thus the overall analysis lasts for 1 hour.

Considering the configuration parameters, the evaluation methodology can be greatly simplified. First, assuming an arbitrary user demand D and considering the practical bounds to the LTE spectral efficiency, $\eta = 5$ bps/Hz, the minimum amount of resources (and so transmit power) required by an LTE connection can be calculated as follows:

Table 4.2: Mapping between distance and data rate applied to the WiFi model, The table includes setup values for both LOS and NLOS situations

	Distance (m)	Data rate (Mbps)
LOS	42.5	54.0
	94.0	48.0
	116.0	36.0
	139.5	24.0
	162.5	18.0
	184.0	12.0
	202.5	9.0
	214.5	6.0
	232.5	5.5
	254.0	2.0
	275.5	1.0
	Distance (m)	**Data rate (Mbps)**
NLOS	18.0	54.0
	45.5	48.0
	60.5	36.0
	70.0	24.0
	77.0	18.0
	83.0	12.0
	88.5	9.0
	92.0	6.0
	99.5	5.5
	111.5	2.0
	122.5	1.0

Table 4.3: Parameters of users and services

	Type	Traffic(Kbps)	ON (s)	OFF (s)
Services	Video (1)	1000	300	1200
	Internet (2)	128	60	80
	Enterprise (3)	128	90	120
	Type	**Services**	**Amount (%)**	**Access**
Users	1	1 + 2	50	All networks
	2	2 + 3	25	All networks
	3	2	25	Cellular only

$$\eta \cdot BW_{RB}[KHz]N_{PRB} = D[Kbps] \rightarrow \quad (4.37)$$

$$N_{PRB} = \frac{D[Kbps]}{900} \quad (4.38)$$

where BW_{RB} = 180 KHz holds for the bandwidth of the physical resources. Hence, the minimum power for an LTE connection is $P_{LTE_{min}}[mW] = \frac{D[Kbps]}{4.50}$. Then, if we compare this expression with the one modeling WiFi, taking into account maximum transmission power and scale factor, the WiFi power is always below the one required for LTE connections for the same traffic demand. Bearing this in mind, the access selection strategy boils down to selecting WiFi when possible, and LTE otherwise. Once the access selection is accomplished, the mutual interference model for LTE is applied, using the simulated annealing solver, since we are not considering a dense setup.

Figure 4.4 shows the statistical distribution of the power transmitted by the whole network. As can be observed, the density of APs does not have a strong impact on the overall power, at least for the defined access policies. Nevertheless, this metric considers the power of the connected services, but it does not show the impact on the QoS.

In order to assess the influence over QoS parameters, Figure 4.5 represents the percentage of services that are rejected (not initiated) and dropped for the various configurations. The results show that, as long as the number of WiFi APs increases, the probability of having unsuccessful services decreases. It is important to point out that the main reason to reject services is the low SINR that users might suffer if they are at the cell edge. In this case the influence of WiFi density becomes evident.

Another aspect worth highlighting is the different behavior observed for the various services and, in particular, *Internet* ones. A relevant percentage of them (70%) are not allowed, according to the policy described in Table 4.3, to use the WiFi network and, as a consequence, the number of interrupted services is less affected by the APs density. However, for the other types of services, the figure yields a clear benefit of increasing the WiFi density.

Finally, we have also assessed the appropriate operation of the access policy. In this sense, Figure 4.6 shows the percentage of time that different services are connected to each
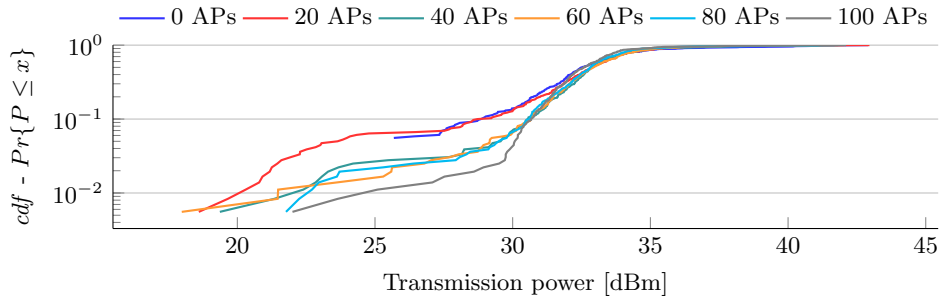
Figure 4.4: CDF of the total transmitted power by the network

(a) Video

(b) Enterprise

(c) Internet
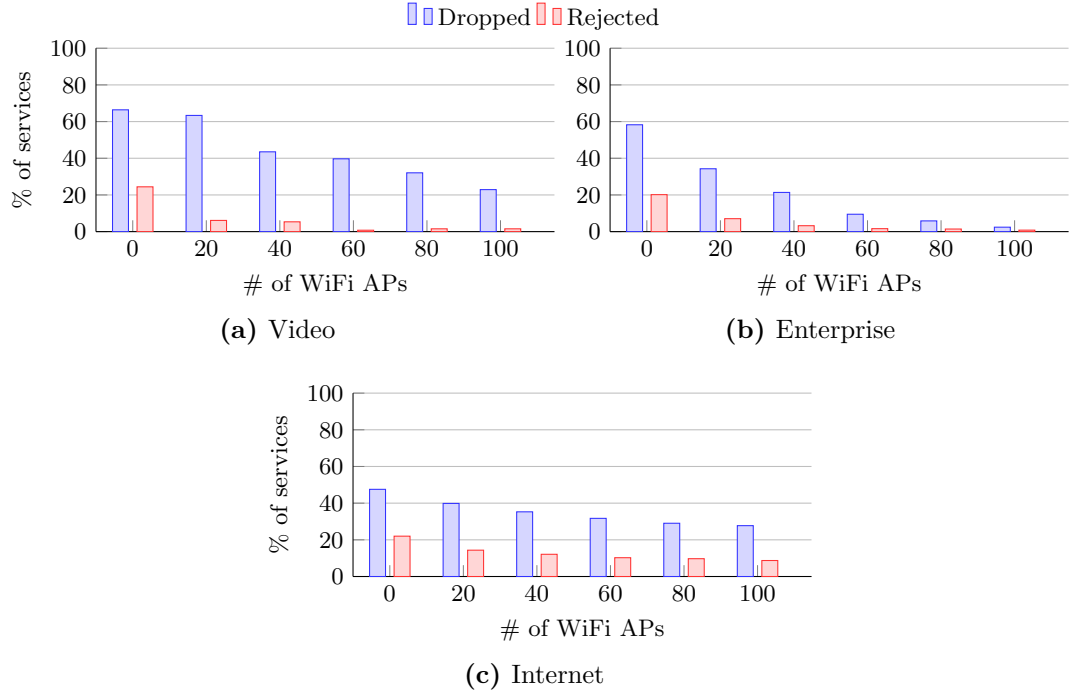
Figure 4.5: Percentage of services that are either dropped or rejected against different WiFi access points densities. The metric is shown for the different types of services

of the two available technologies, including as well services that were dropped at some point during the session. For the *Video* and *Enterprise* services, the figure shows that they connect to the WiFi network as soon as APs become available. In fact, for the scenario embracing 40 APs, more than 80% of these services connect to WiFi. On the other hand, we can see that the defined access policy is able to modulate the behavior of the *Internet* services. Although the usage of WiFi increases as long as more APs are deployed, there is still a relevant number of services using a LTE connection. This behavior is due to the constraint imposed to type 3 users to keep the LTE connections. This leads to a worse QoS for this service, which suffers from a larger reject probability, as was depicted in Figure 4.5.
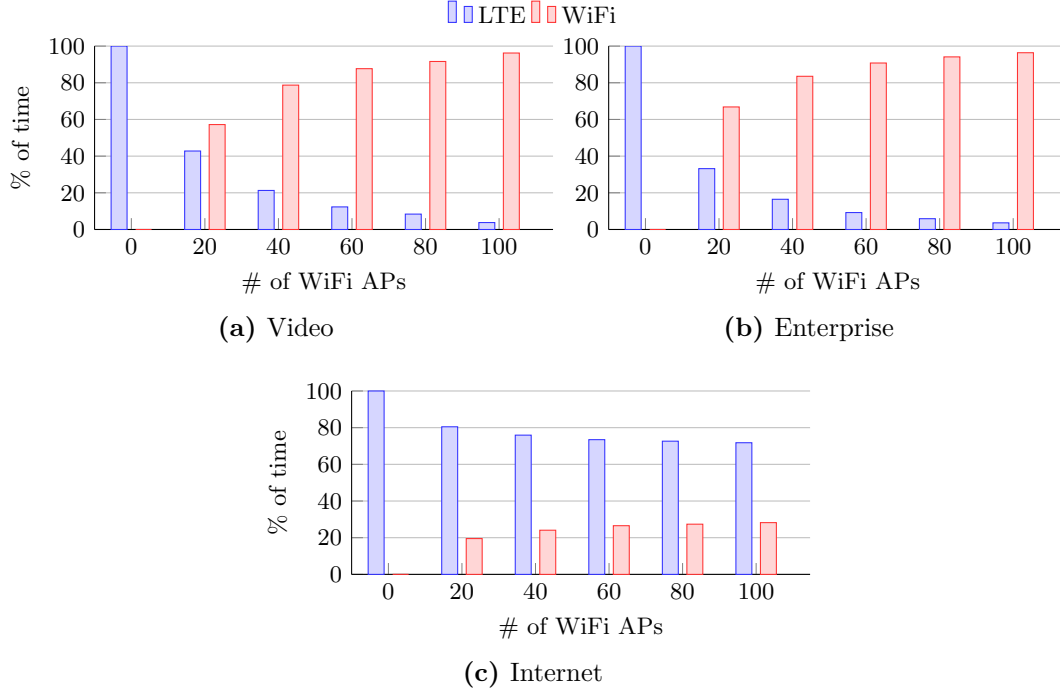
Figure 4.6: Percentage of time that services stay in each technology against different WiFi access points densities. The metric is shown for the different types of services

4.4 Uplink access selection in heterogeneous scenarios

Traditionally users have been mostly consumers of content from the Internet, leading to unbalanced traffic loads between downlink and uplink communications. For this reason, most of the works until now have mainly focused on the communication from the access elements to users, while little attention has been paid to the uplink.

However, the proliferation of new services and applications, such as video streaming or gaming applications, cause users to create more and more content, and it becomes relevant to look at the communication from user terminals to the base station, in terms of resource management.

In cellular networks, a user terminal connects, both for uplink and downlink, to the base station from which it receives the highest signal level. In the case of LTE, this means the highest RSRP. Nevertheless, in heterogeneous environments, where base stations with high (macro-cells) and low (small-cells) transmission power co-exist, the access selection based on RSRP makes users to select, in most cases, the access elements with highest transmission power. It is worth highlighting that the received power from macro-cells can be much higher than from small-cells, even when the small-cell it closer to the user terminal. In this sense, the potential capacity improvement brought by small-cells is not fully exploited. On top of it, RSRP policy only accounts for the signal quality from the base station to the user terminal, so that the transmission power in the uplink can be far from being optimum.

In order to establish access selection policies able to take advantage of heterogeneous scenarios, CRE [Dam11] techniques can be exploited, so that the reference power coming from small-cells is virtually incremented by applying a bias factor. This way, access selection using CRE favors connection to elements with lower transmission power, thus improving load balancing.

In both cases, RSRP and CRE approaches, the access selection policy assumes a traditional network architecture in which each base station is operated in a standalone way. However, new architectures that exploit NFV techniques, such as Cloud Radio Access Network (C-RAN), will enable centralizing some network functionalities, allowing tighter coordination between network elements. Following this trend, the separate access selection of downlink and uplink, DUDe [Boc16], has been recently proposed, so that the usage of radio resources can be optimized.

In the following, exploiting the uplink power model proposed in Section 4.2, different access selection strategies will be analyzed with GWNSyM, in terms of both transmission power and QoS.

4.4.1 Scenario evaluation

The scenario analyzed in this section is again the one described in Table 4.1, over which the aforementioned access selection policies are applied, namely RSRP, DUDe and CRE. For the latter, three different values of the bias are applied: 6 dB (CRE_1), 9 dB (CRE_2), and 12 dB (CRE_3). Furthermore, the number of small-cells is modified, so as to study the impact of the network density.

Over this scenario, 100 users are deployed with the same requirements as in Section 4.2. This is, UEs are configured with a target SINR of 5 dB, and each user runs a service with uplink traffic demand of 1 Mbps. Users move according to a random waypoint pattern with speed uniformly distributed in the range $[1, 3]$ m/s. Besides, services follow an ON/OFF model with average session time of 180 seconds and average time between sessions (i.e. time elapsed since a session finishes until a new one starts) of 60 seconds.

Besides, it is assumed that signaling from the serving cell to the UE is always active, regardless the access selection policy. For that reason, after selecting the serving cell, we check that the downlink SINR is above 7.5 dB. Although this approach may not be very realistic, since ensuring downlink control communications would be prioritized in real systems, this setup allows us to fairly compare the different access selection policies. Besides, since in this scenario we focus on the uplink performance, for the sake of simplicity, the downlink SINR is computed assuming full load of the access elements.

Figure 4.7 depicts the average transmission power allocated to each resource block, when applying the different access selection policies and for a number of network densities. In general, as expected, we observe an overall reduction of the transmission power as the number of small-cells increases. However, this trend strongly depends on the access selection scheme. While in the case of DUDe we observe a ten-fold reduction, for RSRP-based access selection it is around 5.

In this sense, we observe quite remarkable differences of transmission power when comparing the different strategies for a given network density. In general, RSRP is the approach that needs more power, which seems sensible since it favors connections with

macro cells. Furthermore, we observe that CRE, when configured with high bias values, requires less power than DUDe. This behavior is unexpected, since DUDe was precisely designed to achieve the optimal behavior. In order to better understand the reason behind it, we need to analyze how the different strategies impact the QoS.

In this sense, we have studied the probability that, in a given snapshot, a user can effectively perform the connection according to the access selection strategy. It is worth noting that the outage probability is only constrained by the downlink received signal (i.e. downlink SINR).

The results shown in Figure 4.8 complement those seen before in Figure 4.7. Indeed, we can observe that the outage probability decreases as the density of small-cells gets higher, thus reducing the average transmission power. Additionally, the results yield a strong dependency of the outage probability with the access selection strategy. In this sense, RSRP and DUDe always connect to the desired cell, showing the very same behavior. Indeed, it was expected that DUDe would behave exactly as RSRP, since they use the same policy for downlink attachment (i.e. RSRP).

As for CRE, it can be observed that high bias values (e.g. CRE_3) lead to poor performance in terms of outage probability. In fact, this performance degradation explains the power reduction observed in Figure 4.7. When using high bias values, users relatively far from the small-cells can eventually get connected to them, even if it causes downlink outage. This way, only users close to small-cells correctly perform the connection, which explains the rather low uplink transmission power. Figure 4.9 illustrates the service performance statistics for the different configurations, including the ratio of rejected and dropped services[1].

As can be observed, the results are quite aligned with the connection probability shown in Figure 4.8. If we compare the selection techniques, we again see that high values of the bias value used for CRE degrades the service performance. Furthermore, we can observe that the number of small-cells does not have a strong impact on the service performance. In fact, when using the highest density, the service performance gets slightly worse, due to the increase of interfering cells.
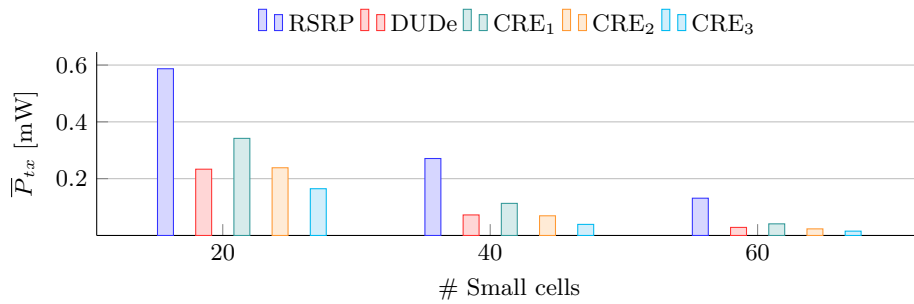
Figure 4.7: Average uplink transmission power per PRB when using different access selection techniques and for different small-cells densities

---

1 The difference between dropped and rejected services has been defined in Section 3.2.1
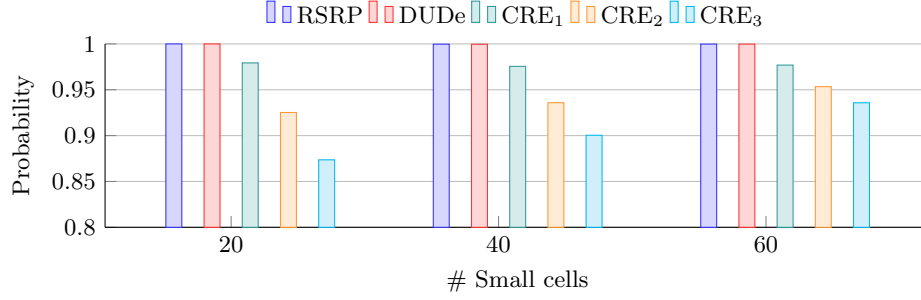
Figure 4.8: Average connection success rate when using different access selection techniques and for different small-cells densities
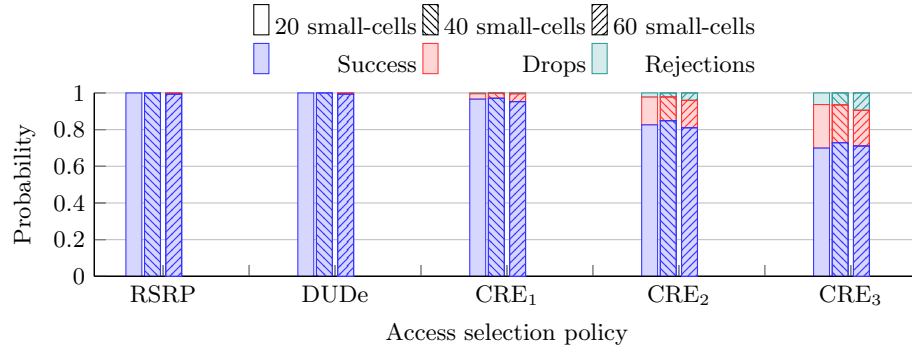
Figure 4.9: Service performance when using different access selection techniques and for different small-cell densities. For each configuration, the ratio of dropped, rejected and successful services is indicated

Finally, Figure 4.10 depicts the average number of handovers during a service session. As we have seen before, when using DUDe we can obtain the same outage probability as with RSRP, while substantially reducing the uplink transmission power. Hence, DUDe provides the best tradeoff between service performance and transmission power. However, these improvements happen at the expense of a much larger number of handovers during service sessions.

Although the number of handovers has traditionally played an important role as performance metric, it may not be as relevant in novel architectures. As was mentioned at the beginning of this chapter, leveraging new paradigms based on network functions centralization, such as C-RAN, the cooperation between access network elements can be greatly improved. In this sense, several cells can be operated by the same controller, and therefore a handover would not cause any service disruption. This centralization approach will be further studied in the following section.
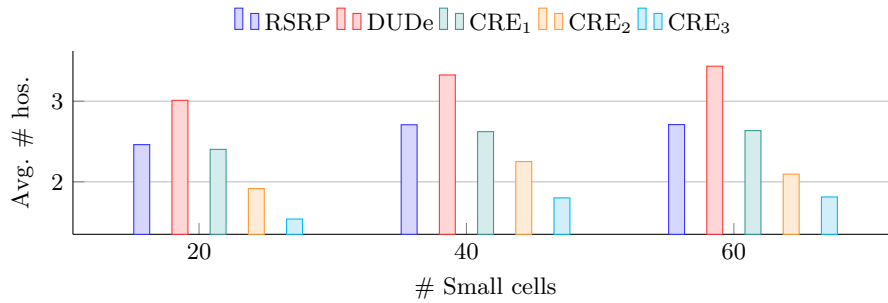
Figure 4.10: Average number of handovers per service when using different access selection techniques and for different small-cells densities. The figure only accounts for changes of base station in the uplink

4.5 Scheduling for multi-RAT heterogeneous networks

In previous sections, access selection models and strategies have been proposed and analyzed, taking sensible yet simple assumptions about the underlying scheduling procedures. Furthermore, as was mentioned earlier in this chapter, in 5G multi-connectivity scenarios several technologies will provide radio resources to tackle the foreseen traffic demand. These technologies, in turn, will be operated in different frequency bands, sharing a unified air interface [Qua16], most likely based on OFDMA.

In particular, we focus on the multi-connectivity scenarios depicted in Figure 4.11. The first one, Figure 4.11a, consists of a typical carrier aggregation deployment, where one eNodeB schedules resources from different frequency bands, potentially discontinuous. We also consider the possibility of having heterogeneous RATs combining licensed and unlicensed technologies, in a distributed or co-located manner, as shown in Figures 4.11b and 4.11c. Finally, the possibility of centralized architectures is also considered, see Figure 4.11d, where additional delays from the access network to the controller might occur. In turn, these delays imply that the information at the network controller, and then at the scheduler, may suffer from uncertainty.

Besides common features, such as QoS and stability guarantees, a multi-connectivity scheduler must also be able to face some additional challenges, as a consequence of the heterogeneity of the considered scenarios in order to ensure seamless operation of different technologies:

1. A desirable feature in multi-RAT systems is the ability to deactivate secondary RATs during low-load regimes to reduce operational costs. However, real hardware needs some time to switch back and forth between operational and sleep states, as considered in cellular standards [TS36331]. Therefore, it should be guaranteed that resources are not assigned to RATs while they are in sleep mode or are being activated.

2. The different co-deployed RATs can be heterogeneous in terms of: available number of resources, both in time and/or frequency, available modulation levels, and duty cycles when using unlicensed bands. In this sense, a multi-RAT scheduler should

(a) Macro-cell with CA

(b) Co-located RATs

(c) Heterogeneous network

(d) C-RAN
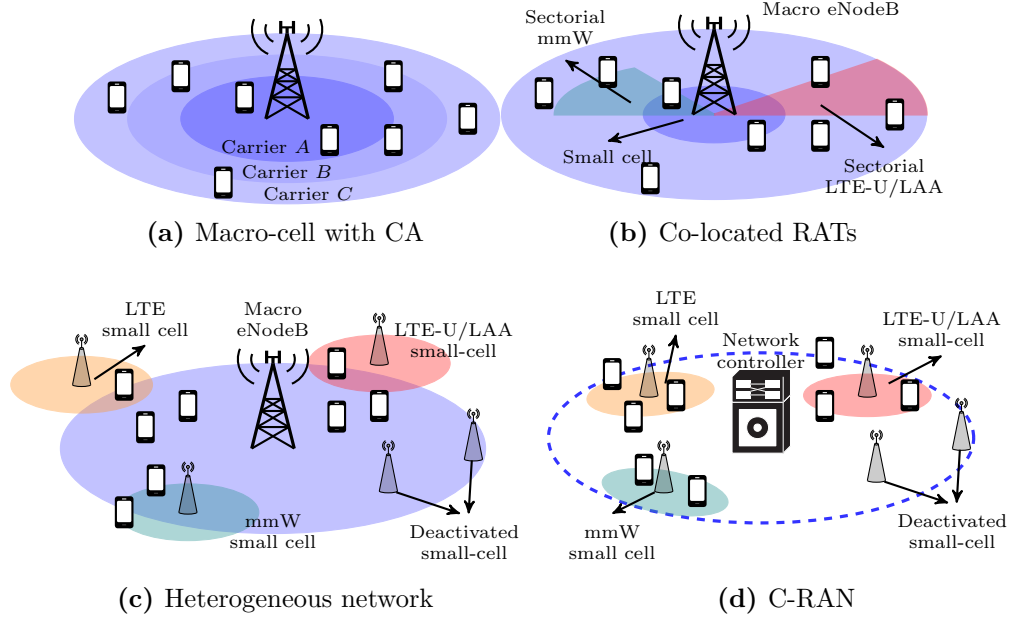
Figure 4.11: Deployments relevant for 5G and beyond

manage this multiple heterogeneity.

3. Related to the previous item, real systems impose limitations to the scheduling decision in terms of granularity. In fact, scheduling information is encoded following some specific formats, so as to get a balance between signaling overhead and scheduling granularity.

   As an example, Figure 4.12 depicts different resource allocation types allowed in LTE, which are indicated in the Downlink Control Indicator (DCI) so that UEs are aware of how to decode the data. Resource allocation in LTE can be done in three ways [Ku15], the so called type 0, 1 and 2. In type 0, consecutive PRBs are grouped into resource block groups (RBGs), so that resources are granted per RBG, thus losing granularity.

   In type 1, RBGs are grouped in subsets, and users are allocated individual PRBs within one subset. In this way, a scheduler is constrained to a subset of PRBs per user. Finally, in type 2, we can allocate any number of virtually continuous PRBs via an offset/length pair. Then, these PRBs can either be physically continuous or distributed by a permutation function defined in the standard. In either case, a scheduler will also be constrained to assigning PRBs in groups. In summary, a practical scheduler should consider this type of practical constraints imposed by protocols.

Resource allocation in OFDM-based systems has been widely studied [And07; Cap13] in the past. In many cases, an infinite backlog model has been adopted [Kwa09; Pra14], so as

(a) Resource allocation type 0

(b) Resource allocation type 1

(c) Resource allocation type 2
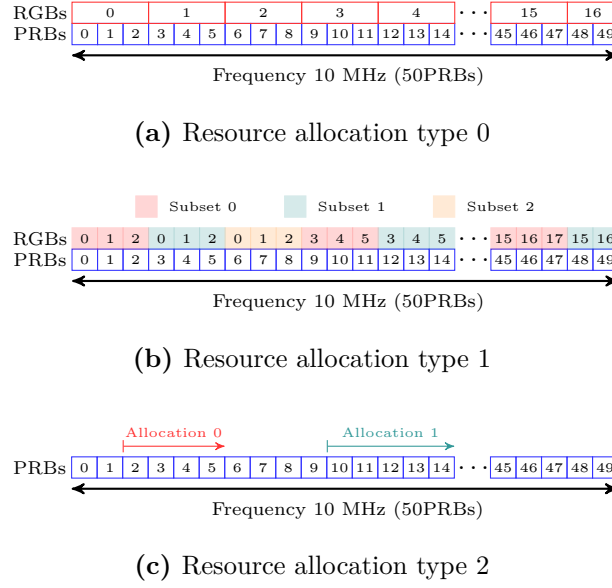
Figure 4.12: Resource allocation types for 10 MHz band

to provide overall performance bounds of the system. While relevant to understand the practical limitations of the network, this approach does not allows the implementation of practical applications.

Other works have proposed practical applications, by making some assumptions of different aspects. In some cases, the solution is tailored to some specific arrival process [Got13; Riz16], thus limiting its application scope. It is specially relevant in dense scenarios, where the traffic load in the small cells can be highly volatile and unpredictable. In other cases, the simplification is taken at the scheduling problem. For instance, in [Dec13; Nee13] the problem is modeled as a Markov Decision Process (MDP), rendering a practical algorithm but suffering from curse of dimensionality. Other practical approaches, such as [Kal15], propose empirical solutions, leading to suboptimal behavior without stability guarantees.

In all aforementioned cases, the proposed schedulers do not consider multi-connectivity scenario. Only recently, a few works consider it by using carrier aggregation [Abd15; Faz16; Sha15; Sha13; Yu15]. In [Sha15; Sha13], authors apply a proportional fairness approach for different traffic classes. Afterwards, this work was extended in [Abd15] to minimize user costs, such as energy consumption which is also considered in [Yu15]. Finally, in [Faz16] a game theoretic approach is adopted, where the power allocation for different component carriers is modeled as a competitive game between teams of access elements. Although all these works are very valuable, they lack either applicability to multi-connectivity scenarios, stability guarantees, or consideration of the practical and specific constraints detailed above.

Bearing in mind the new challenges brought by multi-connectivity scenarios, this section complements the previous ones by proposing a novel scheduling algorithm, Lagrange approximation Supple Radio (LaSR), specially designed for 5G multi-connectivity scenarios.

LaSR is a practical application of the mathematical framework proposed in [Val17; Val16]. Like some prior works [Mar12; Ned08], we exploit the connection between queue states and Lagrange multipliers to make optimal decisions. However, most of this prior work operates with continuous-valued variables, while in our case we propose a discrete control system. This is of paramount importance because it gives us a high degree of flexibility to design policies that accommodate the practical constraints and requirements of real systems, like the ones mentioned above, without compromising stability guarantees. In particular, the scheduler has been designed with the following features:

- It operates using instantaneous state information of network queues with no assumption about the stochastic properties of user traffic or load regime, while ensuring system stability. Hence, LaSR is appropriate to handle the potential unpredictable traffic handled by small cells.

- It is designed so that preferences can be added to the different technologies involved in the system. In this sense, the scheduler allows RATs to be kept at sleep state or to define policies related to the preferred modulation levels.

- The scheduler algorithm accepts decision constraints keeping stability guarantees. It can take into account constraints associated to protocols (e.g. DCI information), those related to the dimensioning of radio resources of the different technologies or the the unavailability of some RATs when switching from sleep to operational states.

- The algorithm can be tailored to support different types of QoS. In particular, we will evaluate LaSR performance to guarantee proportional fairness among elastic traffic, and delay constraints for non-elastic services.

- As was commented before, centralized architectures may cause additional delay, and thus uncertainty about the access network state. Hence, LaSR performance has been also analyzed under uncertain information.

In the following, LaSR will be described and analyzed in a number of scenarios so as to assess its performance and ability to fulfill the aforementioned requirements. It is worth noting that, although the scheduler evaluation will focus on downlink communication, its design makes it suitable also to schedule uplink resources.

4.5.1 System Model

We use conventional notation for the set of numbers. Let $\mathbb{R}$ and $\mathbb{Z}$ denote the set of real and integer numbers. We use $\mathbb{R}_+$, $\mathbb{R}^n$, and $\mathbb{R}^{n \times m}$ to represent the sets of non-negative real numbers, n-dimensional real vectors, and $m \times n$ real matrices, respectively. The rest of sets are represented in calligraphy letters.

In the following, vectors are represented in column form and written in bold font, while matrices are in upper-case font. Superscript T represents the transpose operator, $\mathbf{1}$ or $\mathbf{0}$ indicate a vector where all elements are 1 or 0, respectively, and $\boldsymbol{x} \preceq \boldsymbol{y}$ represents component-wise comparison, so that $x_i \leq y_i, \forall i$. Finally, $[\cdot]^+$ denotes the projection of a vector onto the non-negative orthant, i.e., $[\boldsymbol{x}]^+ = [\max\{0, x_1\}, \ldots, \max\{0, x_n\}], x \in \mathbb{R}^n$. Besides, sequences are represented between angle brackets $\langle \rangle$. Subscripts represent elements

in a vector and superscripts elements in a sequence. For instance, $\langle \boldsymbol{x} \rangle$ is a sequence of vectors, being $\boldsymbol{x}^t = [x_1^t, \ldots, x_n^t]^T$ a vector from $\mathbb{R}^n$ at the position t of the sequence. In turn, x_i^t is the i'th component of the t'th vector in the sequence.

We consider a multi-RAT system comprising a set of access elements $\mathcal{B}$ belonging to different RATs and a set of $\mathcal{U}$ mobile users. For simplicity, we assume that each user has one traffic flow, although the current setup may be easily extended to support multiple flows per user, by adding virtual users bearing one traffic flow.

We are not constrained to any particular hierarchy, so that RATs can follow any of the scenarios presented in Figure 4.11. Besides, we assume that any RAT can be deactivated except the primary one, to guarantee continuity of the control channel. We assume that all RATs are based on OFDMA, with a pool of resources divided in PRBs that can be allocated to users. Furthermore, we assume that each PRB can be assigned a different modulation level from a discrete set $\mathcal{M}$, according to the radio channel state. This way, $m_{u,b} \subseteq \mathcal{M}$ is the subset of modulations available to connect user $u \in \mathcal{U}$ to access element $b \in \mathcal{B}$, and $m_{u,b}^i$ represents the i'th element of the set $m_{u,b}$. This way, LaSR is designed as a network scheduler that jointly assigns PRBs and modulation levels accomplishing the aforementioned requirements

We assume that time is slotted, so that the scheduler allocates resources at the beginning of each slot for all the access elements. Although OFDMA-based RATs are considered, the actual dimension, temporal and frequential, of physical resources may differ between technologies. In order to homogenize the system, we divide the frequency/time pool of resources of all RATs into the same number of VRBs, as depicted in Figure 4.13a.

(a) virtual resource block (VRB) mapping

(b) Slotted system
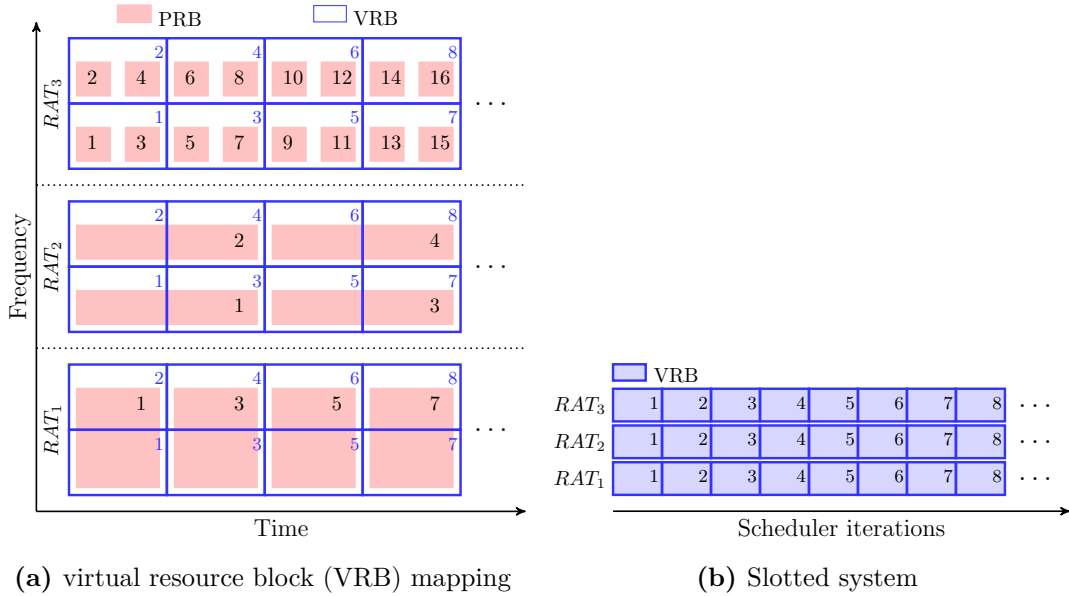
Figure 4.13: Example of mapping from PRBs to VRBs with 3 RATs having different capacities. Figure 4.13a depicts the mapping from PRBs to VRBs. Figure 4.13b shows the slotted scheme perceived by the scheduler

As can be observed, a VRB may contain a number of PRBs or a fraction of them, both time and/or frequency-wise. The definition of the VRB is a design parameter that has to trade off between the computational requirements and the loss on assignment precision. On the one hand, small VRBs allows high granularity at the expense of requiring a high number of decisions per unit time. On the other hand, if the VRB is large compared to the PRB of a RAT, several PRBs will be lotted, thus reducing the granularity of the decision making process. However, this approach allows us to homogenize the pool of resources from the potentially heterogeneous set of RATs, which simplifies the model. Furthermore, this scheme also permit us to consider different resource allocation types in a natural way. After mapping the PRBs into VRBs, we obtained a slotted system as depicted in Figure 4.13b.

The scheduler design is based on a queue system as depicted in Figure 4.14. We assume a system with $\boldsymbol{Q}$ queues, with in which a user is represented by a pair of queues for incoming and outgoing traffic, $|\boldsymbol{Q}| = 2 \times |\mathcal{U}|$. The update of the queues is defined as follows:

$$\boldsymbol{Q}^{t+1} = [\boldsymbol{Q}^t + \boldsymbol{\Delta}^t]^+, \quad t = 1,2,\ldots \tag{4.39}$$

where $\boldsymbol{Q}^t = \left[Q_1^t, Q_2^t \ldots Q_{2|\mathcal{U}|}^t\right]^T \in \mathbb{Z}_+^{2|\mathcal{U}|}$ is a vector that stores the size of all queues in the system at time slot t, and $\boldsymbol{\Delta}^t \in \mathbb{Z}^{2|\mathcal{U}|}$ is a vector containing the queues increments/decrements of data units in slot t.

Then, we link each pair of queues by a set of links that represent the possible scheduling decision (i.e. access element and MCS), leading to $N_L = |\mathcal{U}| \times |\mathcal{B}| \times |\mathcal{M}|$ links. These are represented by means of an incidence matrix $A \in \mathbb{Z}^{2|\mathcal{U}| \times N_L}$, so that element $A_{i,j}$ indicates the amount of data units per VRB departing from (if negative) or arriving to (if positive)
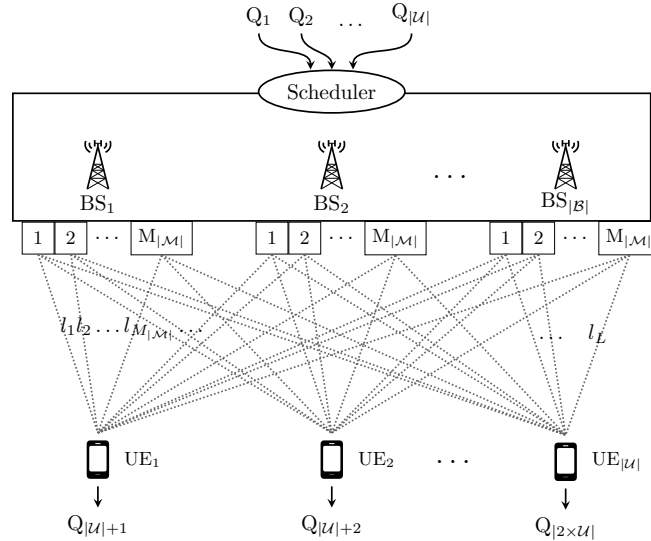
Figure 4.14: Generic system model example embracing a set of users $\mathcal{U}$ and base stations $\mathcal{B}$, where each base station has a set of modulation levels $\mathcal{M}$

queue i through link j. Note that the amount of bits transported on each VRB depends on the set of available modulation levels at the corresponding slot.

As an example, Figure 4.15 illustrates an scenario where 2 users are allocated resources from 2 different access elements, with 2 possible MCSs each. The pairs of queues Q_1/Q_3 and Q_2/Q_4 represent the incoming/outgoing queues of each user. Then, the set of links $l_i : i \in [1,8]$ connect incoming and outgoing queues with different capacities.

The capacity communications are summarized in the incidence matrix shown in Eq. 4.40. It has as many rows as queues and each column corresponds to a link. In order to represent traffic departure as a consequence of scheduling, incoming queues are given negative values, while positive ones are used to indicate traffic arriving to the outgoing queues. This way, by selecting link l_1, Q_3 would receive 1 traffic units per VRB from Q_1, while link l_8 would allow to move 20 traffic units from Q_2 to Q_4 every VRB. At each slot, the controller makes a scheduling decision to enable or disable the different links, so that an amount of traffic is allocated from the access elements to users with the corresponding MCS levels.

$$
\begin{array}{c}
\begin{array}{cccccccc}
l_1 & l_2 & l_3 & l_4 & l_5 & l_6 & l_7 & l_8
\end{array} \\
\begin{array}{c}
Q_1 \\ Q_2 \\ Q_3 \\ Q_4
\end{array}
\left(
\begin{array}{cccccccc}
-1 & -2 & -10 & -20 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & -2 & -10 & -20 \\
1 & 2 & 10 & 20 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 2 & 10 & 20
\end{array}
\right)
\end{array}
\tag{4.40}
$$

Furthermore, we define the sequence of possible actions as the binary matrix $\langle Y \rangle$, so that a vector $\boldsymbol{y}^t \in Y$ represents the action taken at slot t, and, in turn, the element of the vector at position $\boldsymbol{y}_i^t$ indicates whether link i is active or not in slot t. This way, at each slot, the queues are updated as follows:

$$
\boldsymbol{Q}^{t+1} = \left[\boldsymbol{Q}^t + A\boldsymbol{y}^t + \boldsymbol{b}^t \right]^+
\tag{4.41}
$$

where $A\boldsymbol{y}^t$ is the outcome of the scheduling decision, and $\boldsymbol{b}^t \in \mathbb{Z}^{2N}$ is a vector that represents
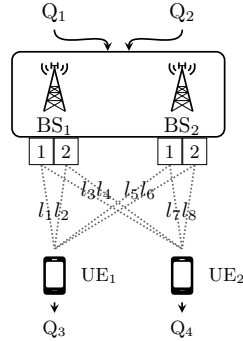
Figure 4.15: System example with 2 users, 2 access elements and 2 available MCS levels

the amount of bits that enter or leave the system. In practice, we will assume immediate service (infinite processing capacity) at the outgoing queues since they are the destination of the data. In this sense, these queues will be always empty, so that $b_i = -\infty, \forall i > N$.

Altogether, for the scheduling decision $\boldsymbol{y}^t = [1, 0, 0, 1, 0, 0, 0, 0]^T$, the queues of the example would be updated as indicated in Eq. 4.42.

$$Q_1^{t+1} = \left[Q_1^t - 21 + b_1^t\right]^+ \qquad Q_2^{t+1} = \left[Q_2^t + b_2^t\right]^+$$
$$Q_3^{t+1} = \left[Q_3^t + 21 + b_3^t\right]^+ \qquad Q_4^{t+1} = \left[Q_4^t + b_4^t\right]^+ .$$

The scheduling decision indicates that links l_1 and l_4 are activated for the current VRB, so that the first user is given resources from both access elements, while no resources are allocated for the second one. As a consequence, 21 traffic units (1 and 20 form the first and second access elements respectively) are moved from Q_1 to Q_3. Besides, incoming queues will be also updated with the amount of data entering the system, while outgoing ones (Q_3 and Q_4) are emptied in each time slot.

4.5.2 Scheduler design

As mentioned before, the objective of LaSR is to select a sequence of actions that ensures stability of the queue system upon any random arrival process at the incoming queues. Besides, the scheduler has to be able to accommodate practical constraints, and provide different types of QoS.

This way, at each time slot, the scheduler selects an action from a discrete set that tends to the optimum behavior of the system. The fact that the set of possible solutions is discrete makes any model of the underlying optimization problem non-differentiable, and then non-convex [Boy04], what would make it difficult to solve.

The main contribution of the LaSR algorithm is the definition of a practical methodology that selects discrete actions from the set of possible ones, so that the optimality is preserved in the long term. Besides, we will describe how the aforementioned requirements can be integrated within the algorithm, while still ensuring optimality.

We start by defining a convex model of the scheduling optimization problem, whose outcome will be referred to as fluid solution, as follows:

Problem 7 (LaSR fluid problem)**.**

$$\min_{\boldsymbol{x} \in \mathcal{X}} \quad f(\boldsymbol{x}) \tag{4.42}$$

$$\boldsymbol{s.t.} \quad A\boldsymbol{x} + \boldsymbol{b} \le 0 \tag{4.43}$$

where $f : \mathcal{X} \to \mathbb{R}$ is the convex objective function, so that $\boldsymbol{x} \in \mathbb{R}_+^{2 \times |\mathcal{U}|}$. Besides, A is the incidence matrix as described in Section 4.5.1 and $\boldsymbol{b} \in \mathbb{R}^{2 \times |\mathcal{U}|}$ represents the traffic random arrival to the queue system. We assume that the set $\mathcal{X}_0 \coloneqq \{\boldsymbol{x} \in \mathcal{X} | A\boldsymbol{x} + \bar{\boldsymbol{b}} \le 0\} \ne \varnothing$, so that Slater condition holds [Boy04] and Problem 7 is feasible.

As can be seen, the problem aims to minimize the cost decision modeled by the objective function (e.g. QoS, assignment cost, etc.), while ensuring that the queue system is stable, which is represented by the inequality constraint 4.43 (note the sign of the A matrix elements).

We start by relaxing the problem by means of the Lagrange dual problem. In short, this relaxation transforms the original, or primary, constrained problem into an unconstrained one (dual problem), in which the initial constrains are given a cost (Lagrange multipliers). Then, when solving the dual problem we obtain approximate solutions to the primal one, where violations of the initial constrains are penalized. This way iterative methods are applied to the dual problem, so that the sequence of solutions tend to the optimal one. We refer the reader to [Boy04] and [Val17] for a more detailed description on this topic.

$$L(\boldsymbol{x}, \boldsymbol{\psi}) = f(\boldsymbol{x}) + \boldsymbol{\psi}^T (A\boldsymbol{x} - \boldsymbol{b}) \tag{4.44}$$

We define the Lagrange dual function as indicated in Eq. 4.44, where $\boldsymbol{\psi}$ is the vector of the Lagrange multipliers associated to the inequality constraint of the primal problem. Then we can defined the Lagrange dual problem as follows:

Problem 8 (LaSR dual problem)**.**

$$\max_{\boldsymbol{\psi} \geq 0} \quad h(\boldsymbol{\psi}) \tag{4.45}$$

$$where \quad h(\boldsymbol{\psi}) \coloneqq \inf_{\boldsymbol{x} \in \mathcal{X}} L(\boldsymbol{x}, \boldsymbol{\psi}) \tag{4.46}$$

which provides a lower bound for the optimal value of the primal problem. We note that solving the dual problem is equivalent to solving the primal one when strong duality holds, which is always the case when Slater conditions is satisfied. We refer the reader to [Boy04] for a succinct explanation of convex duality.

Besides, the dual problem is always concave in $\boldsymbol{\psi}$, and then it can be solved by well known methods. In particular we will exploit the sub-gradient method, which allows converging to the optimal solution even when random perturbations are added [Val17], as it is our case.

At each iteration t, the sub-gradient method updates the Lagrange multipliers as follows:

$$\boldsymbol{\psi}^{t+1} = \left[\boldsymbol{\psi}^t + \alpha \left(A\boldsymbol{x}^t + \boldsymbol{b}^t \right) \right]^+ \tag{4.47}$$

where $\boldsymbol{x}^t \in \arg\min_{\boldsymbol{x} \in \mathcal{X}} L(\boldsymbol{x}, \boldsymbol{\psi}^t)$ is the fluid solution of the current iteration and $\alpha > 0$ is the update step, whose value can be selected according to the problem. Although more complex updates exist, we will only use constant step size. We define the finite set of possible actions $\mathcal{Y}$, and assume that the domain of the primal problem is $\mathcal{X} \subseteq Conv(\mathcal{Y})$. Besides, we will regard selecting a point $\boldsymbol{y}$ from the finite set $\mathcal{Y}$ as selecting a discrete scheduling action.

Then, we note that the update of the Lagrange multipliers has queue form, and that if we replace the fluid solution $\boldsymbol{x}^t$ by the discrete one $\boldsymbol{y}^t$ we can write the following update:

$$\boldsymbol{\mu}^{t+1} = \left[\boldsymbol{\mu}^t + \alpha\left(A\boldsymbol{y}^t + \boldsymbol{b}^t\right)\right]^+ \tag{4.48}$$

$$\text{where} \quad \boldsymbol{\mu}^t = \alpha \boldsymbol{Q}^t \tag{4.49}$$

with $\boldsymbol{\psi}^1 = \boldsymbol{\mu}^1 \geq 0$ and the update step $\alpha > 0$. As shown in [Val17], provided the differences between the fluid and discrete solutions, $\|\sum_{i=1}^t \boldsymbol{x}^i - \boldsymbol{y}^i\|_2$, are uniformly bounded, then $\boldsymbol{\mu}^t$ is an approximation of the Lagrange multipliers, so that the queue occupancies can be seen as scaled versions of the Lagrange multipliers. Having this in mind, it can be proven that a sequence of discrete actions $\langle\boldsymbol{y}\rangle$ tends to the optimal solution of the problem. This observation has two main implication:

1. We can solve the Problem 7 by using discrete actions.

2. Any sequence of discrete actions $\langle\boldsymbol{y}\rangle$, that ensure $\|\sum_{i=1}^t \boldsymbol{x}^i - \boldsymbol{y}^i\|_2$ to be uniformly bounded can be used to solve the problem.

All in all, the work presented in [Val17] provides us with a mathematical framework to accomplish the requirements previously stated for the scheduler. This way, each scheduling decision, upon the random arrival of traffic, is seen as one iteration of the sub-gradient method in which the fluid solution is obtained. Then, based on that solution and considering the bounding requirements, a discrete solution is chosen from the set of possible ones.

In the following, we will present the practical implementation of the algorithm that ensures the convergence of the problem. Afterwards, we will provide a set of possible utility functions to implement the QoS requirement.

4.5.3 Scheduling algorithm implementation

We start by defining the matrix W as the collection of points in $\mathcal{Y}$ as columns. Then we define sets

$$\mathcal{E} := \left\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{|\mathcal{Y}|}\right\} \tag{4.50}$$

$$\mathcal{U} := Conv(\mathcal{E}) = \{u \in [0,1]^{|\mathcal{Y}|} \mid \mathbf{1}^T\boldsymbol{u} = 1\} \tag{4.51}$$

where $\boldsymbol{e}_i$ is a basis vector of dimension $|\mathcal{Y}|$, so that the element at the position i is equal to 1 and the other elements are 0. Bearing in mind that $\mathcal{X} \subseteq Conv(\mathcal{Y})$, we know that any point $x \in \mathcal{X}$ can be obtained as a convex combination of the points in $\mathcal{Y}$. Thus, there must exist at least one vector $\boldsymbol{u}^t \in \mathcal{U}$ so that $\boldsymbol{x}^t = W\boldsymbol{u}^t$, which can be obtained by solving $\min_{\boldsymbol{u}\in\mathcal{U}}\|\boldsymbol{x}^t - W\boldsymbol{u}\|_2^2$. Likewise, there exists a vector $\boldsymbol{v}^t \in \mathcal{E}$ so that $\boldsymbol{y}^t = W\boldsymbol{v}^t$

Then, we can reformulate the differences between the fluid and discrete solutions as

follows:

$$\left\| \sum_{i=1}^{t} \boldsymbol{x}^i - \boldsymbol{y}^i \right\|_2 = \left\| \sum_{i=1}^{t} W(\boldsymbol{u}^i - \boldsymbol{v}^i) \right\| \leq \|W\|_2 \left\| \sum_{i=1}^{t} \boldsymbol{u}^i - \boldsymbol{v}^i \right\|_2 \tag{4.52}$$

so that ensuring that $\left\| \boldsymbol{\Delta}^t \right\|_2 = \left\| \sum_{i=1}^{t} \boldsymbol{u}^i - \boldsymbol{v}^i \right\|_2$ is bounded will suffice to prove the convergence of the scheduler.

For the scheduler design, we consider the construction of a sequence of discrete values $\langle \boldsymbol{v} \rangle$ in a greedy manner, so the selection of $\boldsymbol{v}^t$ only depends on $\boldsymbol{v}^{t-1}$ and $\boldsymbol{u}^t$. From [Val17], we know that the update

$$\boldsymbol{v}^t \in \arg\min_{\boldsymbol{v} \in \mathcal{E}} \left\| \boldsymbol{\Delta}^{t-1} + \boldsymbol{u}^t - \boldsymbol{v} \right\|_\infty \tag{4.53}$$

where $\boldsymbol{\Delta}^t = \sum_{i=1}^{t} \boldsymbol{u}^t - \boldsymbol{v}^t$, ensures that $\left\| \boldsymbol{\Delta}^t \right\|_2$ is bounded even with additional constraints imposed to the set of available discrete decisions. In particular, Theorem 5 from [Val17] states that if the number of slots during which an action from $\mathcal{Y}$ is upper bounded by $\sigma \in \mathbb{Z}^+$, then

$$\left\| \boldsymbol{\Delta}^t \right\|_2 = \left\| \sum_{i=1}^{t} \boldsymbol{u}^i - \boldsymbol{v}^i \right\|_2 \leq \sigma \sqrt{|\mathcal{Y}|}(|\mathcal{Y}| - 1) \tag{4.54}$$

which indicates that provided a particular discrete action is not infinitely unavailable, the update in Eq. 4.53 will ensure that the sequence of discrete actions tends to the optimum solution of Problem 7.

We present the main LaSR operation in Algorithm 2. As can be observed, LaSR makes a scheduling decision for every VRB in a temporal slot, which, in practical systems, might correspond to the Transmission Time Interval (TTI).

The algorithm makes use of two auxiliary vectors: $\boldsymbol{\psi}$ and $\boldsymbol{\Delta} \in \mathbb{R}^{|\mathcal{Y}|}$. The first one, $\boldsymbol{\psi}$, represents the Lagrange multipliers, and it is used to perform the sub-gradient method needed to solve Problem 8. Then, $\boldsymbol{\Delta}$ maintains the aggregate deficit or surplus of fluid actions taken when mapping $\boldsymbol{u}^t$ to $\boldsymbol{v}^t$. Furthermore, the algorithm assumes a predefined set of discrete actions $\mathcal{Y}$ grouped in matrix W, and the standard basis of the discrete actions set, $\mathcal{E}$.

The algorithm takes into account that some actions might be unavailable during certain iterations. In this sense, in step 7, function SET-STATE() updates the set of available actions. For instance, if no resource from a RAT is allocated during a time period, then it might be assumed that the RAT turns to sleep state. Then, when at least one link related to such RAT is selected, the RAT needs to get activated, what in real systems may take some time. The activation/deactivation policy is shown in Algorithm 3.

Then, in step 8, the algorithm updates the set of available actions $\hat{\mathcal{E}}$, according to the RATs availability and other possible constraints. In the following we will further describe

---

Algorithm 2: LaSR algorithm. Temporal superscripts are not indicated as they refer to vector or matrix variables, so that an update refers to change of value of such variable. Names are reused to ease readability.

1: $W := \left[\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{|\mathcal{Y}|}\right]$ ▷ Matrix of elements in $\mathcal{Y}$
2: $\mathcal{E} := \left\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}^{|\mathcal{Y}|}\right\}$ ▷ Standard basis of dimension $|\mathcal{Y}|$
3: $\boldsymbol{\Delta} \leftarrow \boldsymbol{0}$
4: $\boldsymbol{\psi} \leftarrow \boldsymbol{0}$
5: **procedure** LaSR ALGORITHM($\boldsymbol{\Delta}, \boldsymbol{\psi}$)
6:     **for** $t = 1, 2, \ldots$ and $VRB = 1, 2, \ldots$ **do**
7:         SET-STATE $(t, \boldsymbol{\Delta})$ ▷ Activate/deactivate RAT
8:         $\hat{\mathcal{E}} \subseteq \mathcal{E}$ ▷ Update available set of actions, this can depend on several parameters and done for every TTI and VRB
9:         $\boldsymbol{x} \leftarrow$ SOLVER($\boldsymbol{\psi}$)   ▷ Solution to Problem 8
10:         $\boldsymbol{u} \in \arg\min_{\boldsymbol{u} \in \mathcal{U}} \|\boldsymbol{x} - W\boldsymbol{u}\|_2^2$
11:         $\boldsymbol{v} \in \arg\min_{\boldsymbol{v} \in \hat{\mathcal{E}}} \|\boldsymbol{\Delta} + \boldsymbol{u} - \boldsymbol{v}\|_\infty$
12:         $\boldsymbol{\Delta} \leftarrow \boldsymbol{\Delta} + \boldsymbol{u} - \boldsymbol{v}$
13:         $\boldsymbol{\psi} = \left[\boldsymbol{\psi} + \alpha(A\boldsymbol{x} + \boldsymbol{b})\right]^+$
14:         $\boldsymbol{y} \leftarrow W\boldsymbol{v}$ ▷ Discrete solution
15:     **end for**
16: **end procedure**

---

the procedure to activate and deactivate RATs, taking into account the scheduler operation. We also note that other scenario specific constraints may be implemented, as we will see in the evaluation section.

Next, step 9 provides the continuous solution to the dual problem, and, based on that, the convex combination of the standard basis of the set of solutions $\mathcal{Y}$ is obtained in step 10. The minimization problem in step 10 is the well known least squares problem, that can be solved by using a linear system of equations (see [Boy04, Section 6.2]).

Afterwards, the actual element of the standard basis of the discrete elements is obtained in step 11, taking into account the current constraints. In fact, the selection of the basis can be implemented by simply selecting the highest element in vector $\boldsymbol{\Delta} + \boldsymbol{u}$. Finally, the auxiliary variables, $\boldsymbol{\Delta}$ and $\boldsymbol{\psi}$, are updated in steps 12 and 13, and the discrete action is obtained by simply computing $\boldsymbol{y} = W\boldsymbol{v}$ in step 14.

The implementation of the SET-STATE() function depends on the RATs and type of unavailabilities. For instance, a Channel Quality Indicator (CQI) report may indicate that certain MCS index is not available, a cognitive radio controller in unlicensed band alerting that the current slot corresponds to a sleep period, a hardware controller notifying a RAT is not yet activated, or the DCI mapper notifying of some signaling constraints.

In particular, we have considered two different kind of constraints: unavailability of MCS indexes (links) due to poor link quality, and activation or deactivation of RATs as a consequence of inactivity. The former is performed by simply using lookup tables and current

link quality levels to know which set of MCS indexes are available. The later, on the other hand, requires a more complex procedure. In a simple scenario the activation/deactivation of RATs would only depend on the decisions applied by the scheduler $\boldsymbol{y}$. However, as shown in Algorithm 2, the scheduler is aware of the accumulated deficit/surplus of decisions (i.e. traffic assigned to links) kept by vector $\boldsymbol{\Delta}$. In order to avoid that a RAT whose links have a traffic deficit is deactivated, we have implemented Algorithm 3.

This algorithm checks every T time slots whether the RATs need to be deactivated, based on the accumulated deficit and surplus. To this end, the procedure makes use of two internal binary variables $\boldsymbol{m}$ and $\hat{\boldsymbol{m}}$. The first one indicates those links having a deficit in the current slot. On the other hand, $\hat{\boldsymbol{m}}$ identifies those links that have not had any deficit at any slot in the current interval. It is worth noting that different deactivation intervals might be defined per RAT.

As shown in Algorithm 3, the procedure to activate and deactivate RATs takes as input parameters the current time slot and the vector $\boldsymbol{\Delta}$ that carries the deficit/surplus of the actions. First, the algorithm stores in $\boldsymbol{m}$ the links that have some deficit at the current slot and, based on that, updates the value of $\hat{\boldsymbol{m}}$ in step 11. Later on, those RATs having some deficit are kept awake in step 14. Finally, at the end of time interval T, RATs without deficit in the last T slots are deactivated. As can be observed, the current activation/deactivation algorithm is rather conservative, since it keeps active RATs that have had some deficit (i.e. should have been selected) regardless whether the deficit has been compensated at the end of the interval. More aggressive policies might be used, for instance using only the state at the end of the interval.

Design of the Objective Function

In this section we define the objective function that allows LaSR to provide different levels of QoS and operator policies. From the QoS viewpoint, the function is able to provide different priorities to elastic and non-elastic services, and fairness among elastic ones. Besides, from the operator policy perspective, it provides flexibility to manage heterogeneous networks. For instance a network operator might want to aggregate as much traffic as possible into a subset of the available RATs, in order to save costs by switching off other access elements.

We define the objective function as a weighted cost function, as follows:

$$f(\boldsymbol{x}) = \rho \cdot \boldsymbol{\omega}^T \boldsymbol{x} - (1 - \rho) \cdot \sum_{i=1}^{|\mathcal{U}|} \log\left(U_i(x_i)\right) \tag{4.55}$$

where $0 \le \rho \le 1$ is a design parameter that controls the relative weight given to the different cost components. First, the vector $\boldsymbol{\omega} \in \mathbb{R}_+^L$ assigns different costs to the links (i.e. RATs and MCS indexes), so that, according to the specific policy, the objective function can encourage or deter the selection of specific RATs and modulation schemes. Then, the second component aims to foster different levels of QoS, according to the traffic type. It is defined as the sum of the utilities corresponding to each traffic queue $i = 1, 2, \cdots, |\mathcal{U}|$. As can be seen in Eq. 4.55, we consider a utility proportional fairness, so that, if active, different

---

Algorithm 3: Activation/Deactivation algorithm. $\vee$, $\wedge$, $\neg$ are logic *OR*, *AND* and *NOT* operators respectively. $\mathbb{T}(\cdot)$ is an element-wise function that returns 1 if $(\cdot) > 0$ and 0 otherwise.

1: $T \in \mathbb{Z}_+$ $\triangleright$ Epoch interval to make sleep decisions
2: $C \in \{0,1\}^{R \times L}$ $\triangleright$ C maps links to RATs
3: $\boldsymbol{m} \in \{0,1\}^L$ $\triangleright$ Binary vector where the element i'th is 1 if resources are due to link i
4: $\hat{\boldsymbol{m}} \in \{0,1\}^L$ $\triangleright$ Binary vector where the element i'th is 1 if link i has not been used in the current interval T
5: **procedure** SET-STATE$(t, \varDelta)$
6:     **for** $i \in \{1, \ldots, |\varDelta|\}$ **do**
7:        $\varepsilon \leftarrow \boldsymbol{0}^{|\varDelta|}$
8:        $\varepsilon_i \leftarrow \varDelta$
9:        $\boldsymbol{m} \leftarrow \boldsymbol{m} \vee \mathbb{T}(W\varepsilon > \boldsymbol{0})$ $\triangleright$ Update only with deficit
10:     **end for**
11:     $\hat{\boldsymbol{m}} \leftarrow \hat{\boldsymbol{m}} \wedge (\neg \boldsymbol{m})$ $\triangleright$ Update un-used links in current interval
12:     SET-AWAKE$(\mathbb{T}(C\boldsymbol{m}))$ $\triangleright$ Awake used RATs
13:     **if** $t \pmod{T} == 0$ **then**
14:        SET-SLEEP$(\mathbb{T}(C\hat{\boldsymbol{m}} > 0))$ $\triangleright$ Sleep un-used RATs
15:        $\hat{\boldsymbol{m}} \leftarrow \boldsymbol{1}^L$
16:     **end if**
17: **end procedure**

---

utility functions are proportionally considered.

All in all, for a given ρ, a solution $\boldsymbol{x}^*$ to Problem 7 corresponds to a point in the Pareto optimal trade-off between utility satisfaction and cost minimization. For example, by setting $\rho = 0$ we would only consider QoS satisfaction regardless of how the infrastructure is used. On the other hand, if $\rho = 1$ the scheduler would not provide QoS, so that it would foster the actual network management policy. In the following we define the utility functions that have been considered, and we prove that the overall objective function is convex, so that the underlying sub-gradient method implemented by the scheduler will converge to the optimum.

Since the utility function does not consider resources but traffic, we first define the vector of traffic rate as $r := -A\boldsymbol{x} = [r_1, r_2, \cdots, r_{2N}]^T$. It is worth noting that, in practice, only the first N elements of the rate vector are used, since the last N elements refer to the outgoing queues, see Figure 4.14.

The first QoS criterion aims to ensure that the non-elastic (or delay-sensitive) traffic is prioritized. To this end, we make use of the following sigmoid-shaped function [Lee05; Sha15]:

$$U_i(x_i) := g_i(r_i) := \frac{1 + e^{a_i b_i}}{e^{a_i b_i}} \left(\frac{1}{1 + e^{-a_i(r_i - b_i)}} - \frac{1}{1 + e^{a_i b_i}} \right) \tag{4.56}$$

where $a_i \in \mathbb{R}^+$ and $b_i \in \mathbb{R}^+$ are defined according to the system configuration. We can observe that the function is unary bounded for $r \in \mathbb{R}^+$, so that $g_i(0) = 0$ and $g_i(\infty) = 1$.

This function resembles a step-like one taking very low utility below a threshold, and a large and constant value once that threshold is reached. In this sense, parameter a_i establishes the slope of the function and b_i the inflection point. For example, if $a_i \approx b_i \approx r_v$ this function is a good approximation to a step function for voice traffic with rate requirement r_v. On the other hand, when $a_i \ll r_r \ll b_i$, it can be used to model the utility of adaptive real-time applications with mean rate r_r.

The second utility function is a logarithmic one defined as follows:

$$U_i(x_i) = h_i(r_i) := \frac{\log\left(1 + c_i \cdot r_i\right)}{\log\left(1 + c_i \cdot \hat{r}_i\right)}, \tag{4.57}$$

where $\hat{r}_i$ represents the maximum aggregated throughput for the particular service, and $c_i \in \mathbb{R}^+$ is the satisfaction growth rate. Alike in the previous case, this utility function satisfies the unary bounding for the r domain, $r \in [0, \hat{r}]$. In practice, this function is useful to model the utility for elastic services, so that the higher rate, the larger utility up to a certain limit. We also consider the case $U_i := 0$, so that the traffic corresponding to the ith queue does not have any QoS requirement. It would resemble a best-effort case.

The following lemma formalizes the suitability of the utility function for Problem 7:

Lemma 1. *For any $\boldsymbol{x} \in \mathbb{R}_+$, the function $f(\boldsymbol{x})$ defined in Eq. 4.55 is convex.*

Proof. To prove the convexity of the objective function, we make use of convexity properties of common functions and transformations. In particular, we know that summation and affine transformations preserve convexity (and so concavity) [Boy04].

First, we note that the utility function is a summation, so that we need to prove that each term of the objective function is convex.

It follows that the first component, $\boldsymbol{w}^T\boldsymbol{x}$, is convex, since it is an affine transformation. As for the second term, taking into account that it is a negative summation, we need to prove that $\sum_{i=1}^{N} \log\left[U_i(x_i)\right]$ is concave, which boils down to proving that each component $log[U_i(x_i)]$ is concave.

In turn, utilities $U_i(x_i)$ are defined as functions of $\boldsymbol{r} := A\boldsymbol{x}$, which is an affine transformation of $\boldsymbol{x}$. Thus, we only have to prove that functions $\log[g_i(r_i)]$ and $\log[h_i(r_i)]$ are concave functions in the domain of $\boldsymbol{r}$.

For a real valued function, f, of a single variable, to be strictly concave (and then concave) on an interval, it suffices that the function is twice differentiable and its second derivative $f'' < 0$ within the interval.

We can reformulate $g(\boldsymbol{r})$ as follows:

$$\log[g_i(r_i)] \quad = \log\left[A_i\left(\frac{1}{1 - e^{-\alpha}} - B_i\right)\right] \tag{4.58}$$

$$where \quad A_i = \frac{1 + e^{a_i b_i}}{e^{a_i b_i}}; \quad B_i = \frac{1}{1 + e^{a_i b_i}}; \quad \alpha = a_i(r_i - b_i) \tag{4.59}$$

where $A \in \mathbb{R}^+$ and $B \in \mathbb{R}^+$. We also observe that [Sha15]

$$0 \leq A_i \left(\frac{1}{1 - e^{-\alpha}} - B_i \right) \leq 1 \tag{4.60}$$

$$B_i \leq \frac{1}{1 - e^{-\alpha}} \leq \frac{1 + A_i B_i}{A_i} \tag{4.61}$$

$$\frac{1}{B_i} \geq 1 - e^{-\alpha} \geq \frac{A_i}{1 + A_i B_i} \tag{4.62}$$

$$0 \leq 1 - B_i \left(1 - e^{-\alpha} \right) \leq \frac{1}{1 + A_i B_i} \tag{4.63}$$

It follows that the first and second derivatives of $\log[g_i(r_i)]$ are

$$\frac{d \log[g_i(r_i)]}{dr_i} = A a_i \left[\frac{B_i e^{-\alpha}}{1 - B_i \left(1 + e^{-\alpha} \right)} + \frac{e^{-\alpha}}{1 + e^{-\alpha}} \right] \tag{4.64}$$

$$\frac{d^2 \log[g_i(r_i)]}{dr_i^2} = -A a_i^2 \left[\frac{B e^{-\alpha}}{A_i \left[1 - B_i (1 + e^{-\alpha}) \right]} + \frac{-e^{\alpha}}{1 + e^{-\alpha}} \right] < 0 \tag{4.65}$$

such that, the function is concave.

In the case of function $h_i(r_i)$, we first observe that for a function $p(x) = log(q(x))$, we can define $\frac{d^2 p(x)}{dx^2} = \frac{q''(x)q(x) - q'(x)^2}{q(x)^2}$. According to the definition of $h_i(r_i)$, we know that it is a positive valued functions, in fact it is unary bounded. Hence, proving that $\frac{d^2 h_i(r_i)}{dr_i^2} < 0$ would prove its concavity.

$h_i(r_i)$ can be expressed as $h_i(r_i) = \frac{\log(1 + c_i r_i)}{C}$, being $C = log(1 + c_i \hat{r}_i) \in \mathbb{R}^+$. Then, we can calculate the first derivative as $\frac{dh_i(r_i)}{dr_i} = \frac{c_i}{C} \frac{1}{1 + c_i r_i}$, and thus $\frac{d^2 h_i(r_i)}{dr_i^2} = \frac{-c_i^2}{C(1 + c_i r_i)} < 0$. $\qquad\square$

4.5.4 Performance Evaluation

In this section we validate, by means of simulation, the most relevant features of LaSR, according to the design objectives previously described. In a first scenario, we will study the stability of the scheduling algorithm under different traffic distribution and its response to the costs in the objective function. Afterwards, the scheduler performance will be analyzed when applying the QoS utility functions.

Later on, LaSR will be evaluated under different types of constraints and in comparison with an alternative scheduling algorithm. In particular, we will use as benchmark the Queue Side Greedy (QSG) scheduling algorithm proposed in [Che13] since, to the best of our knowledge, it is the only work with carrier aggregation support comparable, in its applicability to real systems, to ours. The approach of [Che13] is however a heuristic scheme with no performance guarantees, and its sole goal is to minimize the size of the queues.

In this sense, we will first illustrate the algorithm response to constraints in the set of possible MCSs, thus emulating dynamics of the scenario (e.g. mobility of users). Finally,

we will assess how LaSR adapts to other scheduling constraints and imperfect information, both of the queues occupancy and available resources.

For simplicity, and unless otherwise stated, we carry out most of our tests in a scenario with 3 RATs with fixed bandwidths of 10 MHz (e.g. an LTE macro-cell), 5 MHz (e.g. an LTE small-cell) and 20 MHz (e.g. a mmW), respectively. Furthermore, we adopt standard LTE resource dimensioning, so that each PRB occupies 180 KHz and the TTI is 1 ms. This way, the RATs have pools of resources of 50, 25 and 100 PRBs respectively. In general, the number of VRBs is 50, so that each resource of the small cell consists of 2 VRBs and 2 resources from the mmW technology are grouped at each allocation, as depicted in Figure 4.13.

Besides, the LTE macro cell plays the role of primary RAT, and thus it is always available. On the other side, secondary RATs have activation and deactivation times of 8 and 1280 ms, respectively [TS36331]. In this sense, if a link of a secondary RAT has allocation deficit (see Algorithm 3), the RAT initiates the activation process, but the resources would not become available until the activation time is elapsed (i.e. 8 TTIs).

Finally, we will use the configuration of MCS levels and associated costs indicated in Table 4.4. It is worth noting that the cost is associated to a VRB and TTI, so that objective function defined in Eq. 4.55 is unitless. As can be observed, this configuration provides 3 values of spectral efficiency for each RAT, emulating different channel conditions (e.g. bad, medium, good). Furthermore, the system is configured to foster the usage of the macro cell (i.e. primary RAT) over the others, and to exploit LTE over mmW.

In particular, the values indicated in Table 4.4 are obtained from the MCS indexes $\{0, 14, 28\}$ in [TS36213], and the modulations used for the mmW RAT correspond to the indexes $\{0, 9, 23\}$ from those proposed by project MiWEBA [Miw41]. All in all, a scheduling decision is performed for every VRB, so that the algorithm is executed 50 times every TTI.

Stability assessment

We start by evaluating a scenario with 5 input queues, which may represent users or traffic types, and using the setup described in Table 4.4. In this sense, the objective function only considers allocation costs, so that no QoS criteria is applied to the traffic.

The scheduler performance is analyzed under different traffic distributions, that model the arrival bit rate, and system loads. The maximum system capacity is calculated assuming that the full capacity of the RATs is used with the largest MCS indexes, thus providing

Table 4.4: Default system configuration of available MCS indexes, indicated by their spectral efficiency, and associated costs. The spectral efficiency has units of bps/Hz and the cost of bits^{-1}

	Spectral efficiency/Cost		
Macro	0.2344/1	2.1602/1	5.5547/1
Small	0.2344/2	2.1602/2	5.5547/2
mmW	0.5/6	2.0/6	4.875/6

the highest spectral efficiency. Afterwards, the mean value of the distribution is defined so that the traffic aggregated by the incoming queues corresponds to the desired system load. In particular, we use a *Poisson*, a *Uniform* and two *Normal* distributions, namely *Normal*$_1$ and *Normal*$_2$, with variances of 1000 and 100 bits respectively. The different configurations are evaluated during 20 seconds (i.e. 20000 TTIs).

The goal of this analysis is twofold. First, to validate that LaSR is able to stabilize the queues up to reaching the boundary of the system capacity regardless of the traffic pattern. Second, to show that LaSR performs optimally, according to our objective function, in all load regimes.

Figure 4.16 shows the average delay and queue size of the incoming queues during the simulation, upon different system loads and traffic patterns. In order to provide a fairer comparison under different load regimes, the queue size is normalized by dividing the number of bits in the queue by the average incoming rate. This way, a value of 1 would represent that the queue size is, in average, the same as the incoming rate.

In general, we can observe that the scheduler keeps the queues stable even when the system is in the capacity limit. Indeed, Figure 4.16a shows almost zero delay for all configurations, and the size of the queues depicted in Figure 4.16b is therefore slightly above the unit, reflecting the minimum achievable value. In fact, we see some delay when the system is not saturated, due to the costs used in the objective function, as will be shown later.

Furthermore, we can observe that LaSR performs worse with traffic distributions with higher variance, specially when the system is close to the saturation point. This effect is evident when using a *Uniform* distribution, where the delay increases to above 2 ms.

In order to illustrate the scheduler dynamics, Figure 4.17 depicts the temporal evolution of the incoming queues for the different system load regimes with *Uniform* traffic pattern, which represents the worst case, according to the results shown in Figure 4.16. Firstly, we can observe that, when the load is in the boundary of the system capacity, LaSR needs more time to stabilize the queues, leading to the higher average delay observed before. Furthermore, the results show that, after a transitory, queue occupancies tend to a stable value, which increases with the total load. The reason of the peaky transitory lies is the usage of the sub-gradient method to solve the underlying optimization problem, and provides an intuition of the LaSR operation.

According to Algorithm 2, when LaSR is called, it obtains the fluid solution by solving the dual problem, Problem 8, for a given update of the Lagrange multipliers, $\boldsymbol{\psi}$. Besides, for the current configuration of the objective function, Problem 8 trades off between the cost associated to resources usage and the Lagrange multipliers[1]. In turn, as explained before, the value of $\boldsymbol{\psi}$ is a scaled version of the queues occupancy, so that LaSR prevents the allocation of resources until the queues grow enough, and thus an inflexion point is reached.

Then, the smoothness of the convergence, as well as its time, can be easily tuned by

---

1 Note that the value of $A\boldsymbol{x} - b$ is negative for the incoming queues. Besides, the outgoing ones will not affect the problem solution, since they are defined to have infinite processing capacity

(a) Mean traffic delay

(b) Mean queue size
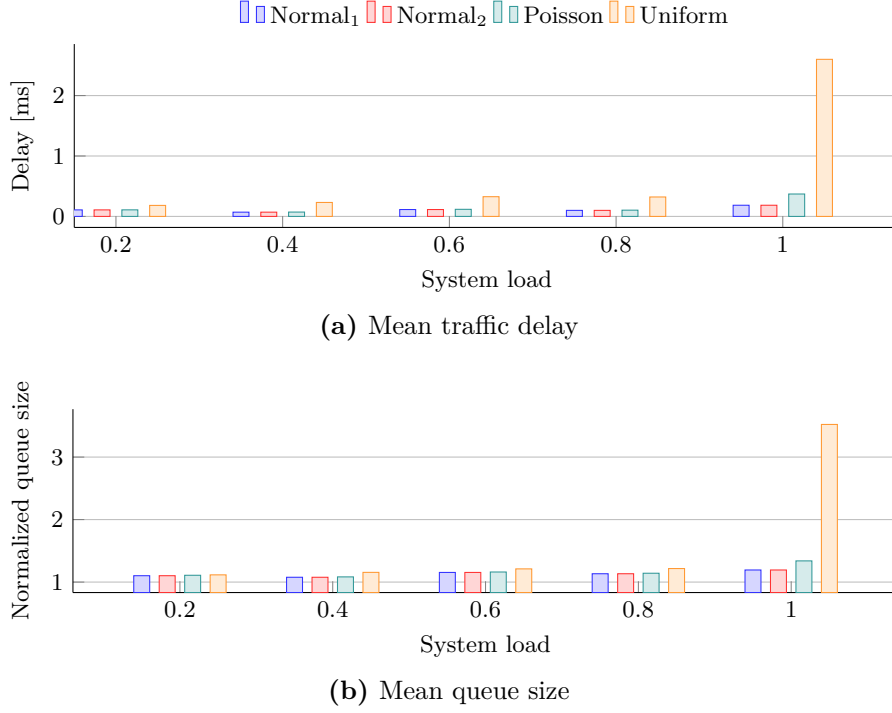
Figure 4.16: Stability metrics of the queues. Delay is measured as the time elapsed since a bit arrives to the queue until it is allocated. Queue size is normalized to the average arrival rate of the incoming queues

modifying the value of the update step α, see line 13 of Algorithm 2. This way, smaller α values would yield slower convergence time, while avoiding the peaky transitory. On the other hand, a more aggressive configuration of the update step reduces the convergence time at the expense of quite unpredictable behavior at the beginning, even when the remaining system capacity is high[1].

In order to validate the configuration of the objective function, Figure 4.18 illustrates the average RAT load obtained in the different scenario setups. As can be observed, the scheduler behaves according to the configuration favoring the allocation of resources from the macro cell, then the small one, and finally it allows the usage of the RAT based on mmW. Furthermore, we can see that this metric is rather independent of the traffic pattern, not showing any noticeable difference between the results obtained with the different traffic distributions.

As a complement the previous results, Figure 4.19 depicts the temporal evolution of the RAT load for different system loads. We have omitted the case of full load, since it does not provide any information. We can see that, when the system is lightly loaded, only resources from the LTE macro cell are allocated, while the other RATs are hardly used. When the system load increases, LaSR exhausts the resources of the macro cell and starts

---

1 All the results obtained for the evaluation of LaSR have been obtained with the configuration $\alpha = 1$.

(a) System load 0.2

(b) System load 0.4

(c) System load 0.6

(d) System load 0.8
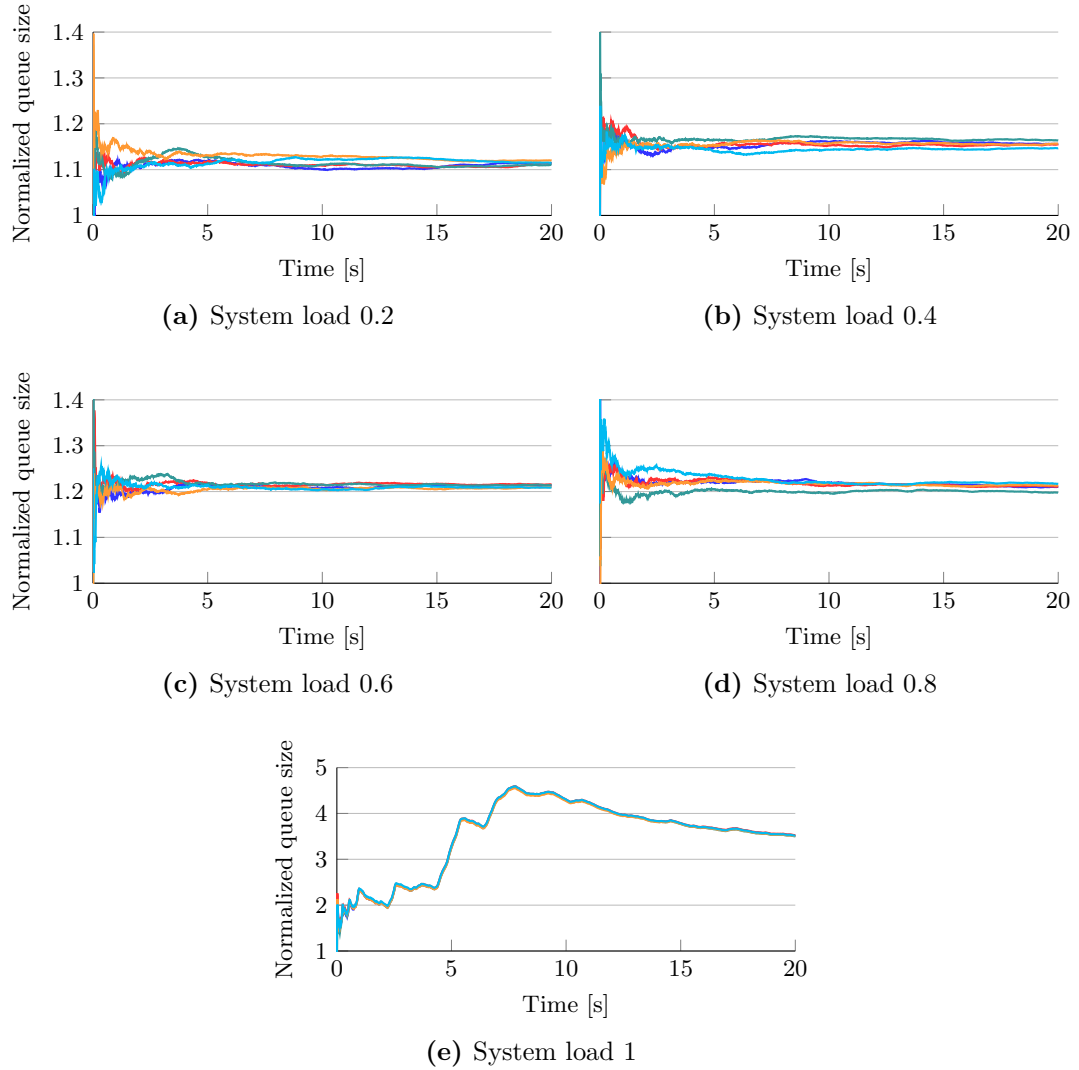
(e) System load 1

Figure 4.17: Temporal evolution of the normalized queues size upon different system loads when the *Uniform* distribution is used

using the small one. Likewise, the resources from the mmW technology are allocated when the other RATs are saturated[1].

Hereinafter, only *Poisson* traffic pattern will be use to assess other features of the LaSR scheduler.

---

1 Figure 4.19 shows a moving average to improve visibility. For this reason, it may seem that RATs are used even if others with lower cost are not saturated.

(a) Normal₁

(b) Normal₂

(c) Poisson
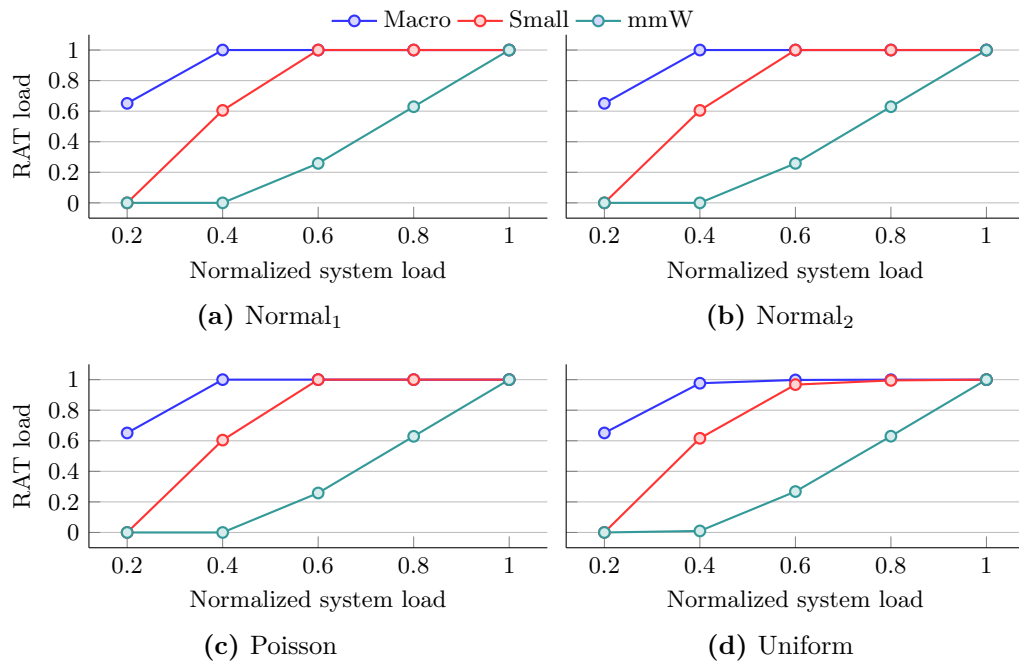
(d) Uniform

Figure 4.18: Average relative RAT load upon different system load configuration.

(a) System load 0.2

(b) System load 0.4

(c) System load 0.6
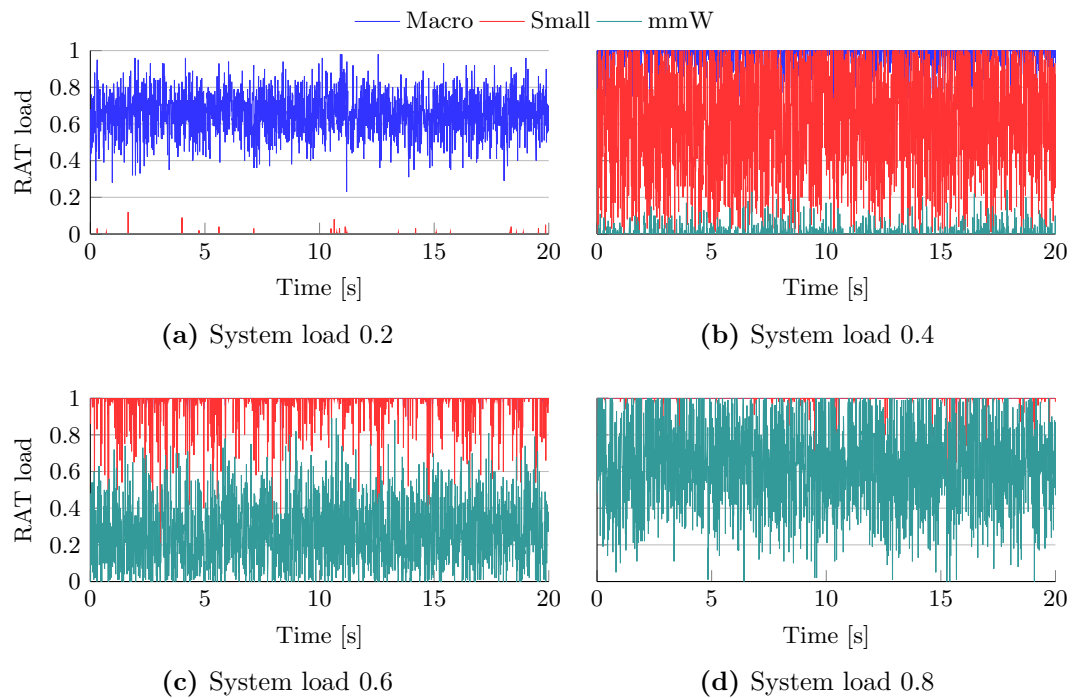
(d) System load 0.8

Figure 4.19: Temporal evolution of the normalized queues size upon different system loads when *Uniform* traffic distribution is used. For the sake of visibility, the graphs show the moving average the of RAT load with a window of 5

Quality of service

We now evaluate the ability of LaSR to satisfy different QoS requirements, implemented by the utility functions previously defined. In all cases, the objective function is tweaked so that the QoS is its main optimization parameter, see Eq. 4.55. In particular, we set ρ so that the relative weight of the QoS utility is thousandfold greater than the cost associated to RAT usage. In addition, the simulation time for each metric is the same as before, 20000 TTIs.

First, we study how non-elastic traffic may impact the system stability. To this end, we analyze a scenario comprising 5 traffic flows, from which 2 correspond to non-elastic traffic. We define the utility function according to Eq. 4.56, with parameters a = 1 Kbps and b equal to the average incoming traffic at each queue.

Figure 4.20 depicts the average delay and relative queue size experienced by the different traffic flows. In general we can assert that, for the defined scenario, our scheduler is able to provide the QoS required for non-elastic traffic (and also non-delay-tolerant) while keeping the system stability. In fact, we can observe that the delay for the non-elastic queues is almost zero, regardless of the system load, and the relative queue size is therefore around the unit. On the other hand, the results indicate that the elastic traffic, labeled as *best effort*, suffers higher delay than the one observed in previous results. As expected, LaSR prioritizes the allocation of non-elastic traffic, and it uses the remaining resources for the best-effort one.

Then, we have analyzed the response time of LaSR when non-elastic traffic appears in the system. In this sense, we have deployed a scenario similar to the previous one has, but in this case the non-elastic traffic enters the system during the simulation (at TTI 10000), so that the aggregated traffic almost reaches the total system capacity. The reason not to show total saturation is the long time required for the system to get stable.

Figure 4.21 shows the temporal evolution of the delay. This metric is computed every TTI as the time the bits transmitted in the corresponding TTI have stayed in the queues. As can be observed, the non-elastic traffic hardly experiences delay throughout the whole simulation. Conversely, the delay experienced by the best effort traffic increases when the system gets saturated. Nevertheless, we can observe that the delay gets stable within the simulation time. Although particular behaviors, and response times, strongly depend on the scheduler configuration, the results demonstrate that, even when it is close to the system capacity limit, LaSR is able to provide QoS, while maintaining the system stability.

In the following we study the algorithm behavior when using Eq. 4.57 as utility, in order to improve the fairness among traffic flows. In this case, we have modified the scenario so as to create a situation where fairness among flows can be better analyzed. The scenario comprises 20 traffic flows and 1 LTE macro-cell with 100 PRBs, so that each VRB correspond to exactly one PRB. In addition, we divide the flows in 2 groups with different channel conditions. The first group, made of 10 flows, have poor average channel conditions, so they can only use modulation indexes $\{11, 12, 13\}$, whereas the remaining flows experience better link conditions and they can thus use higher modulation indexes $\{21, 22, 23\}$ [TS36213]. In addition, each flow generates traffic with an average value of 1 Mbps and *Poisson* distribution and we set the configuration parameters c_i = 100 and

(a) Traffic delay

(b) Relative queue size
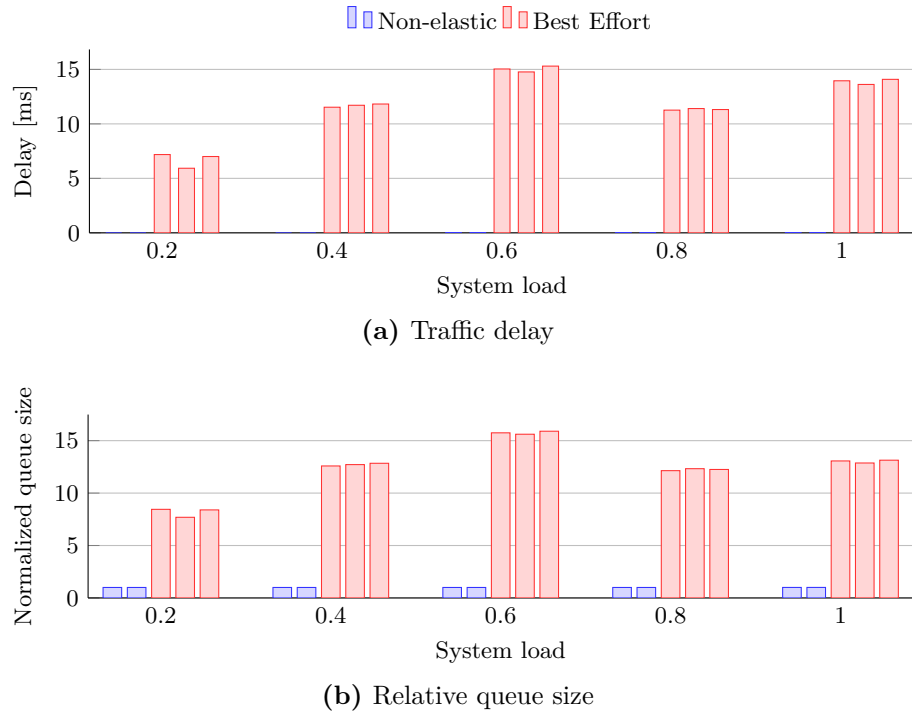
Figure 4.20: Stability analysis in a static scenario with inelastic and best effort traffic. Both traffic delay and relative queue size is shown for different system load regimes
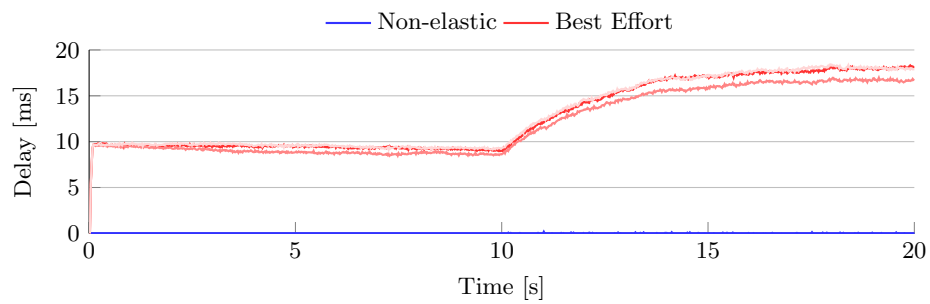
Figure 4.21: Temporal evolution of the traffic delay in a dynamic setup consisting of 3 best effort and 2 inelastic traffic flows. Inelastic traffic starts in the middle of the simulation ($t = 10000$)
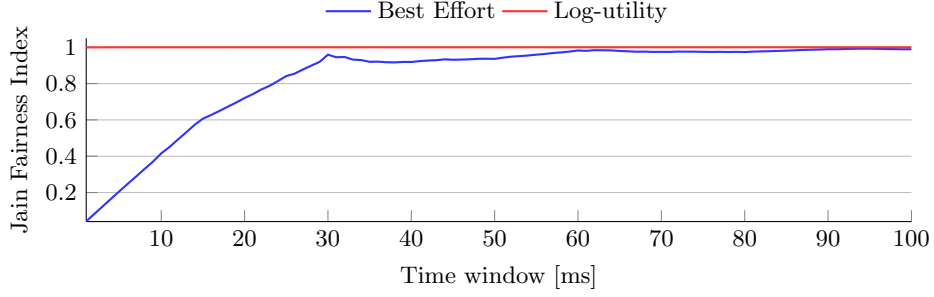
Figure 4.22: Jain Fairness Index [Jai84] of the users throughput measured along different time windows. Short(long) time windows represent a measure of short(long)-term fairness

$\hat{r}_i$ = 20 Mbps.

Over this scenario we evaluate both the short and long term fairness. To this aim, we measure the individual throughput achieved by each user (with equal long-term mean data rate requirements) in different time windows, and calculate Jain's fairness index [Jai84][1]. Computing a fairness index across short time windows helps us to understand short-term fairness performance [Bar02].

In this sense, Figure 4.22 illustrates how LaSR performs when using both logarithmic utility and best effort configuration. According to the results, the usage of logarithmic utility yields high fairness index regardless of the time window size. In addition, the results also show that, without logarithmic utility, LaSR brings about poor short-term fairness and high fairness in the long-term. It is expected behavior since the algorithm guarantees queue stability and all the queues have the same, average, incoming rate.

Mobility scenario

So far we have analyzed the scheduler behavior on various mostly static scenarios. We now evaluate the LaSR response to dynamic scenarios, where either the system load varies, as a consequence of the appearance of new flows (e.g. addition of new users), or the channel conditions change due to, for instance, mobility. Since they are dynamic scenarios, simulation times are defined according to the dynamic effects we want to study, so that specific temporal values will be provided in the results.

The first scenario uses the same network setup as the static ones, namely 3 RATs with the configuration described in Table 4.4 without using QoS utilities (ρ = 1). Over this scenario the load varies over time by emulating users joining and leaving. We start from an empty scenario in which new flows are added every 100 ms, leading to a load increment of 2.5% of the system capacity per flow. Then, after a short time the capacity boundary is reached, flows start leaving the system at the same pace they arrived.

Figure 4.23 depicts the usage of the RATs resources, in two different ways. First, Figure 4.23a shows how the load of the network elements evolves over time. For comparison

---

1 Jain's fairness index is a well known metric to evaluate how *equitable* a resource allocation is, spanning from 1 (equal share) to $1/n$, where n is the number of users.
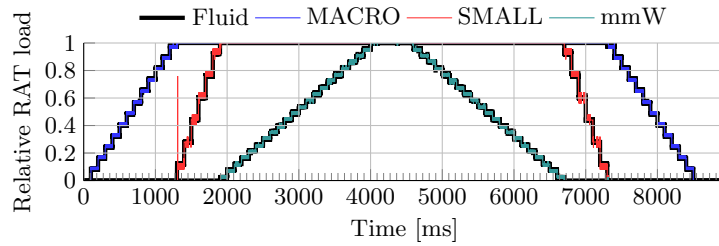
(a) Normalized RAT load over time of the different RATs and the unconstrained solution, labeled as *Fluid*
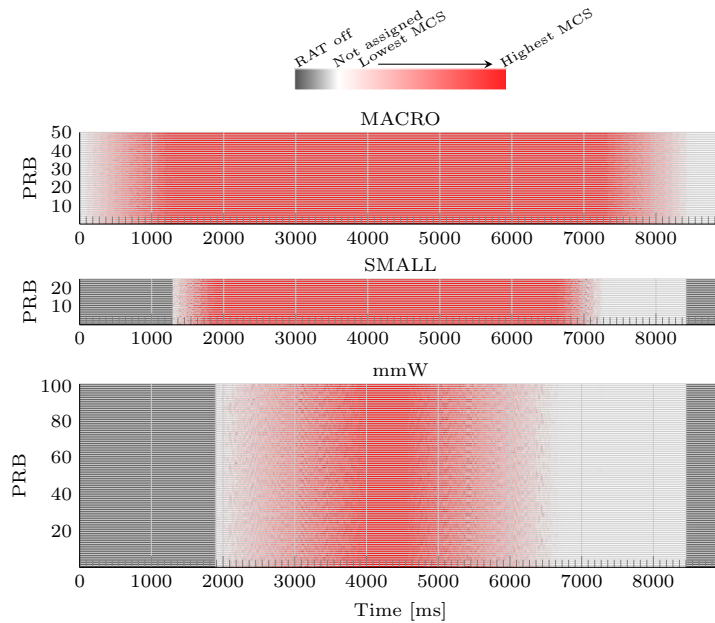
(b) Temporal evolution of the MCS selection for the different RATs. Color scheme indicates the state of the PRB: black if the RAT is off, white if the RAT is on but resource is not assigned, and different red color intensities according to the selected MCS

Figure 4.23: LaSR response to load variation with costs defined in Table 4.4

purposes, the figure also illustrates the optimum bound (i.e. fluid solutions) brought about by Problem 8. Then, Figure 4.23b depicts how the scheduler chooses the MCS levels from those defined in Table 4.4.

First, we note that there is a rather tight match between the scheduler outcome and the optimum solution. In addition, we observe that LaSR is able to quickly adapt to load fluctuations, so that we can perfectly differentiate when new flows join and leave the system. Finally, and aligned to previous results, we can observe how the costs given to the RATs permit us to control the usage of resources. In fact, we see that, for this configuration, LaSR uses resources from the macro cell until it is fully loaded. Then

it begins assigning resources from the small cell, which quickly saturates due to its low bandwidth, and eventually the scheduler starts using mmW.

Furthermore, Figure 4.23b illustrates the usage of resources by means of a time/frequency grid with the state of each PRB of the corresponding RAT. Similarly to the previous results, this figure shows how the secondary RATs (labeled as SMALL and mmW) are disabled until the macro cell is fully loaded. It is worth noting that the allocation of resources alternates between not-assignment and highest modulation level. This effect is reflected in the figure by lack of gradient color. The behavior is due to the configuration described in Table 4.4, which does not distinguish between MCS levels within the same technology. In this sense, the default configuration used during the LaSR analysis fosters an overall reduction of the number of allocated resources.

In order to study how the costs associated to the MCS levels affect the usage of resources, we have repeated the previous analysis with a configuration where modulation levels of a particular technology are given different costs, as described in Table 4.5. Differently to the default configuration, which only established priorities between technologies, the new configuration fosters the selection of low modulation indexes, regardless of the technology.

In contrast to the previous results, Figure 4.24 does not show a gradual increase of the RAT load but most resources are allocated as soon as the technology is needed. In addition, we can also observe that resources of secondary RATs start being allocated before. Indeed, according to Figure 4.23, the resources of the mmW technology were first allocated around TTI 2000, while in this case it happens before TTI 1000.

This behavior is better explained by Figure 4.24b. We observe that, although all resources are used, low or medium MCS levels are selected until all access elements are saturated. In this way, at the expense of more allocated resources, we obtain a better distribution between the RATs. According to the configuration, the results yield that, as soon as the macro cell gets saturated with its lowest MCS, resources from the small cell start being allocated. In a similar way, the mmW technology is used before the highest modulation level of other technologies are selected. This highlights the versatility of LaSR to foster tailored resource utilization according to different policies.

We present next a mobility experiment with two access elements, $Cell_1$ and $Cell_2$ with identical characteristics. In this case, the scenario embraces 5 traffic flows, where users move together (e.g. commuting) from the proximity of $Cell_1$ towards $Cell_2$, with a constant traffic demand, and the cost configuration indicated in Table 4.4.

Furthermore, at each time slot, all users have the same 3 available MCS. This way, at

Table 4.5: System configuration with MCS dependent cost. The spectral efficiency has units of bps/Hz and the cost of bits^{-1}

	Spectral efficiency/Cost		
Macro	0.2344/1	2.1602/5	5.5547/10
Small	0.2344/2	2.1602/10	5.5547/20
mmW	0.5/6	2.0/30	4.875/60
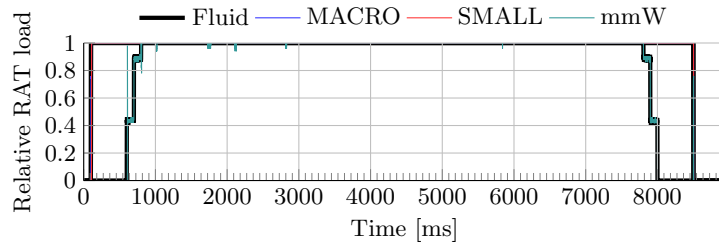
(a) Normalized RAT load over time of the different RATs and the unconstrained solution, labeled as *Fluid*
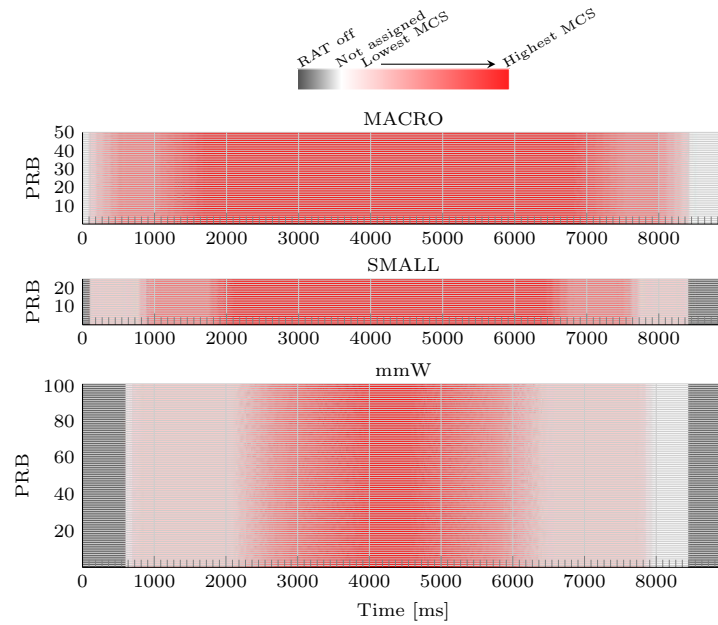
(b) Temporal evolution of the MCS selection for the different RATs

Figure 4.24: LaSR response to load variation with costs defined in Table 4.5

the beginning of the simulation, users have the best possible channel conditions with $Cell_1$, and, accordingly, the set of available MCS is the highest from LTE specification [TS36213], namely indexes $28, 27, 26$. On the other hand, we assume that users are out of the $Cell_2$ coverage range, so that the set of modulations with this cell is empty.

Then, we vary the set of modulations for both cells, every 500 TTIs, using a sliding window such that users get into the coverage range of the second cell at TTI 3000. This way, around TTI 9000 the same channel conditions are experienced with both cells[1].

We show in Figure 4.25, the traffic rate served by each RAT, along with the aggregated

---

1 Between TTIs 8500 and 9000, the set of MCS indexes is 13,12,11 and 12,11,10 for $Cell_1$ and $Cell_2$, respectively. Afterwards, in the next 500 TTIs, the sets are the other way around.

(a) Traffic rate with LaSR

(b) Traffic rate with QSG

(c) Relative RAT load with LaSR

(d) Relative RAT load with QSG
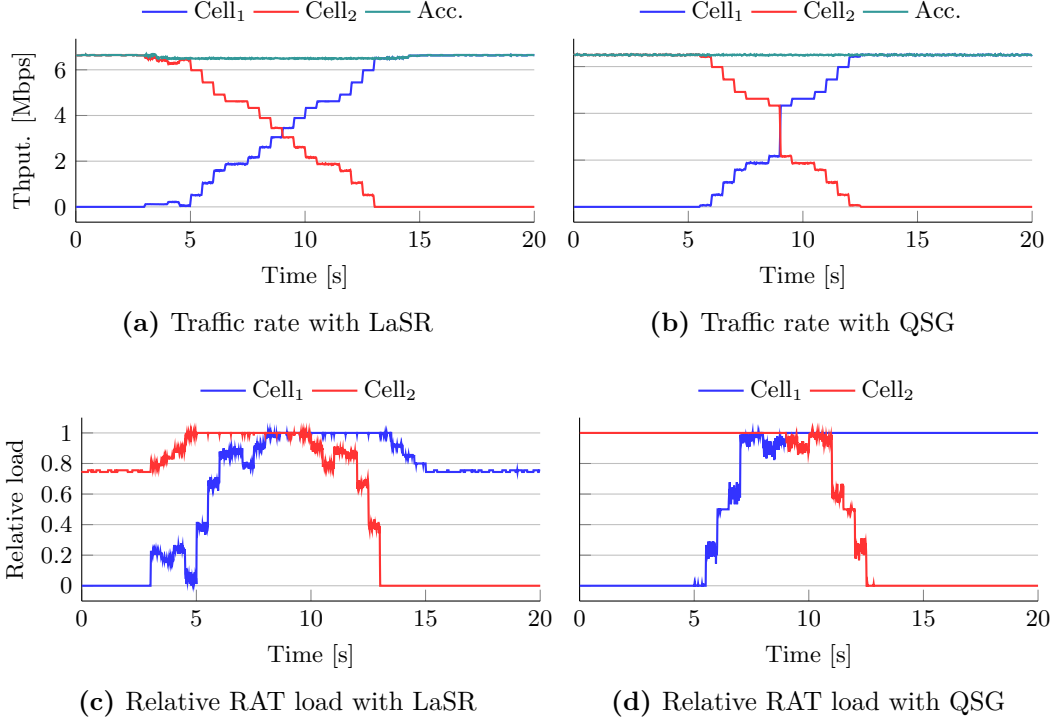
Figure 4.25: Users join/leave every 100 ms

one, and the normalized load of both cells over time. In this case we compare the behavior of LaSR with that obtained when using QSG. The comparison with QSG has not been done before due to the lack of support for utility/cost functions of this algorithm.

The results evince that both algorithms satisfy the throughput requirements of the system (around 6.5 Mbps of aggregated throughput). However, as can be observed in Figure 4.25, QSG achieves this at the expense of higher utilization of resources, due to its greedy nature. In contrast, LaSR is able to balance the load between the two RATs as needed, and minimizes the allocated resources with no compromise in throughput.

System constraints

We now analyze the behavior of LaSR considering additional constraints derived from the design of the communication system. In order to broaden the multi-RAT analysis performed in this section, we further study the impact that the VRB design has over the scheduler design.

Furthermore, according to the reference scenarios depicted in Figure 4.11, LaSR has been designed as a centralized scheduler, either for carrier aggregation or cooperative scenarios. Although centralization permits tighter cooperation between network elements, it can also cause additional delays, so that the information at the scheduler may be inaccurate. This effect may be particularly important in C-RAN scenarios. In this sense, we will also analyze the impacts of delay on the queues backlog information. In both cases, the performance of

LaSR will be compared with that provided by the QSG algorithm.

The definition of the system VRB is studied over a simple scenario embracing one single RAT with 64 PRBs per TTI. We impose different degrees of granularity to the scheduler. To this end, we define ζ as the size of the VRB, so that it indicates the amount of PRBs that must have the same assignment and modulation within a TTI. For instance $\zeta = 1$ would indicate that each VRB has just one PRB, thus allowing individual assignment of physical resources. On the contrary, $\zeta = 64$ would force to assign all the physical resources at once.

In Figure 4.26 we compare the utilization of physical resources when using LaSR and QSG. In the case of our algorithm, we have applied two different configurations, to observe the impact of the proposed cost associated to the resources.

The results evinces that, while QSG always over utilizes resources for low and mid-load ranges, LaSR matches well the utilization of physical resources to the average demand when it is required. A greedy scheduler, such as QSG, tends to use the full system capacity, since its goal is to minimize the queues occupancy. Therefore, it allocates resources as soon as the queues are not empty. In turn, this greedy behavior adds padding in most of the resources. We can also observe that this effect is more clear as the system load increases.

In contrast, LaSR is able to adapt the utilization to the actual demand of resources at the cost of user delay. Note that, conversely, when we reduce the cost of using physical resources, see $LaSR_2$, the utilization is higher. This illustrates the potential adaptation of LaSR to different scenarios, according to the particular requirements.

We next illustrate the robustness of LaSR to imperfect backlog information. To this aim, we make use of the default scenario, in which some delay Δ_{bl} (TTIs) is added to the backlog information. Figure 4.27 depicts the mean queue size of the users overtime, for both LaSR and QSG, when we modify the system load.

First, we can observe that QSG is very sensitive to uncertain queue information, regardless of the system load, and it is not able to recover over time. In contrast, LaSR barely suffers of performance degradation at low and mid-load regimes, being able to reduce the queue size (and so traffic delay) over time. From the figure, we can observe how the performance worsens as control information delay and mean load increase.
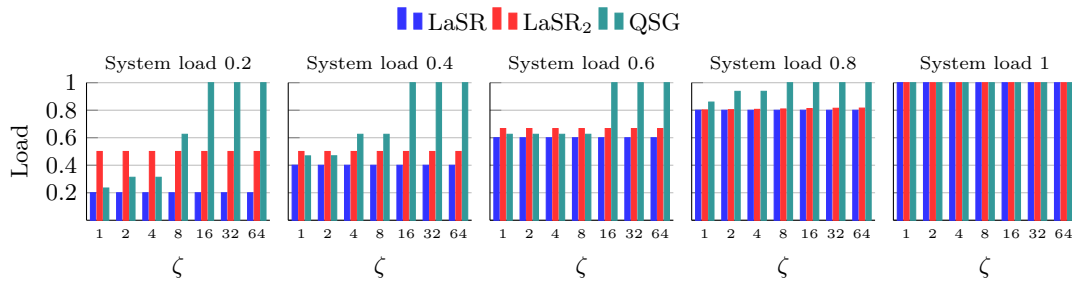
Figure 4.26: Average system load as a function of ζ. The figures show the metric upon different system load regimes. The results labeled as $LaSR$ uses a cost of 100 bit^{-1}, as those labeled as $LaSR_2$ have an associated cost of 1 bit^{-1}

(a) LaSR

(b) QSG
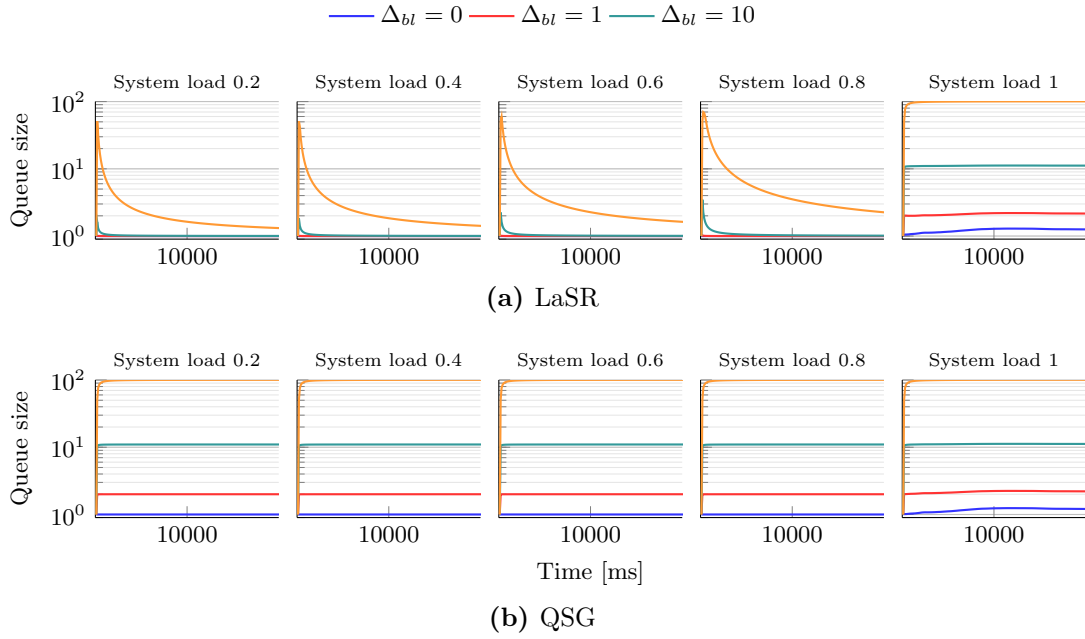
Figure 4.27: Mean user queue size in the default scenario with delayed backlog information equal to Δ_{bl} TTIs

4.6 Conclusion

The increasing demand of data traffic over wireless networks will lead to the deployment of very dense and heterogeneous networks, giving rise to multi-connectivity scenarios where access elements of the same or different technologies will provide seamless connectivity services. In these novel scenarios, access selection and resource management techniques will become both more complex and flexible. At the same time, the high density of network elements, and the way they are exploited, will have a remarkable impact over the EMF population exposure.

Exploiting the simulation methodology presented in Chapter 3, we have analyzed existing techniques from the EMF viewpoint, and we have proposed new ones, able to tackle the requirements of future scenarios.

Traditional simulation models do not have into account the system dynamics when modeling physical resources. In particular, network load and the impact that access selection policies have over resource efficiency, due to mutual interference, is usually neglected. In this sense, the first contribution of this chapter has been the definition and evaluation of simulation models that consider mutual interference in the access selection process, so as to accurately mimic the system behavior.

Later on, the simulation models have been exploited to analyze heterogeneous and dense scenarios, which are believed to have a high impact on both downlink and uplink EMF exposure. First, traffic offloading to WiFi has been analyzed, with different access point densities, and considering both the EMF exposure and QoS. This evaluation has shown a significant decrease of the population exposure when WiFi connections are favored. On the

other hand, we have also observed that operator policies can greatly impact the quality of services in those cases where the additional capacity brought by WiFi is not exploited.

Then, access selection techniques have been analyzed over dense LTE scenarios. In this case, since the proximity to access points is likely to reduce the downlink exposure, the analysis has focused on the potential improvement on the uplink side. In this sense, the evaluation has shown that new access selection policies, such as CRE or DUDe, can significantly reduce the exposure perceived by the population. Nevertheless, constant bias values used in CRE to foster the usage of small-cells can greatly impact the QoS. On the other hand, by leveraging the total decoupling of downlink and uplink selection procedures, DUDe is able to keep the exposure low while ensuring that the QoS is not affected. This improvement is achieved at the expense of increasing the number of handovers, which might also jeopardize the QoS, depending on the system architecture.

Finally, the last section of this chapter has been devoted to the definition and evaluation of a resource management scheduler specially designed for multi-connectivity scenarios. Exploiting the mathematical framework presented in [Val17], the LaSR scheduler is able to accommodate system constraints or operator policies while ensuring stability even close to saturation conditions. The scheduler has been analyzed in a number of scenarios. According to the results, LaSR can accommodate operator policies that foster the usage of one technology over the others, which might have a strong impact on the population exposure. At the same time, the scheduler is also able to provide QoS to different types of traffic, or to improve fairness among users.

CHAPTER 5

EMF aware network management techniques

While resource management and access selection play an essential role to mitigate the population exposure induced by cellular networks, new topologies and the evolution of services require the analysis of other layers in the protocol stack, above MAC.

In this sense, in this chapter we will propose modifications for two network management techniques that may highly impact the exposure. On the one hand, we assess the impact that alternative topologies could have over the population exposure, in particular multi-hop wireless networks. Furthermore, considering the rapid growth of data traffic, it seems also relevant to tackle improvements on content delivery techniques. In particular, we will focus on video services, since they are foreseen to carry most of the data traffic in the near future.

Although the research on routing protocols for multi-hop wireless networks started already in the mid 90's, in recent years there has been an upturn on the investigation about this area. Their original purpose, as communication alternatives in rather particular scenarios (for instance, natural disasters), has turned to more innovative applications and use cases. One key example is D2D communications, which has been specified by the 3GPP standards [Liu15] for commercial applications (i.e. proximity services). In addition, it is also worth highlighting the key role that multi-hop topologies can play for M2M communications [Dom13], and wireless sensor networks. Considering the significant impact that these new scenarios may have over the population exposure, due to the body proximity to the involved devices, it becomes necessary to study how EMF exposure can be taken into account within the routing process.

Traditional routing metrics have mainly focused on QoS parameters, such as delay, hop count or power [Gue07; You14], and they can be divided into two main categories [Iqb14]: single-radio and multi-radio. From the first category, the Expected Transmission Count (ETX) [Cou05] and Expected Transmission Time (ETT) [Dra04] are two of the most widespread alternatives, and they have been enhanced and broadened with the Weighted Cumulative Expected Transmission Time (WCETT) [Dra04], Exclusive Expected Transmission Time (EETT) [Jia07], and with other solutions that consider the interference. Metric for Interference and Channel Switching (MIC) [Yan05] or Interference Aware (iAWARE) [Sub06] are examples belonging to the second category.

There are, as well, routing protocols that focus on load balancing, which might have a strong effect on the distribution of the exposure over a particular area. The reader might refer to the survey by Toh et al. [Toh09] for a succinct discussion of the most interesting load balancing proposals. They classify the existing solutions into three main groups,

depending on the metric they foster. The traffic-based approach looks at a dynamic load balancing. For instance, Hassanein and Zhou [Has01] proposed Load-Balanced Ad-hoc Routing (LBLAR), which used the degree of nodal activity (number of active paths) as a novel routing metric. They showed that their proposal outperformed legacy approaches in terms of delay and throughput, although the paper does not study how the traffic is actually balanced between the nodes. Likewise, Altalhi and Richard [Alt04] proposed Traffic-Size Aware (TSA) routing protocol, which was based on the combination of the number of packets belonging to each node and to its neighbors. More recent works on load-balancing schemes for multi-hop and mesh networks are those from Li et al. [Li12], which use a greedy routing scheme, based on geographical information, and Yamamoto et al. [Yam12], which introduces a rate adaptation mechanism, instead of changing routes. To the best of our knowledge, while some protocols consider parameters that might shape the induced exposure, the EMF exposure, per se, has not been yet considered for routing solutions.

Presumably, the closest solutions to the EMF exposure are those under the Green Networking motto. The main idea behind it is fostering an energy-aware behavior, mostly by means of reducing the transmitted power, increasing at the same time the battery life-time. Most of the existing proposals deal with cellular communications [Aue11; Cor10], while not much attention has been paid to wireless multi-hop topologies.

As for data traffic, current predictions [Cis17] estimate that for 2021 the mobile data traffic will experience a sevenfold increase with respect to the values in 2016. From the different data services, forecasts predict that video streaming will represent, approximately, 78% of the Internet traffic in 2021 worldwide. Hence, it is relevant to study how current video content delivery mechanisms impact the EMF exposure and how they interact with the lower layers of the most widespread technologies, so as to propose alternatives to mitigate such exposure.

In this sense, Dynamic Adaptive Streaming over HTTP (DASH) [Sod11] stands out as one of the most popular protocols for video streaming over the Internet. For instance, it is used by *Netflix* [Adh12] and it is also implemented on the YouTube platforms. DASH has been developed to optimize the Quality of Experience (QoE) perceived by the users, adapting the delivery of audio-visual content to the particular constraints of both networks and terminals. It has been proved to appropriately adapt the delivery of contents to networks whose bandwidth fluctuates, as it is usually the case of wireless networks. Besides, DASH has been standardized by the 3GPP, as reported in [TS26244; TS26234].

All in all, this section will present and evaluate novel solutions for the two aforementioned network techniques: routing over multi-hop wireless networks and video content delivery in cellular networks. In both cases, exposure metrics, such as the LEXNET EI, are not well suited for legacy network management solutions. Having this in mind, it becomes necessary to define proxy metrics, such as transmission power or time required for a transmission, to estimate the potential impact over the population exposure. Accordingly, the evaluation of the proposed techniques will include the definition of suitable proxy metrics.

5.1 EMF aware routing for multi-hop networks

This section is devoted to the integration of the EMF exposure into the routing decision process. To this end, a theoretical framework is defined, considering the exposure from a routing perspective. Then, based on this study, a novel routing protocol, namely EMF Aware Routing Protocol (EARP), is developed.

Before defining an actual routing protocol, we formulate an optimization problem, and we propose an algorithm to solve it, setting the performance baseline for the protocol. Afterwards, we strictly define proxy indicators for the EMF exposure, which will be used to assess the performance of routing solutions. Then, based on the previous analysis and definitions, the routing protocol will be defined and evaluated. In particular, its performance is studied by means of an extensive simulation campaign carried out over the *ns3* [Ns317] framework, where we compare EARP with traditional approaches, based on the proxy indicators.

5.1.1 EMF aware routing algorithm

The routing algorithm is based on an optimization problem that takes into account the exposure in a twofold manner. On the one hand, the algorithm aims to minimize the transmission power when establishing the routes, since it is a key factor to reduce the EMF exposure according to the EI definition in Section 2.1.3. On the other hand, the reduction of transmission power, although leading to an overall exposure reduction, would not prevent the appearance of exposure peaks, as well as high variability between different network areas. Hence, it is also necessary to consider the accumulated exposure induced by the nodes.

Once the network model and optimization problem are defined, we propose an algorithm that can be used to solve such problem, exploiting graph theory and transformation techniques, which are introduced to take into account both the transmission power and accumulated exposure. Afterwards, the algorithm will be evaluated, providing an insight about the potential performance of routing protocols based on it.

Network Model

The network is modeled as a graph $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of vertices and $\mathcal{E}$ the set of corresponding edges. We assume that two nodes are connected if the euclidean distance between them is shorter than the coverage range of the underlying communication technology. The aforementioned metrics, accumulated exposure and transmission power, can affect either nodes or edges of the graph as follows:

- Node accumulated exposure ω_{acc}: this cost considers the overall exposure that has been induced by a particular node up to a certain time. It is worth highlighting that it is independent of the amount of traffic traversing a node at a particular time, according to the routes at that moment, since it represents the accumulated exposure.

- Edge power ω_P: this reflects the transmission required to send a packet to a neighbor. It is therefore assigned to every edge within the network. In this sense, ω_P scales up according to the traffic assigned to the link.

(a) Initial Graph

(b) Transformed Graph
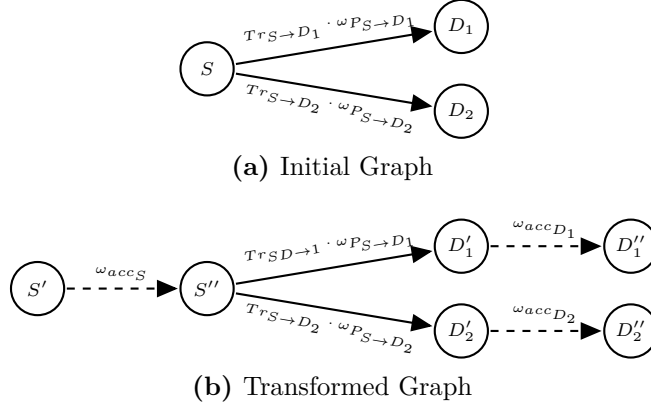
Figure 5.1: Simple example of graph transformation with 3 nodes. Initial graph G and transformed graph $G^{'}$ are shown in Figures 5.1a and 5.1b respectively

Optimization algorithms applied to graphs usually assume that cost functions are applied to edges proportionally to the flow traversing them. Since the approach presented here assigns costs to both nodes and edges, network transformations are needed so that the resulting network fits into well known algorithms. To this end, we introduce virtual edges to bear the cost ω_{acc}. In this sense, each node n is split into two virtual vertices n' and n'', ingress and egress respectively, which are then connected with a virtual edge having a cost ω_{acc_n}. The incoming edges at node n are now connected to n', while the outgoing ones start at n''.

As an illustrative example, Figure 5.1 shows the transformation applied to a simple graph of three nodes. Here, $Tr_{i \to j}$ and $\omega_{P_{i \to j}}$ are the traffic volume and the cost related to the transmission power between nodes i and j respectively. On the other hand, ω_{acc_i} models the accumulated exposure induced by node i up to the current instant. As shown in Figure 5.1b, initial nodes are split so that virtual links bear the cost related to the accumulated exposure. Besides, the figure shows that the proposed transformation increases the complexity of the graph. In particular, from an initial graph $G = (\mathcal{V}, \mathcal{E})$, a transformed one $G^{'} = (\mathcal{V}', \mathcal{E}')$ is obtained, where $|\mathcal{V}|^{'} = 2|\mathcal{V}|$ and $|\mathcal{E}'| = |\mathcal{V}| + |\mathcal{E}|$.

Optimization Problem and algorithm

Once the network model has been defined, an optimization problem is posed to establish routes, considering both instantaneous power and accumulated exposure. Since the overall cost is proportional to the flow traversing an edge, we can state a problem similar to a Minimum Cost Flow (MCF) [Ahu93], whose main goal is to distribute the overall traffic throughout the graph, so that the global transportation cost is minimized. Thus, the problem can be analytically expressed as follows:

<div align="center">

Problem 9 (Minimum Cost Flow)**.**

</div>

$$\boldsymbol{min} \quad \sum_{\forall (i,j) \in \mathcal{E}} \omega_{ij} \cdot x_{ij} \tag{5.1}$$

$$\boldsymbol{s.t.} \quad \sum_{j:(i,j) \in \mathcal{E}} x_{ij} - \sum_{j:(j,i) \in \mathcal{E}} x_{ji} = Tr_i \qquad\qquad \forall i \in \mathcal{N} \tag{5.2}$$

$$c_{ij} \geq x_{ij} \geq 0 \qquad\qquad \forall (i,j) \in \mathcal{E} \tag{5.3}$$

$$\sum_{i=1}^{n} Tr_i = 0 \tag{5.4}$$

where ω_{ij} and c_{ij} represent the cost per flow unit and the edge capacity, respectively. Besides, Tr_i accounts for the traffic sent ($Tr_i > 0$) or received ($Tr_i < 0$) at each node, and x_{ij} is the traffic flow traversing every link of the graph. Constraint 5.2 corresponds to the flow balance at every node (if Tr_i equals 0, the node is neither a source or a sink). Besides, constraint 5.3 ensures that edge capacities are respected, while constraint 5.4 enforces that the global flow balance is maintained.

Various algorithmic solutions exist for the previous problem. Since the main contribution os this section is the definition of the network model rather than a particular technique to solve it, a classical solution, namely the Cycle Canceling Algorithm (CCA), is adopted. This algorithm leverages the properties of cycles in a graph. In particular, the algorithm exploits the optimality property, which establishes that a feasible solution $x^\star$ is optimal if, and only if, the residual graph $G(x^\star)$ has no negative cycles [Ahu93]. Its basic operation is described in Algorithm 4, where r_{ij} is the residual capacity of the arc (i,j) and x represents the flow allocation. The algorithm starts by establishing a feasible allocation and studying the presence of negative cycles by studying the residual graph. Then more flow units are allocated according to the minimum capacity of the negative cycle. The process continues until no negative cycles exist. The reader may refer to Appendix B for a practical example of the algorithm execution.

<div align="center">

Algorithm 4: *Cycle Canceling* Algorithm

</div>

---

procedure CCA($G(\mathcal{N},\mathcal{E})$)
    Establish one possible allocation for the flows x ▷ Any arbitrary solution
    while $G(x)$ contains a negative cycle **do**
        Identify one negative cycle $\mathcal{W}$
        $Tr_r := \min\{r_{ij} : (i,j) \in \mathcal{W}\}$
        Increment Tr_r flow units in the flow $\mathcal{W}$ and update $G(x)$
    end while
end procedure

The aforementioned algorithm was originally introduced for networks having one single source-destination pair. In order to handle more realistic scenarios, where a number of nodes (gateways) might provide access to external networks, we propose adding two
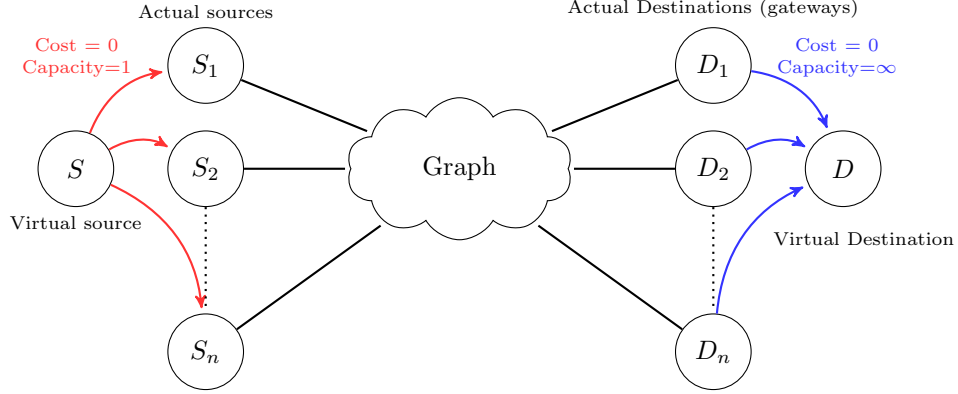
Figure 5.2: Addition of virtual origin and destination to accommodate arbitrary graphs to single origin/destination algorithms

additional virtual nodes. As depicted in Figure 5.2, a virtual source is connected to the actual traffic sources, while real gateways are attached to the virtual destination. Without loss of generality, each flow is assumed to bear one traffic unit. We fix the capacity of links between virtual and real sources to one traffic unit, ensuring that sources are not generating more than one flow. Besides, there is no capacity limit on the links between gateways and virtual destinations. Also, it is assumed that the additional links from and to the virtual source and destination do not have any cost, so that they do not impact the optimization algorithm. As can be observed, although we could consider an arbitrary number of sources and gateways, the algorithm would just need to solve a problem with one single source-destination pair.

As mentioned before, the problem is modeled as a MCF one, where the cost of the edges is proportional to the amount of traffic allocated to them. Nevertheless, according to the graph transformation presented before, the cost of the virtual edges (the ones connecting ingress and egress virtual nodes) is not proportional to the flow but constant, since it is meant to model the accumulated exposure. In order to ensure that the cost related to the exposure remains constant, we define the generic cost ω_{ij} as follows:

$$\omega_{ij} = \begin{cases} \rho \cdot P_{ij} & \text{for original edges} \\ \frac{\omega_{acc_i}}{\sum_{n \in \mathcal{E}} x_{in}} & \text{for new (virtual) edges} \end{cases} \tag{5.5}$$

where P_{ij} is the transmission power used by node i to reach node j, ρ is a constant factor and $x_{i,n \in \mathcal{E}}$ is the amount of flow traversing the node i. As can be observed, the cost of virtual edges is normalized by the flow traversing the nodes, so that it remains constant once multiplied by the traffic flow. In order to implement such modification, we heuristically modify the cost of the virtual links during the algorithm execution. Although this could break the isotonicity of the corresponding graph, the obtained results empirically show that this approach was appropriate given its characteristics.

Algorithm Evaluation

The algorithm has been evaluated in two different ways, comparing its performance with a legacy solution used as reference. In particular it has been compared with a traditional minimum cost search algorithm, where costs are assigned to links according to the required transmission power. This algorithm is a shortest-path search, in which the cost of the edge represents the transmission power to get the next hop, without taking into account the amount of flow used by that particular edge.

In the first set of results, we have assessed the behavior of the proposed solution with a static configuration over a number of scenarios with different topologies, so as to study the impact that the network density might have on it. Then, the evaluation has been extended, analyzing how the configuration of the cost associated to accumulated exposure impacts the algorithm performance.

In all cases, slotted time is assumed and the algorithms are sequentially applied to the graph at the beginning of each slot. Then, at the end of the slot, the information regarding the accumulated exposure is updated, to appropriately capture its temporal evolution. This way, the entire procedure will reflect the behavior of an eventual routing protocol. This methodology allows the use of an ad-hoc implementation, that considers the temporal evolution due to different traffic flows, without using heavier tools (such as network simulators), that would greatly impact the computation time. In order to ensure the statistical validity of the results, 100 independent experiments, comprising 100 temporal slots, are executed for each configuration.
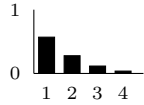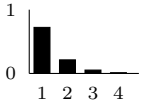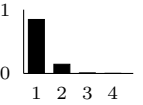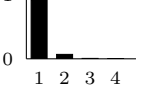
The evaluation of the algorithms has been carried out using metrics that account for both transmission power and accumulated exposure. The former is modeled as a linear function of the distance between nodes. The latter is defined by a generic parameter ϵ, which represents the increment of accumulated exposure in a node when it is part of one active route during a time slot. In particular, we will assume that all flows are alike, hence the accumulated exposure of nodes belonging to one route during T time units would increase $\epsilon \cdot T$. Similarly, if a node is selected to take part in 2 routes during T time slots, such node would induce an accumulated exposure $2\epsilon \cdot T$ and so on.

As for the costs, in all the experiments, the one associated to the transmission power, ω_P, is equal to the power metric. This is, linearly proportional to the distance between the involved nodes. Concerning ω_{acc}, which accounts for the accumulated exposure, its value has been tailored for each scenario. It is worth noting that it can be indeed adjusted to foster different algorithm behaviors.

As a common setup, 50 nodes equipped with wireless technology and a coverage range of 15 m, are randomly deployed over a squared area following a Poissson Point Process (PPP). According to this configuration, both the power metric ω_P, and the corresponding cost, take values in the range $[0, 15]$, as a linear function of the distance, so that the longer the distance, the higher the power required to send packets over such link.

The different configurations of the first set of results are summarized in Table 5.1. On top of the common setup, 4 different network topologies are defined, where the side of the scenario is modified so that we can analyze the impact of the network density, as well as the connectivity degree of the graph.

Table 5.1: Configuration of the static topology. For each topology To_1, To_2, To_3, To_4, the PMF of the number of gateways that can be reached is depicted

	To_1	To_2	To_3	To_4
Sources	{2,4}	{2,4}	{2,4,6,8}	{2,4,6,8}
Area m^2	120×120	130×130	140×140	150×150
PMF of reachable gateways				

For each topology the results are obtained with different number of traffic sources, while it is always assumed that 4 traffic destinations are deployed following a grid pattern, thus covering the maximum area. The routes can be established towards any of the 4 possible destinations. Furthermore, the cost associated to the accumulated exposure ω_{acc} has been defined to be equal to ϵ, which defines the exposure metric increment.

First, we studied the accumulated exposure at the different nodes during one experiment. Figures 5.3 and 5.4 show, for both algorithms, the Probability Mass Function (PMF) of this metric, considering all the nodes and when using topologies To_1 and To_2 respectively. To better understand the graphs, we may consider that the value $\epsilon \cdot T$ corresponds to the exposure accumulated by a single node if it were active in one path during the whole experiment (duration T). Hence, a multiple of $\epsilon \cdot T$ would reflect nodes that were active in more than one route. For instance, if the PMF at $2\epsilon \cdot T$ equals 0.1, that would imply that 10% of the nodes have been active in two routes during the whole experiment.

In both figures, we can observe that the minimum power algorithm shows a discrete behavior, since the PMF is greater than 0 only in multiple values of $\epsilon \cdot T$. As expected, it indicates that routes are kept static during the whole experiment, not changing in spite of the high accumulated exposure. On the other hand, the proposed EMF aware algorithm balances the exposure. For instance, in Figure 5.3 we can observe a relevant decrease on the PMF for $\epsilon \cdot T$, which is around 30% and 20% for the scenario with 2 and 4 sources

(a) 2 sources   **(b)** 4 sources
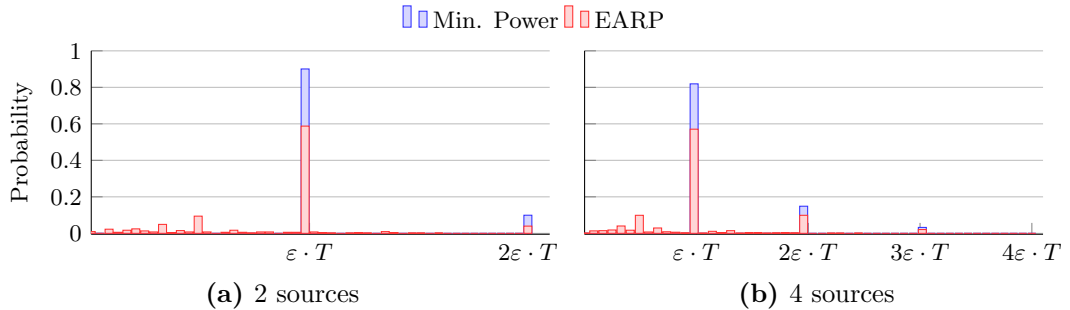
Figure 5.3: PMF of the exposure at each node in topology To_1

(a) 2 sources **(b)** 4 sources
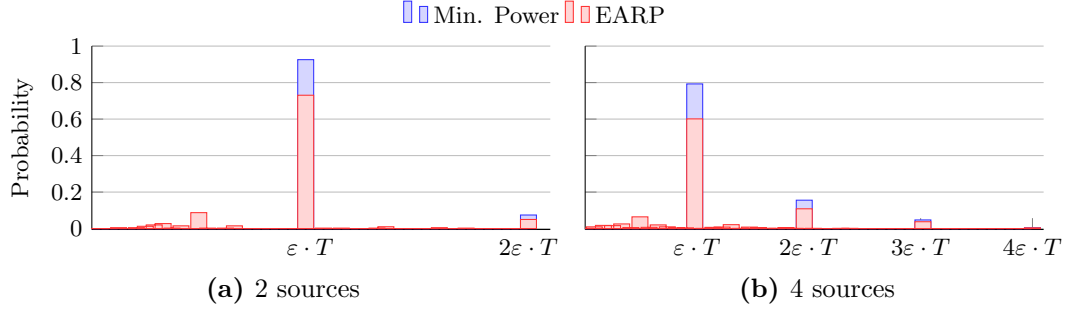
Figure 5.4: PMF of the exposure at each node in topology To_2

respectively. This decrease is reflected on the smaller bins that can be seen for values within the interval $[0, \epsilon \cdot T]$, which do not appear in the legacy solution. Furthermore, if we compare both topologies, we can also observe that the density of the scenario has an impact on the results. In this sense, the less dense topology, Figure 5.4, yields a slightly lower difference between both algorithms as a consequence of a reduction in the number of potential routes.

Once assessed the behavior of the proposed algorithm, the analysis is broadened to consider different topologies and traffic loads. In this sense, Figure 5.5 shows the average value of the exposure after 100 independent experiments. This metric has been calculated considering only nodes that have induced exposure, so that it gives an idea of how the accumulated exposure is balanced. The bars correspond to the values observed for the minimum power solution, while the horizontal line represents the results of the proposed algorithm. As can be seen, the gain reduces as long as the node density decreases. Indeed, for dense scenarios the number of potential routes is higher, and, as a result, there are more alternatives to exploit the wider optimization goal that is sought by the proposed solution. Aligned to this, the gain for topology To_1 is greater than the one observed for To_4. On the other hand, the figure also yields that there is not a clear impact of the number of traffic flows (simultaneous number of sources) on the difference between both algorithms.
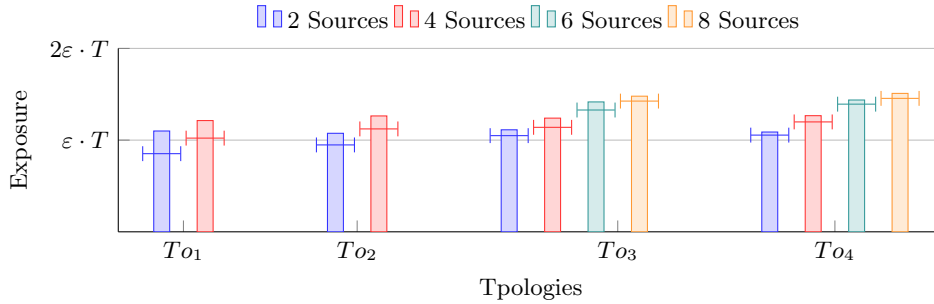
Figure 5.5: Average exposure per experiment. The markers correspond to the value obtained when using the EMF aware algorithm, while the columns represent the results of the *Min Power* alternative

Furthermore, Figure 5.6 shows the average value of the overall required transmission power per route. It is worth recalling that, in the case of the minimum power algorithm, the transmitted power does not change during the experiment, since the routes remain static, while this might not be the case for the EMF aware approach. Note that when a change of route is enforced, the required transmission power might be different. The results show that, although there is a slight penalty regarding overall transmitted power, the increase is quite small, and it even becomes negligible for the low density scenarios. In fact, we can observe that the proposed solution somehow adapts to the network topology so that its behavior is closer to the minimum power one as the network density decreases, while it takes advantage of the additional routing opportunities when possible, to reduce the average accumulated exposure.

The algorithm has been afterwards analyzed in more realistic scenarios, where traffic sources dynamically change every 25 time slots (simulation experiment lasts 100 slots). In this case, one single density is analyzed, assuming a squared area of 140 m each side and 4 traffic sources.

Opposed to the previous case, the scenario has been studied for different configurations of the accumulated exposure cost, ω_{acc}. In particular, the algorithm behavior has been analyzed applying an *aging* function to this cost when the corresponding node does not belong to any route. This way, it is incremented an amount Δ_{acc} when the node takes part in a route, while it is reduced ∇_{acc} otherwise. Table 5.2 shows the particular set of values that have been used for both Δ_{acc} and ∇_{acc}.

Firstly, Figure 5.7 depicts the average accumulated exposure for each configuration
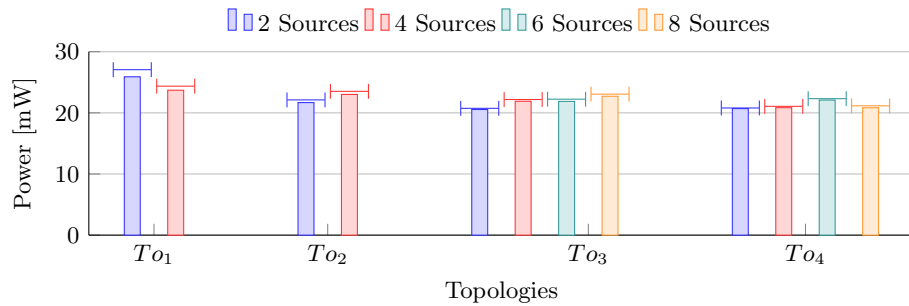
Figure 5.6: Average power per route. The markers correspond to the value obtained when using the EMF aware algorithm, while the columns represent the results of the minimum power alternative

Table 5.2: Configurations of the cost ω_{acc}. The table shows both the increment and decrement of the cost ω_{acc} for the different configurations

Cost parameter	Value set
Δ_{acc}	$\{3,6,9,12,15\}$
∇_{acc}	$\{0,...,\Delta_{acc}\}$

defined in Table 5.2. As before, this metric has been obtained for those nodes that have been selected to take part in, at least, one route. As can be observed, the gain in terms of accumulated exposure, when compared with the minimum power algorithm (dashed line), is notably lower than in the previous analysis, specially when the *aging* function is disabled (i.e. $\nabla_{acc} = 0$). The increment of the exposure is caused by the modification of the traffic sources during the experiment, since this would involve more nodes in the routes, thus incrementing the value of ω_{acc}. In turn, it would eventually force the algorithm to use alternative paths, leading to longer routes. In this sense, we can also see that the *aging* function has a positive effect, in particular when the cost increment and decrement happen at the same pace.

Furthermore, Figure 5.8 depicts the average power emitted by the nodes, per route and time unit, for the different configurations shown in Table 5.2, and the value obtained when using the minimum power algorithm. As can be observed, the variation of the traffic sources yields to a rather similar behavior of all configurations, which behave slightly worse than the minimum power solution in terms of transmitted power. Besides, it is worth noting that the use of the *aging* function does not have a relevant effect on this metric.

Altogether, the proposed solution extends minimum power algorithms by including the induced EMF exposure. In this sense, an appropriate calibration the EMF aware algorithm would allow us to prioritize either the accumulated exposure or the instantaneous transmission power, leading to the same behavior as the minimum power approach. Besides, it has been observed that the trade-off between costs and the penalty, in terms of total transmitted power, of a particular configuration of the algorithm strongly depends on the network topology. Finally, we have observed that the use of an *aging* function for the accumulated exposure has positive effects in more realistic scenarios.

5.1.2 Theoretical Exposure Model

In the previous section, the performance of the EMF aware algorithm has been analyzed by using a simple model of the EMF exposure. However, a more precise evaluation of a routing protocol based on the algorithm would require a more realistic metric of the exposure.

Similar to previous sections, we use conventional notation for the set of numbers where
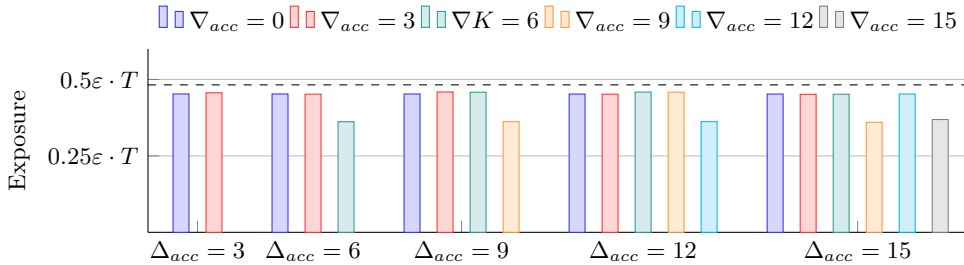
Figure 5.7: Average exposure at the end of each experiment. The value is obtained by averaging the accumulated exposure of the nodes that have taken part in a route. Dotted line indicates the value yielded by the minimum power algorithm
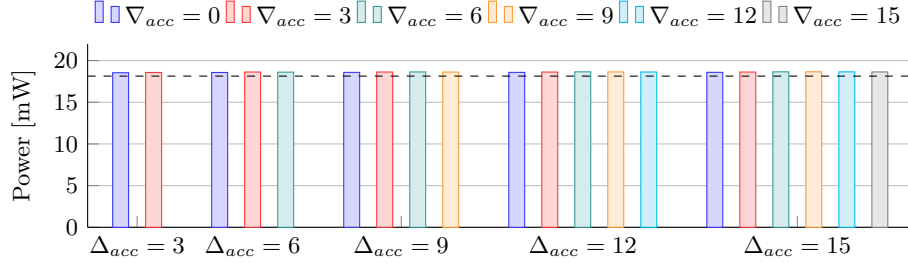
Figure 5.8: Average power at the end of each experiment. The value is obtained by averaging the power transmitted by nodes taking part in a route and normalized by route and time unit. Dotted line indicates the value yielded by the minimum power algorithm

$\mathbb{R}$ denote the set of real, and $\mathbb{R}_+$ and $\mathbb{R}^n$ represent the sets of non-negative real numbers, n-dimensional real vectors, respectively. Besides, the rest of sets are represented in calligraphy letters. Furthermore, vectors are written in bold font and sequences are represented between angle brackets $\langle \rangle$. Subscripts represent elements in a vector and superscripts elements in a sequence.

The exposure model defined here also assumes that the time is slotted, so that in a particular time slot each node can either transmit one packet or stay idle. We consider a scenario in which there is a set of nodes $\mathcal{N}$ and define $\langle \boldsymbol{\epsilon} \rangle$ as the sequence of exposure induced by all nodes over time. This way, $\boldsymbol{\epsilon}^t \in \mathbb{R}_+^{|\mathcal{N}|}$ is the vector of exposure at the time slot t, while ϵ_n^t is the exposure induced by a particular node $n \in \mathcal{N}$ at that slot. Besides, it is assumed that nodes are able to dynamically adapt the transmission power for each transmitted packet, being constant transmission power one particular case. In this sense, the exposure induced by a node grows every time it transmits a packet and remains constant otherwise. In general, it can be assumed that the growth rate (first derivative) of exposure of node, in a particular time instant can be expressed as a function $f()$ of its transmission power as follows:

$$\epsilon_n^t{}' = \begin{cases} f(P_{tx}) & \text{if node } n \text{ transmitted a packet in slot } t \\ 0 & \text{otherwise} \end{cases} \tag{5.6}$$

where t indicates the current time slot and P_{tx} holds for the chosen transmission power. For instance, $f()$ can be a linear function (as was considered in the previous section to evaluate the EMF aware algorithm), or more complex functions that consider other factors such as the antenna pattern.

Assuming that the first and second derivatives of the exposure function ($\epsilon_n^t{}', \epsilon_n^t{}''$, respectively) exist, we can establish the following properties:

1. Since the exposure is accumulative, $\epsilon_n^{t+\Delta} \geq \epsilon_n^t, \quad \Delta > 0$

2. Consequently, $\epsilon_n^t{}' \geq 0$, corresponding to this growing function over time.

3. Depending on the value of $\epsilon_n^t{}''$, we can establish the following possibilities:

- If $\epsilon_n^t{}'' > 0$, the rhythm of accumulating exposure on node n is increasing.
- If $\epsilon_n^t{}'' = 0$, the pace of accumulating exposure on node n is constant.
- If $\epsilon_n^t{}'' < 0$, the rhythm of accumulating exposure on node n is decreasing.

As can be observed, first derivative maps to transmission power, while the second one is related to the time a particular node is used in a route. Thus, based on these properties, appropriate exposure metrics can be defined to design EMF aware routing mechanisms. Besides, it would be also interesting to calculate the EMF exposure induced by a node from the routing metrics. In the following, it is demonstrated how this calculation can be performed.

We start by assuming continuous time, so that the exposure induced by a node $n \in \mathcal{N}$ at time t can be calculated as a function of its first and second derivative as defined in Theorem 1.

Theorem 1.

$$\epsilon_i(t) = \epsilon_n(t_0) + \epsilon_n'(t_0)(t - t_0) + \int_{t_0}^{t} \int_{t_0}^{x} \epsilon_n''(t)dtdx \tag{5.7}$$

where $\epsilon_n(t_0)$ and $\epsilon_n'(t_0)$ are known constants.

Proof. It is well-known that the area between 2 points, a and x, below the curve of any one variable function h that is derivable, can be expressed formally by means of the Riemann integral. As indicated in Eq. 5.8, if we consider a to be a fixed initial real value and x to be variable, we can define the value of the curve at each point as a function of the integral from the considered initial value.

$$\int_{a}^{x} h'(t)dt = h(x) - h(a) \iff h(x) = h(a) + \int_{a}^{x} h'(t)dt \tag{5.8}$$

Then, if the function h is two times derivable (as it was assumed for the exposure function), considering that we know the expression or shape of the function h'', we could similarly express h' as indicated in Eq. 5.9.

$$h'(x) = h'(a) + \int_{a}^{x} h''(t)dt \tag{5.9}$$

Finally, if we integrate this once again, between a and b we get:

$$\int_{a}^{b} h'(x)dx = \int_{a}^{b} h'(a)dx + \int_{a}^{b} \int_{a}^{x} h''(t)dtdx \iff \tag{5.10}$$

$$h(b) - h(a) = (b - a)h'(a) + \int_{a}^{b} \int_{a}^{x} h''(t)dtdx \iff \tag{5.11}$$

$$h(b) = h(a) + (b - a)h'(a) + \int_{a}^{b} \int_{a}^{x} h''(t)dtdx \tag{5.12}$$

$\square$

Since we have assumed that time is discretized, we can express the double integral in Theorem 1 by summing the values of ϵ_i'' in the various slots. Besides, if we assume that the the exposure is zero at slot 0, we can express the exposure induced by node $n \in \mathcal{N}$ at slot t as indicated in Theorem 2.

Theorem 2.

$$\epsilon_n^t = \sum_{i=0}^{t} {\epsilon_n'}^i = \sum_{i=0}^{t} (t - i + 1) \, {\epsilon_n''}^i \tag{5.13}$$

Proof. Let h be a staircase function, the value of h in a point n, h^n, can be expressed as the double summation of its second derivative as follows as expressed in Eq. 5.14.

$$h^n = \sum_{i=0}^{n} \sum_{j=0}^{i} h''^j \tag{5.14}$$

which, in turn, can be expressed as indicated in Eq. 5.15.

$$\sum_{j=0}^{n} \sum_{i=j}^{n} h''^j = \sum_{j=0}^{n} (n - j + 1) h''^j \tag{5.15}$$

$\square$

According to the previous definitions, the exposure induced by a node grows according to how it sends packets. It is worth recalling that the theoretical model is not constrained by the power used to send each packet, but it only considers the event of sending. Depending on the technology, as well as on the power used to reach the next hop within the route, the accumulated exposure will grow differently.

In order to illustrate the use of the theoretical framework to evaluate the behavior of a routing protocol, Figure 5.9 shows a temporal sequence in an arbitrary node $n \in \mathcal{N}$. As can be observed, the accumulated exposure $\boldsymbol{\epsilon_n}$ is an ever growing function over time, while the lower axis illustrates the pace of the exposure accumulation. From a protocol perspective, the goal is twofold: first that the accumulated exposure of each node is kept as low as possible; and, that the pace of exposure accumulation is compensated among all nodes within the network.

Bearing in mind that time is assumed to be discretized, and that each node knows the transmission power required to transmit a packet to any of its neighbors, we can calculate the second derivative as the variation of the power transmitted during two consecutive slots, ${\epsilon_n^t}'' = {\epsilon_n^t}' - {\epsilon_n^{t-1}}'$.

In order to consider the EMF exposure within a routing protocol, there are two key aspects to take into account. First, minimizing the total transmission power of a route, since it would reduce the overall exposure within the whole scenario. Additionally, by considering the exposure accumulation pace as a routing cost, the protocol should be able
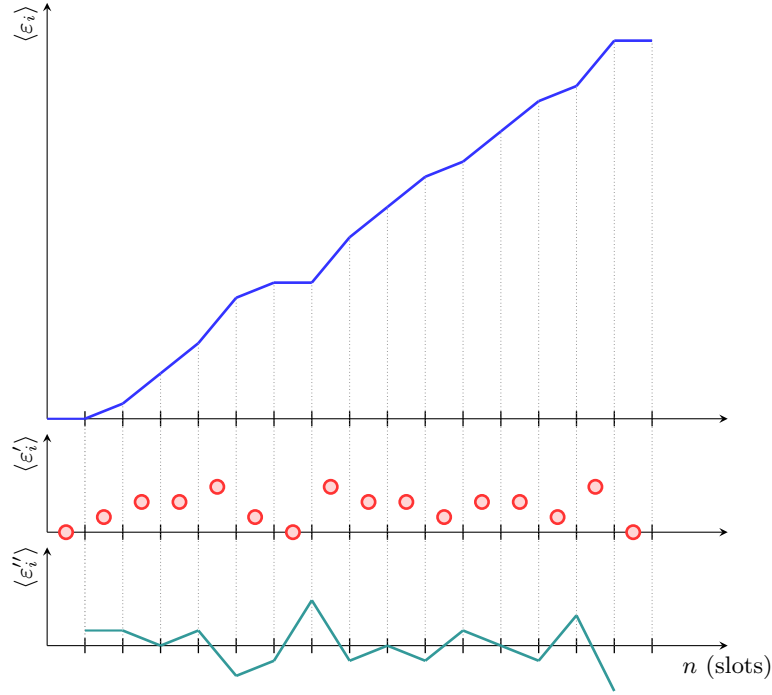
Figure 5.9: Evolution of the exposure and its second derivative upon transmission events

to balance the traffic between various routes, avoiding exposure peaks over a particular area (this would be the consequence of keeping the same route, as it would be the case of traditional routing mechanisms). This exposure model allows considering dynamic route changes, by means of taking into account the temporal evolution of the exposure accumulation pace.

5.1.3 EMF aware routing protocol

In this section we define and evaluate EMF Aware Routing Protocol (EARP), a practical implementation of the algorithm described in Section 5.1.1. In this sense, the protocol fosters the correlation between the population exposure and the route cost, so as to both reduce the overall exposure and avoid areas with higher accumulated exposure.

Legacy routing solutions can be roughly divided in two main groups: reactive protocols, in which nodes trigger route discovery procedures only when they need to; and proactive protocols, where nodes periodically send topology-related information, which is afterwards used to establish the corresponding routes. One of the main advantages of the reactive protocols is that, especially for moderate traffic loads, they exhibit a better energy consumption behavior. Hence, it is sensible to adopt the philosophy of reactive protocols for our solution. In particular, EARP has been designed adopting the basic functionality of one the most widespread protocols, Ad-hoc On-Demand Distance Vector (AODV) [Per03].

Protocol signaling and operation

Being based on a legacy protocol, the design of EARP borrows some of the features of AODV. In the following, we will present the most relevant aspects of the EARP design that make it differ from legacy solutions, leaving aside those that are shared with AODV. The information presented here is complemented with Appendix C, where further information is provided.

Firstly, it is assumed that nodes are able to dynamically adapt their transmission power, according to the link state with their neighbors, and that the protocol is itself aware of such power, leveraging some cross-layer solution. As for the protocol signaling, four messages have been defined:

- *HELLO*: this is a broadcast message with a twofold goal. It allows neighbor discovery and it is also used to notify about a change in the node cost; for instance, upon the event of accumulated exposure increase. Figure 5.10 depicts the fields of this message. As can be seen, it includes two flags to indicate the role of the message: M determines whether the packet is used for neighbor discovery purposes ($M = 0$) or for notifying a change in the corresponding cost ($M = 1$). In the latter case, the S flag indicates the sign of such modification (0 or 1 for positive or negative change, respectively). Finally, the *Cost* field carries the corresponding value.

- *ROUTE_DISC*: it is a broadcast message initiated by a source to find the route to a particular destination. It carries the accumulated cost, so as to disseminate it during the route discovery process. Each node adds a value that is proportional to the transmission power that was required by the previous node to reach it.

0 1 2 3 4 5 6 7	8 9 10 11 12 13 14 15	16 17 18 19 20 21	22	23	24 25 26 27 28 29 30 31
Version	Type	Reserved	M	S	Cost
SRC-ID					
SRC-SEQ-NUMB					

Figure 5.10: Structure of the *Hello* message

0 1 2 3 4 5 6 7	8 9 10 11 12 13 14 15	16 17 18 19 20 21 22 23	24 25 26 27 28 29 30 31
Version	Type	Reserved	Num. Hops
Cost K			
Cost P			
SRC-ID			
SRC-SEQ-NUMBER			
DST-ID			
DST-SEQ-NUMBER			

Figure 5.11: Message structure of *ROUTE_DISC* and *DISC_ACK*

- *DISC_ACK*: this message follows a *ROUTE_DISC*, and it is sent by the destination as a unicast transmission towards the corresponding source. Its format is the same as the one presented for the *ROUTE_DISC* message, the only difference being in the corresponding *Type* field. Since the nodes follow the same procedure as the one depicted for the management of the *ROUTE_DISC* message (i.e. they add the required transmitted power), the protocol is able to handle asymmetric links, in which the power required to send a packet depends on the particular direction.

- *REPORT*: it is used to disseminate a cost change (either an increase or a decrease of the accumulated exposure). Thus, it follows a *HELLO* with *M* = 1. This message is propagated to ensure that all affected nodes receive the information. Figure 5.12 illustrates the format of this message.

One of the main novel aspects of the proposed protocol is its ability to dynamically manage the variation of costs caused by the evolution of the accumulated exposure, mimicking the behavior of the algorithm described in Section 5.1.1. We note that the evolution of the cost with respect to the accumulated exposure does not follow any particular pattern but it is a consequence of the routing decision making, thus resulting in an asynchronous behavior. Furthermore, each node locally stores the corresponding route entries (identified by the *destination/next-hop* pair), which might have one of the following four states (see Appendix C for more information about route entries):

- *InUse*: the route is being used to send locally generated data. This means that the node is the actual source of such entry.

- *Active*: the route is being used to forward data. This is used when the node acts as a forwarding entity.

- *Active_InUse*: this state applies when the two previous conditions apply.

- *Valid*: the entry is not currently being used, but it has not yet expired.

Figure 5.13 shows the message exchange that would be triggered upon a change of a node cost. We assume that there already existed a route from *A* to *D*, *B* and *C* being intermediate nodes. At a certain time, *C* needs to inform about a modification of the

0 1 2 3 4 5 6 7	8 9 10 11 12 13 14 15	16 17 18 19 20 21	22 23	24 25 26 27 28 29 30 31
Version	Type	Type Report	R	S	Valor
SRC-ID				
SRC-SEQ-NUMBER				
Routes Num.				
ROUTE-DST-ID1				
. . .				

Figure 5.12: Structure of the *REPORT* message

accumulated exposure, for instance upon crossing a predefined threshold. This would eventually impact the cost stored at the routing tables of the remaining nodes within the route. Since A is the source, it internally has tagged the route to D as *InUse*, while the state of the corresponding entry in both B and C is *Active*.

First C sends a *HELLO* message with the flag $M = 1$, notifying about a cost change. Upon receiving it, B and D check whether they have any active route using C as the next hop. In this particular example, D does not have any affected entry, and the message is thus silently discarded. On the other hand, when B receives the *HELLO*, it checks that a route entry with C as the next hop has *Active* state, so that it propagates the update by means of a *REPORT* message. Since it is a broadcast message, the *REPORT* bears information about the source of the message in the field *SRC-ID* (i.e. node B) and the list of routes affected by the cost change. In particular, the field *RoutNum* carries the number of affected routes, and it is followed by the identifiers of all the corresponding destinations (i.e. *ROUTE-DST-ID1*). In this example, *RoutNum* equals 1, followed by the identifier of the corresponding destination, D. This way, those nodes that receive the *REPORT* get aware that routes having C as next hop, and whose destination is in the list reported by C (i.e. node D), are affected by the cost change. When B received the message, it updated the route entry towards D accordingly, and it continues propagating the cost change.

Finally, when A receives the *REPORT* message, it parses the *RoutNum* field, checking whether it has any entry towards one of the identifiers that are included within it, having B as the next hop. In Figure 5.13, there would be a match with an entry towards D, and its
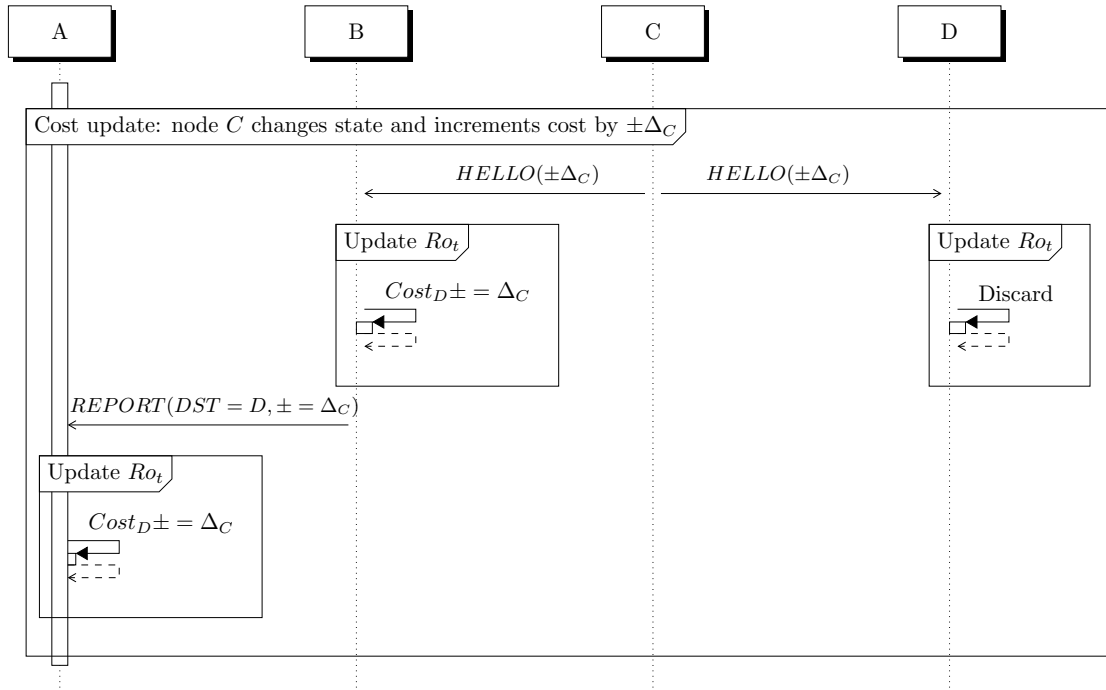
Figure 5.13: Temporal diagram of cost change

cost is therefore increased, according to the information carried in the *REPORT* message. In addition, *A* would also forward it, if the entry would be either *Active* or *Active_InUse*.

Protocol evaluation

First, we have assessed the protocol implementation by comparing it with a legacy solution. Then, its behavior has been further evaluated by using the theoretical model defined in Section 5.1.2. The protocol has been fully implemented in the network simulator *ns3* [Ns317] using *802.11* as communication technology with a maximum transmission power of 16 dBm. The particular transmission power required to reach the next hop has been obtained by an off-line training procedure described in Appendix C.1.

For the first analysis, EARP has been compared with legacy AODV protocol, which minimizes the length of the routes in terms of number of hops. The decision of using AODV as baseline was taken by its availability and simplicity.

We randomly deploy 50 nodes over a square scenario of 100 m each side, which start traffic flows with a fixed duration. During the deployment, it is enforced that the nodes form a connected graph, so that any node can reach each other. In this scenario, all the flows have the same destination, which is the closest route to the center. Table 5.3 summarizes the scenario setup.

In this first analysis, the cost report has been implemented following a simple state machine, where each sate has a different associated cost. Periodically, each node updates its state, so that the cost report is triggered upon state change. It is worth noting that only transitions between adjacent states are allowed. In particular, 4 states have been defined, $S_{\{0,1,2,3\}}$, so that moving from one state to an adjacent one happens when a threshold is crossed. In this sense, for each state S_i, two thresholds are defined, Th_i^{i-1} and Th_i^{i+1}, which indicate the criterion to move to S_{i-1} or S_{i+1} respectively. Besides, nodes increment their cost by a factor 1.6 when moving from state S_i to S_{i+1}. The particular values applied to the thresholds are provided in Table 5.3, where they have been defined in number of packets sent during 10 seconds.

Furthermore, in order to measure the exposure over the whole scenario, a grid of probes are deployed with a separation of 5 m. These have been implemented as passive WiFi nodes, able to measure the received power, as a consequence of overheard data packets.

In a first set of simulations we studied the accumulated power that was received by the grid of probes during a simulation to validate the implementation of EARP and its
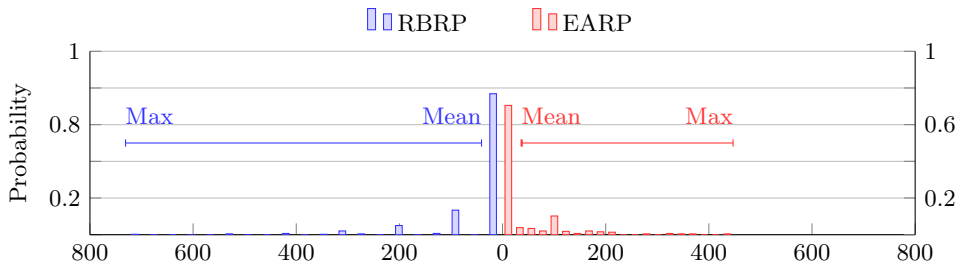
Figure 5.14: PDF of the power received by the grid of probes for both EARP and AODV

Table 5.3: Scenario setup to compare EARP and AODV

Topology	
Size(m^2)	100×100
Max. transmission power	16 dBm
# Routers	50
Traffic	
Protocol	UDP
Packet size (B)	1024
Generation rate (s)	0.15
Duration (s)	25

Cost Report Scheme (window 10 s)		
State	T_i^{i-1} (#packets)	T_i^{i+1}
S_0	0	50
S_1	25	120
S_2	100	190
S_3	150	$--$

configuration. We carried out 100 independent experiments with the two protocols, each of them with a duration of 100 seconds and 4 randomly generated traffic flows. Figure 5.14 depicts the histogram of accumulated power at the probes for each protocol, averaged over the 100 experiments. Furthermore, Table 5.4 shows the statistical values of the power at the probes in this scenario. In the light of these results, we can observe that EARP yields a better distribution of the received power, thus exposure peaks. In fact, as can be observed in Figure 5.14, EARP is able to reduce the difference between average and maximum power received by the probes, as a consequence of presenting more power levels received by the sensing grid.

Furthermore, from the statistics shown in Table 5.4, we can observe a remarkable reduction of the maximum achieved value, coupled with a lower standard deviation. The latter is due to the fact that AODV does not minimize the transmission power, but the number of hops.

After validating the implementation, we have extended the comparison between both

Table 5.4: Statistics of the power received by the grid of probes

Protocol	Max. (W)	Mean (W)	Std. Dev (W)
AODV	730.958	40.7738	91.9429
EARP	446.9462	37.0287	72.8755

protocols in a scenario with more traffic demand. For that, we have run 300 experiments, each of them 1200 seconds long, and having 48 traffic flows that start in each simulation. Again, all flows were sent to the same destination, which was the closest node to the middle point of the scenario.

Figure 5.15 depicts the value of the total power measured by the grid of probes for a single experiment. Aligned with the previous results, EARP is able to reduce the exposure levels obtained when compared with AODV, specially for the highest received power situations. Furthermore, Figure 5.16a represents the CDF of the sum of all the samples during a simulation. The results yield that EARP shows a better performance than AODV in most of the cases ($\approx 80\%$), while for the remaining percentage the difference is rather small.

Although the main goal of this study was to assess the potential exposure reduction brought about by EARP, we also wanted to study whether there is any trade-off that could jeopardize the perceived QoS. Since the flows that were used did not saturate the wireless channel, we analyzed the route lengths for each of the protocols. Figure 5.16b depicts the probability density functions of a such parameter for both AODV and EARP. According to the results, the difference of the routes length is almost negligible for this topology configuration. As expected, for dense scenarios, EARP is able to improve the behavior of legacy routing protocols in terms of exposure, without having a negative impact over other performance metrics.

Once the potential of the protocol has been assessed, a more rigorous assessment has been performed based on the algorithm and theoretical analysis of exposure metrics depicted in Sections 5.1.1 and 5.1.2 respectively. According to the protocol specification, EARP itself does not define the cost computation procedure, but it provides a framework to incorporate a dynamic cost adaptation scheme, including the mechanisms to appropriately propagate the corresponding changes.

In this second analysis, two different costs have been implemented within the protocol, mostly resembling the ones used to evaluate the algorithm. The first one, namely ω_P, corresponds to the transmission power that is required to reach the next hop. Additionally, ω_{acc} accounts for the accumulated exposure. During the simulations, ω_P takes the same value as the transmission power in mW, within the range $[1, 40]$. As for ω_{acc}, it is assessed every T seconds, by monitoring the total transmission power during the last slot. In order to study the impact of assigning a different weight to either of the two costs, ω_{acc} is scaled
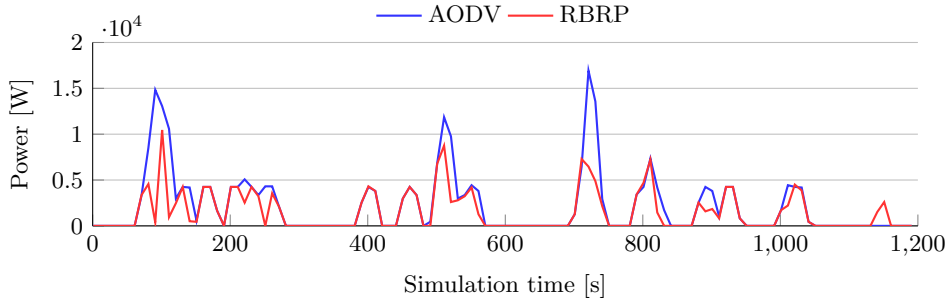
Figure 5.15: Temporal evolution of the total power measured by the grid
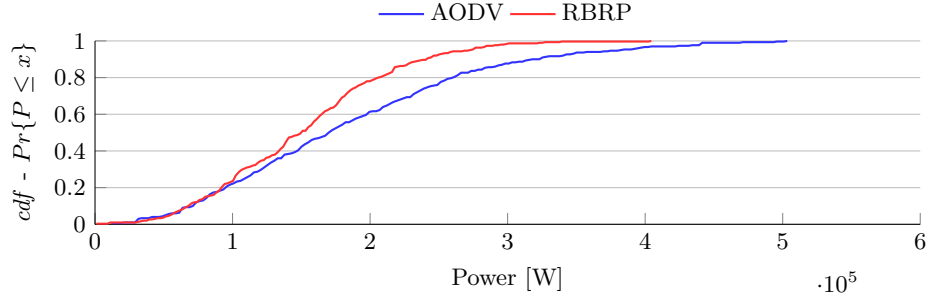
(a) CDF of the total power measured by the grid
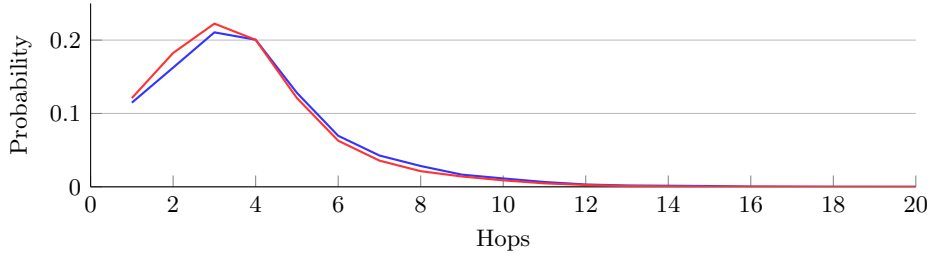
(b) Probability Density Function (PDF) of the number of hops to send a packet

Figure 5.16: Statistical performance comparison between AODV and EARP

by a factor of ρ. We will afterwards analyze how it impacts the protocol performance.

As can be observed, ω_P and ω_{acc} can be seen as a particular definition of the abstract ones that were used in Section 5.1.1. Furthermore, they both have a clear relationship with the theoretical framework presented in Section 5.1.2. Indeed, the exposure metric ϵ can be minimized by reducing the transmission power, i.e. with ω_P. On the other hand, ϵ'' (measured in mW/s), which modulates the pace at which the accumulated exposure varies, is considered by the latter cost, ω_{acc}. In particular, those nodes that have transmitted more power during the last slot (higher mW/s), are penalized in the route decisions process, and thus the pace at which exposure accumulation varies is also taken into account by the proposed protocol.

We have first studied the protocol behavior over a simple network, to analyze the impact that the ρ parameter may have. The corresponding network topology is shown in Figure 5.17. As can be seen, there are 3 traffic sources, which send traffic to 3 potential gateways, with several available routing alternatives. A simulation of 120 seconds has been performed, where sources send traffic at 120 Kbps to any gateway. The duration of the time slot T is set to 5 seconds to evaluate the accumulated exposure.

Figure 5.18 shows the rhythm at which the accumulated exposure grows for the three potential intermediate nodes (I_5, I_6 and I_7) and for different ρ values. The results show that for high ρ values ($\rho = 1$ or $\rho = 5$), EARP is actually able to modulate the exposure increasing pace. However, when the weight given to this cost decreases, the protocols favors the route with the lowest transmission power (minimum ω_P), which is maintained during all the simulation, and thus there is no change on the rhythm at which the accumulated
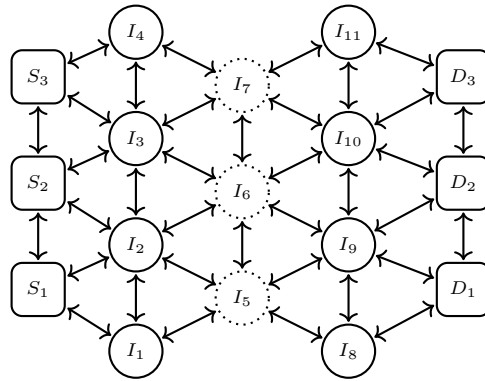
Figure 5.17: Simple network with 3 sources and 3 destinations

(a) Node I_5

(b) Node I_6

(c) Node I_7
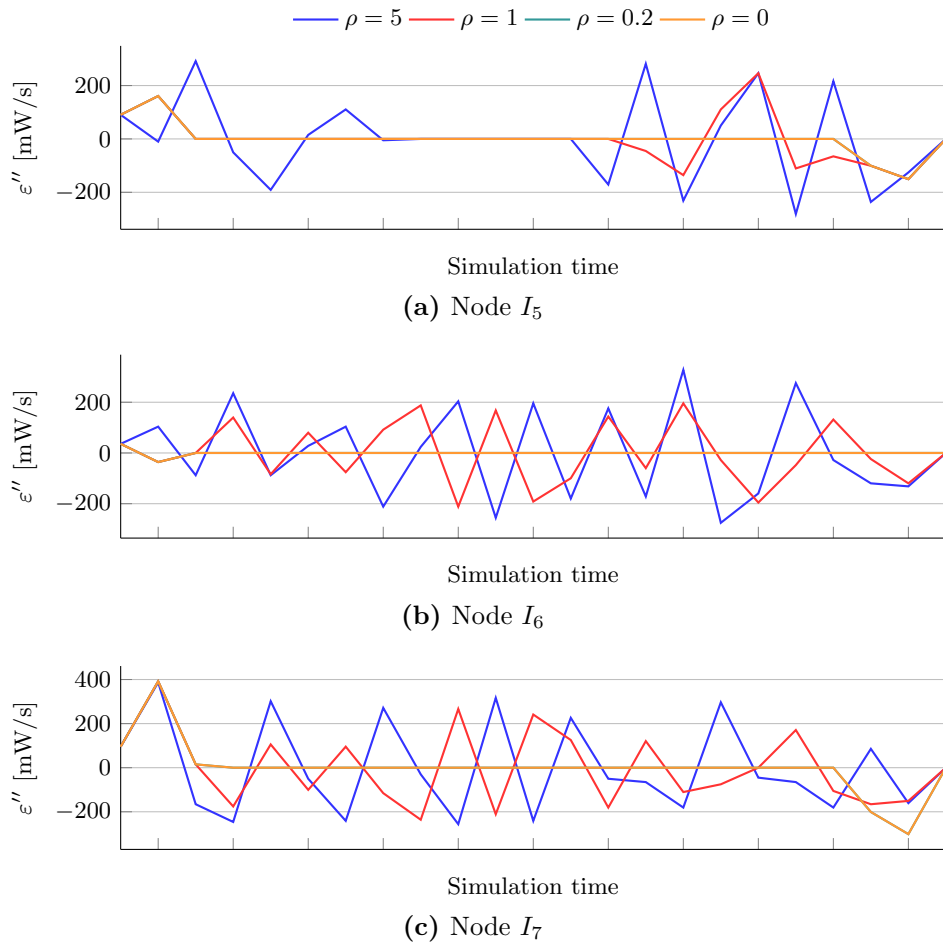
Figure 5.18: Evolution of the exposure accumulation pace ϵ''

exposure grows at the intermediate nodes.

Afterwards, we also evaluated the impact of the ρ parameter over the overall transmitted power. Figure 5.19a shows its standard deviation for different values of ρ. We can see a clear influence of the ω_{acc} cost, resulting in a far more homogeneous exposure when ρ is higher. Besides, Figure 5.19b illustrates the power that was transmitted by the entire network. In this setup, it can be also observed a slight improvement for the highest ρ value, which was not expected, since higher ρ usually involves using routes that are not favoring the minimum power criterion. However, the traffic balancing between the various routes leveraged by this configuration may lead to fewer collision events and retransmissions.

After assessing the correct behavior of the routing protocol, we have analyzed its performance over more complex scenarios, considering random topologies. We assume a square scenario of 500×500 m^2, where 4 arbitrary sources send traffic to any of 4 gateways, which are placed such as to maximize their coverage, similarly to the analysis of the algorithm in Section 5.1.1. We increase the number of intermediate nodes, and thus the network density, from 50 to 200, and the value of ρ is modified taking values $1, 3, 5$. As was done when studying the algorithm performance, we discard those network deployments where a source cannot reach, at least, one of the available gateways. Traffic characteristics and window time (T) are the same as before and 100 independent simulations for each configuration (i.e. density and ρ value) are performed.

First, Figure 5.20 shows the CCDF of the power transmitted by each node during the whole experiment. As can be observed, there is a slight improvement when moderated weight (1 and 3) is given to the metric that considers the accumulated exposure, ω_{acc}. This enhancement is slightly higher for more dense networks, as can be observed on Figure 5.20b. In any case, the difference with the configuration where the protocol just seeks the minimization of the transmitted power ($\rho = 0$) is not very relevant.

Afterwards, we assessed the influence of the ω_{acc} metric over the average and overall transmission power. Figure 5.21a shows the average transmitted power per node. We can observe a monotonic decrease of the average power due to the fact that the denser the network, the more alternatives appear. This effect, in turn, implies that it is more likely to find routes requiring less power. The results show that the use of higher ρ values leads to a clear benefit in terms of the transmitted power. Furthermore, the same effect can be

(a) Standard deviation

(b) Total power transmitted
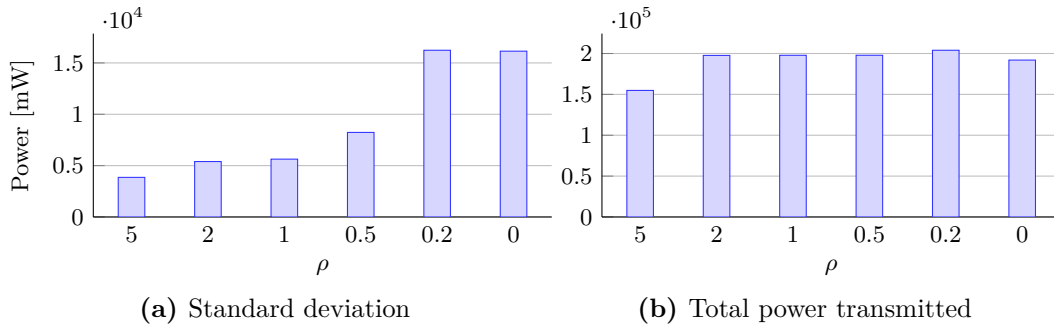
Figure 5.19: Simple topology: power transmission statistics

(a) 50 intermediate nodes **(b)** 200 intermediate nodes
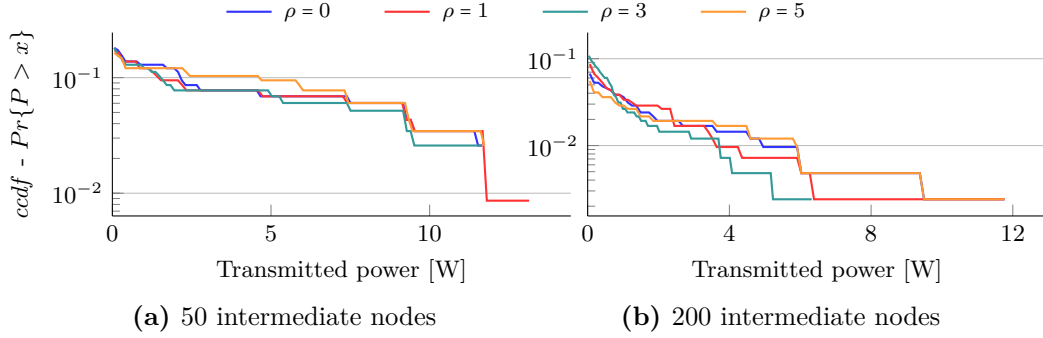
Figure 5.20: CCDF of the power transmitted by each node

observed for the overall transmitted power. Figure 5.21b exhibits a similar behavior than the one seen for the standard deviation of the transmit power.

We can finally conclude that, for the studied topologies, the use of the dynamic metric that accounts for the accumulated exposure yields a reduction of the total power transmitted by the complete network. This reduction is achieved together with a fairer distribution of the exposure induced by the different nodes that, in addition, yields a certain traffic balance between the available paths.

(a) Average transmitted power **(b)** Total transmitted power
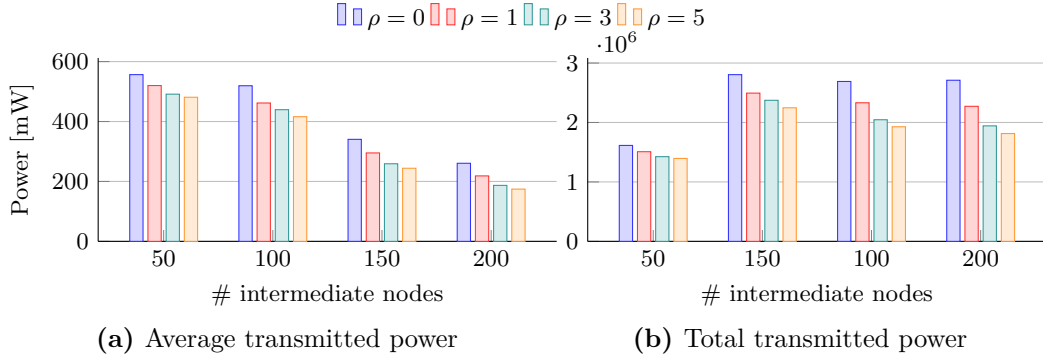
Figure 5.21: Power transmission statistics obtained from the random topology

5.2 EMF aware video content delivery

In this section we propose modifications to traditional video delivery services to potentially reduce EMF exposure. We have established a reference generic scenario, depicted in Figure 5.22, over which the different proposed mechanisms are evaluated.

As can be observed, the setup comprises LTE access network, and a video service provider, that might not necessarily belong to the same operator. On the server side, we assume a widely adopted setup. The video stream is encoded using $H.264/AVC$, the *codec* recommended by the 3GPP, and DASH [Sod11] is used as video streaming protocol.

When using DASH, a number of different representations of the same content, each of them having different coding parameters, are generated and stored in one or more Hypertext Transfer Protocol (HTTP) servers. Each representation is characterized by a
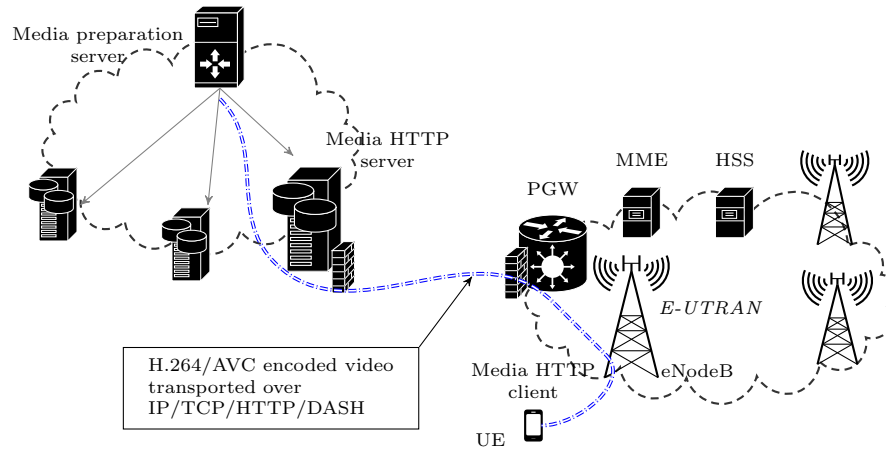
Figure 5.22: Architecture of the reference scenario. It comprises media content delivery network and LTE operator network

different quality, and the data rate required to send them is thus different. In the data plane, a HTTP client at the UE retrieves a Media Presentation Description (MPD) file, which provides the location of the different representations. Afterwards, and by means of HTTP *verbs*, the client is able to obtain the most appropriate video segment, according to the perceived quality. As can be inferred from the previous description, Transmission Control Protocol (TCP) is used for video streaming, since it is the protocol HTTP relies on.

The corresponding video slices are transported over TCP/IP and afterwards forwarded to the LTE protocols, as illustrated in Figure 5.23[1]. As can be seen, Internet Protocol (IP) datagrams are sent to Packet Data Convergence Protocol (PDCP) layer, which performs header compression and integrity checking, amongst other functionalities. Then the PDU
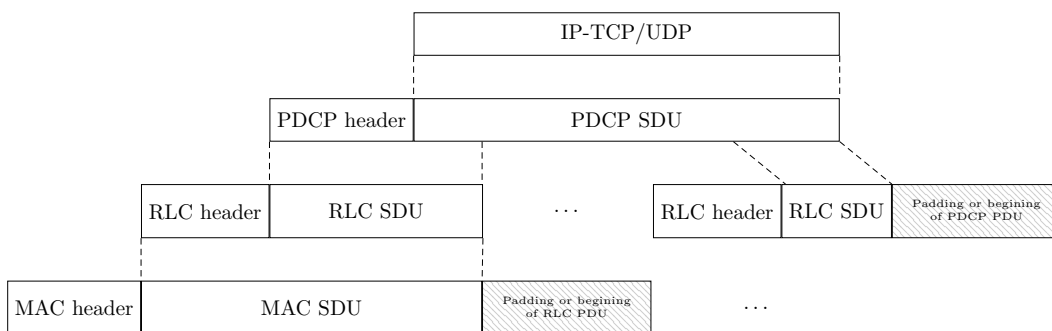
Figure 5.23: Simplified encapsulation diagram of IP and LTE protocols

---

1   Figure 5.23 uses standard Open Systems Interconnection (OSI) model nomenclature. Service Data Unit (SDU) refers to the data unit sent down in the OSI protocol stack from a layer to the one immediately below. Protocol Data Unit (PDU) indicates the data specified by a protocol in a layer.

of the PDCP layer is sent to the Radio Link Control (RLC) protocol, which performs a first fragmentation of the packets. Depending on the configuration, this protocol may implement concatenation, segmentation and reassembly of PDCP PDUs, as well as error protection, using Automatic Repeat Request (ARQ) mechanisms. In particular, RLC presents 3 operation modes [TS25322]: Transparent Mode (TM), in which content goes through this layer without any modification (i.e. only buffering is performed); Unacknowledged Mode (UM), where no error protection is accomplished; and Acknowledge Mode (AM), in which ARQ is implemented. The first mode is used by control channels, while either UM or AM can be used for user data, depending on the traffic type.

Additionally, from an EMF exposure point of view, the target is the reduction of resource utilization without jeopardizing the QoS. To this end, two different and complementary modifications of the reference scenario have been studied, which foster adaptive configurations of both transport and RLC protocols.

As for video streaming, there exist different strategies to adapt the video content to the constraints imposed by networks and terminals. For instance, in [Sim13] the authors consider the QoE perceived by users while the energy consumed by networks and terminals is prioritized in [Kha14]. The strategies can be divided into two types of algorithms. In the first type, the adaptation is focused at the terminal [Mil12]. For example, a HTTP client downloads an audio-visual content stored at a server with two different representations: one at 1000 Kbps and the other one at 500 Kbps. Having obtained and read the MPD file, the client requests a segment of the 1000 Kbps representation. Then, according to a given strategy, it estimates the available preferred data rate, which equals 500 Kbps, and it therefore selects the following segment from such representation. Hence, by knowing the available network capacity, the client can select the most appropriate representation for each segment, while the content is being downloaded. In the second approach, this adaptation is made by the operator managing the radio access networks [Ess13].

Although TCP is used to guarantee the reliability of DASH transmissions, from an exposure angle, there are a number of reasons to favor the use of User Datagram Protocol (UDP), such as its header size (8 bytes), which is smaller than the TCP one (20 bytes), or the fact that a UDP client, unlike a TCP one, does not transmit acknowledgments.

Hence, we propose a dual (TCP/UDP)-based DASH protocol. The initial assumption is that, if the radio access technology itself already provides a certain level of reliability, the use of a reliable transport protocol, like TCP, might be unnecessary, provided that the flow control functionalities can be implemented otherwise. For instance, in a scenario where

Table 5.5: Advantages and drawbacks of using TCP or UDP as transport protocol to media content delivery

PROS	CONS
Uplink resource utilization reduction	Firewall reconfiguration for UDP
EMF exposure reduction	Server and client modifications
Radio power consumption reduction	Radio access equipment modification

the content provider is within the operator system. In this case, a non-reliable transport protocol, such as UDP, may indeed bring about some benefits, as detailed in Table 5.5.

In our scenario, we focus on the resource utilization of uplink communications, with the consequent reduction in battery usage and EMF exposure. Furthermore, the simplification of the interactions between different layers, in particular between the TCP flow control and the RLC entities, brings additional benefits. On the other hand, there are as well some drawbacks that arise when using UDP, being the most relevant one the fact that firewalls usually block UDP connections.

Additionally, if the RLC instance has AM configuration, when one RLC PDU belonging to a particular UDP/IP datagram gets lost, the sender entity retransmits the corresponding PDU, regardless of the type of information carried by the current frame. Otherwise, if RLC uses UM, any lost PDU would not be recovered. As can be seen, the design of RLC responds to an all-or-nothing approach, so that all the data is processed alike.

Nevertheless, in the case of video services, it is well known that the lost of some pieces of information may be affordable, since the video *codec* would eventually be able to interpolate the missing slices without notably impacting the QoS. On the other hand, if the system accepts some losses, continuous transmission would lead to a higher number of packets received by the decoder, yielding an application throughput increase, and so a positive impact on the perceived QoS. An adaptive RLC mode, adaptive RLC-AM (aRLC-AM), has been developed and evaluated. It adapts the error protection of the particular content carried by the RLC. The proposed scheme is able to protect those video frames that carry critical information, such as data needed to perform interpolation. This functionality requires the RLC protocol to be aware of the type of data that each of the RLC SDUs bears, thus leading to a cross-layer approach. The description on how RLC gets aware of the type of information (i.e. critical or non-critical) carried in each packet can be found in [Die15a].

In the following sections, the adaptive solutions for both transport and RLC protocols are presented. First, the dynamic transport protocol selection is described and evaluated in Section 5.2.1. Later on, Section 5.2.2 will describe and study the performance of the proposed RLC modification.

As was done for routing mechanisms, we need to define appropriate metrics that consider the exposure in an indirect way, since parameters such as LEXNET EI can hardly be considered at service level. In this sense, the proposed solutions seek a more efficient usage of the transmitted data at different layers, which would yield a reduction of transmissions without impacting the QoS. In particular, we will assess the exposure reduction by studying the performance of RLC communications, with a direct impact on the physical transmission, leading to a lower average EMF exposure. Considering the definition of the EI metric (Eq. 2.1), the modifications in this section would imply lower average SAR values for video traffic.

5.2.1 Dynamic transport protocol selection

As has been mentioned before, in order to ensure a reliable delivery of audio-visual content, DASH segments are traditionally transported over TCP connections. Although the use of UDP would lead to unreliable communication, we propose the usage of a dual transport

protocol, in which either UDP or TCP can be chosen, depending of the estimated bandwidth in the radio network and considering the fact that RLC, with AM configuration, may ensure service reliability. It is worth noting that we take the hypothesis that the capacity of the wired network is unlimited.

DASH algorithms select the best representation based on the information gathered from the MPD file (bit rate of each representation together with other video parameters [Hor10]) and bandwidth estimation between server and client [Sim13]. Unlike legacy alternatives, the solution proposed here exploits the information that can be obtained from the link layer to adapt the video streaming, thus promoting a cross layer approach.

In this sense, the eNodeB could be able to estimate the available capacity, by monitoring the corresponding RLC acknowledgments. Whenever an acknowledgment is received at the eNodeB RLC, this becomes aware of the amount of bits that have been successfully received by the mobile terminal. If the acknowledgment corresponding to packet k is received by the RLC entity at the eNodeB at time t_k, it implies that the UE has correctly received the corresponding amount of data d_k. Hence, the sample of the instantaneous bandwidth used by the connection can be estimated as:

$$b_k = \frac{d_k}{t_k - t_{k-1}} \tag{5.16}$$

where t_{k-1} corresponds to the time when the previous acknowledgment was received. In order to avoid congesting the link, we need to ensure that the input traffic (our estimate) does not exceed the link capacity. To this end, we have followed the approach presented in [Mas01], which was used to estimate the traffic rate for TCP. We employ a discrete time low-pass filter to average the instantaneous bandwidth and to get the low-frequency components of it as follows:

$$\hat{b}_k = \frac{\frac{2 \cdot \tau}{t_k - t_{k-1}} - 1}{\frac{2 \cdot \tau}{t_k - t_{k-1}} + 1} \cdot \hat{b}_{k-1} + \frac{b_k + b_{k-1}}{\frac{2 \cdot \tau}{t_k - t_{k-1}} + 1} \tag{5.17}$$

where b_k and $\hat{b}_k$ are the sample and filtered values of the bandwidth at t_k, and $\frac{1}{\tau}$ is the cut-off frequency of the filter. For instance, if we consider a constant inter-arrival time $t_k - t_{k-1} = \frac{\tau}{10}$, the filter in Eq. 5.17 has constant coefficients as follows [Mas01]:

$$\hat{b}_k = \rho \cdot \hat{b}_{k-1} + \frac{(1 - \rho)}{2} \left[b_k - b_{k-1} \right] \tag{5.18}$$

where $\rho = 0.9$. This way, Equation 5.18 shows that the new estimate $\hat{b}_k$ is calculated as a weighted sum, where 90% corresponds to the previous value $\hat{b}_{k-1}$ and 10% is due to the arithmetic average of the last two samples, b_k and b_{k-1}.

In the following we will analyze the potential performance of the aforementioned approach with a twofold objective. First, we will study the service QoS when using both TCP and

UDP as transport protocol to assess under which circumstances UDP may be preferable. Then, the performance of RLC will be also studied with different configurations, to evaluate the reduction of RLC transmissions, what in turn is used to indirectly measure the EMF exposure reduction.

Experimental evaluation

The evaluation has been performed through an extensive simulation campaign performed over the *ns3* simulator [Ns317]. As already mentioned, it is focused on the LTE technology, which is integrated within the simulator by the LENA module[1]. The main simulation parameters are summarized in Table 5.6. Furthermore, we use the *hybrid-buildings* propagation loss model, which provides accurate results in terms of coverage and network performance.

The scenario emulates the one illustrated in Figure 5.22. The access segment consists of one eNodeB and one UE, which runs a video service provided by a server that is in fact connected to the Evolved Packet Core (EPC) of the eNodeB. Since the simulator does not provide appropriate video service implementations, the corresponding traffic has been modeled as a Constant Bit Rate (CBR) traffic, whose rate is selected from a set of typical values[2], according to the estimated bandwidth. These values range from 46.980 Kbps to 4.726 Mbps (corresponding to resolutions of 320×240 and 1920×1080 respectively), while the service duration has been fixed to 15 seconds and each of the video chunks are 1 second long, so that the data rate might be updated every second. We gradually increase the distance between UE and the eNodeB, and, at each point, we retrieve the video from the server. In addition, 100 independent simulations are executed per point, so as to ensure the statistical validity of the results.

In order to enhance the filtering procedure, we estimate a new sample of $\hat{b}_k$ every 200 ms, by considering the amount of acknowledged data during such interval at the eNodeB RLC entity. Since we are assuming a fixed *ACK* inter-arrival time, we obtain a filter with constant coefficients as in Eq. 5.18. Hence, by selecting the cut-off frequency value, we can adjust the relative weights of both the last estimated value, $\hat{b}_{k-1}$, and the last measured b_k.

Table 5.6: Scenario setup for dynamic transport protocol selection

Parameter	Value
UE Transmission power	23 dBm
eNodeB Transmission power	43 dBm
UE height	1.5 m
eNodeB height	30 m
RLC mode	AM

---

1 `http://networks.cttc.es/mobile-networks/software-tools/lena/`
2 `www-itec.uni-klu.ac.at/`

As mentioned earlier, we have carried out different simulations, moving the UE away from the eNodeB and studying a number of figures of merit. The distances vary within the interval $[150, 810]$, with a step of 20 m. For shorter distances, the quality of the link leads to almost ideal conditions, while for distances larger than 810 m the communication suffers from very hostile conditions and the simulation can hardly finish. Furthermore, we have used a τ value that is 10 times greater than the inter-arrival time, $\frac{\tau}{10} = t_k - t_{k-1}$, thus leading to a weight of $\rho = 0.9$ for the last estimated value $\hat{b}_{k-1}$.

We analyze the throughput that is observed at the application, see Figures 5.24 and 5.25, for different link conditions, to study the impact of the data rate adaptation and its interaction with the transport protocols. It is measured by dividing the amount of received data by the time between the first and last received packets. Hence, under certain conditions (for instance if the application receives just one burst of packets), the results may be misleading. In order to clarify this aspect, additional information, regarding the actual amount of information exchanged between the user terminal and the network, will be considered. Furthermore, when the rate adaptation is not activated, we assume a constant bit rate of 1663 Kbps, which is the mean value of the available rates.

First, Figure 5.24 shows the throughput of the application as a function of the distance. According to the results, the bit rate adaptation, as it is configured, has a negative impact over the UDP flow when the channel conditions are good (distances lower than 210 m). On the other hand, under more hostile conditions (distances above 400 m) the adaptive technique, applied to UDP traffic, outperforms the other configurations. Hence, for the scenario under study, the rate adaptation mechanisms can favor either TCP or UDP, depending on the particular conditions of the wireless channel. Furthermore, Figure 5.25 depicts the CDF of the throughput after carrying out 1000 independent simulations over a distance of 300 m. We can see that the results leveraging the proposed solution show a notable variability, which might have been expected, since it is a direct consequence of the rate adaptation.

In addition, Figure 5.26 shows the delivery ratio, defined as the percentage of data sent by the server and correctly received at the UE client. First, as depicted in Figure 5.26a, all configurations exhibit a similar performance when the distance between the eNodeB and the UE is below 400 m. For longer distances, the adaptive technique, applied to UDP traffic, outperforms the rest of configurations, as was also seen for the throughput. Figure 5.26b
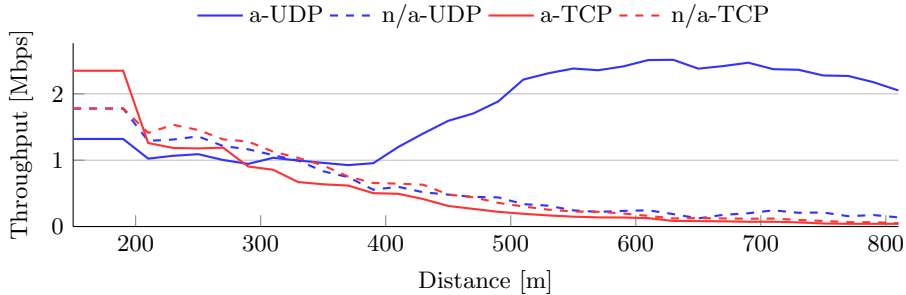
Figure 5.24: Throughput of the video application with and without adaptive data rate
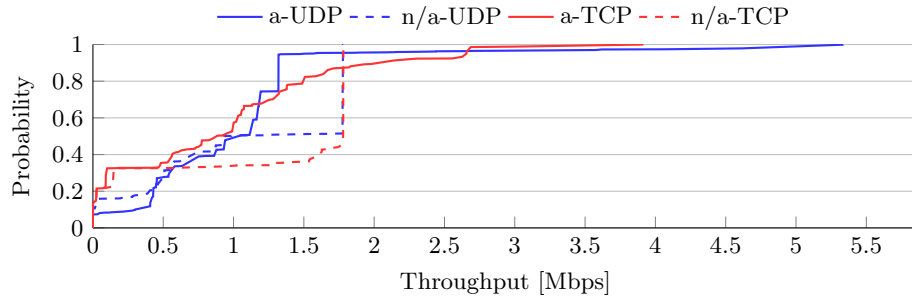
Figure 5.25: CDF of the throughput measurements. These values are obtained from 1000 simulations with a distance of 300 m
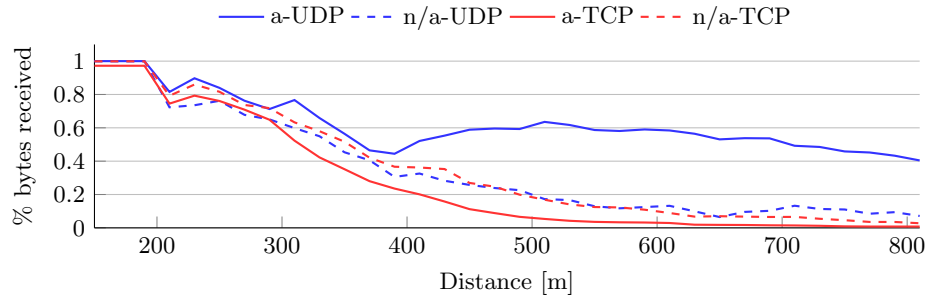
(a) Average values vs. distance
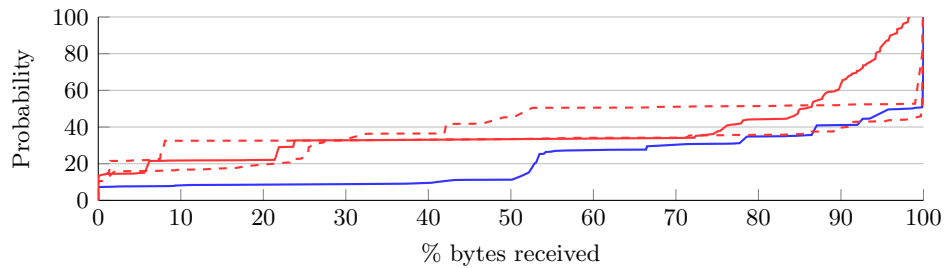
(b) CDF at 300 m

Figure 5.26: Average values of data successfully delivered by the application

shows the CDF of this metric, after carrying out 1000 independent simulations also at a distance distance of 300 m. We can see that all configurations show a great variability.

Finally, Figure 5.27 allows us to better understand the performance of the adaptation scheme. The figure shows the amount of bytes that were sent and received by the server and the client application, respectively. Regarding the packets received at the client, Figure 5.27b yields a behavior similar to the one seen for the other metrics. On the other hand, Figure 5.27a illustrates the behavior of the corresponding filter. The increase of the bytes received with UDP protocol is due to the bit rate increase during those periods in which the channel conditions improve. On the other hand, TCP does not show this

(a) Average transmitted bytes

(b) Average received bytes
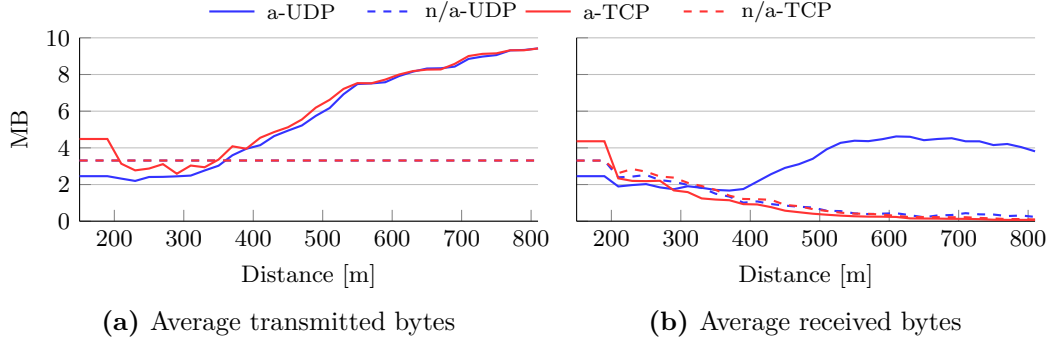
Figure 5.27: Application statistics

behavior due to its flow control (transmission window reduction). Besides, we can observe that for low distances, below 400 m, the filter configuration leads to underestimating the capacity, while for larger distances, the estimation works rather well.

After analyzing the behavior of the different transport protocols under bit rate adaptation, we will analyze its impact over the EMF exposure. In particular, we seek a exposure reduction by improving the efficiency of RLC transmissions, which are a good indicator of transmission at the physical layer. In this sense, Figure 5.28 shows the *RLC-normalized* throughput, a measure of the RLC efficiency, defined as the ratio between the observed throughput (in Kbps) and the number of RLC transmissions that were performed during a session.

This metric is therefore an indicator of the trade-off between QoS and the EMF exposure. Figure 5.28 shows that, when the distance to the base station is short, TCP exhibits a worse performance than UDP. On the other hand, when the distance increases (up to 400 m), the adaptive version of TCP yields better results. Finally, as it happened with the results discussed in Figure 5.24, for even larger distances (> 400 m), the performance of the adaptive version of UDP outperforms the other configurations.

In order to assess the statistical validity of the previous results, Figure 5.29 shows the corresponding CDF, obtained after carrying out 1000 independent simulations for one distance. We again observe the notable variability that characterize this metric, specially
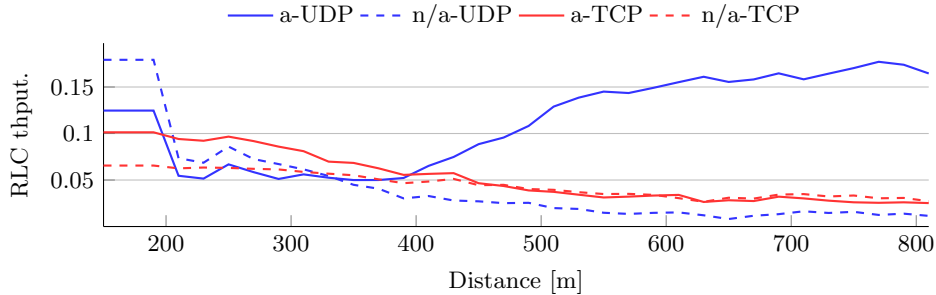
Figure 5.28: *RLC-normalized* throughput as a function of the distance
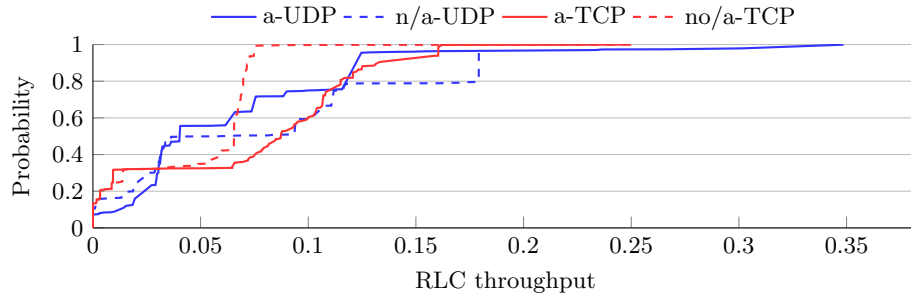
Figure 5.29: CDF of the *RLC-normalized* throughput for a distance of 300 m

in the case of UDP with adaptive traffic.

Finally, we have also studied the impact of the rate adaptation over the RLC protocol. We analyze the evolution of the transmission buffer, i.e. the amount of data that is waiting to be sent. Figure 5.30 represents the CDF of the overall number of bytes at the transmission buffer, for the two transport protocols (with and without the rate adaptation scheme) and at different distances. As can be observed, the adaptive configuration of the traffic flow reduces the buffer length, both for UDP and TCP. The difference between the two transport protocols is a consequence of the congestion control algorithms used by TCP.

We see that for TCP it is more likely to have an empty buffer when adapting the data rate. This indicates that the cut-off frequency that was selected for the experiments leads to an under-utilization of the available bandwidth. On the other hand, the rate adaptation applied to UDP heavily reduces the buffer length. Taking into account that the throughput for the adaptive scheme was higher at some distances, we can conclude that the proposed filtering allows taking advantage of those time periods in which the link quality temporarily improves.

(a) UDP **(b)** TCP
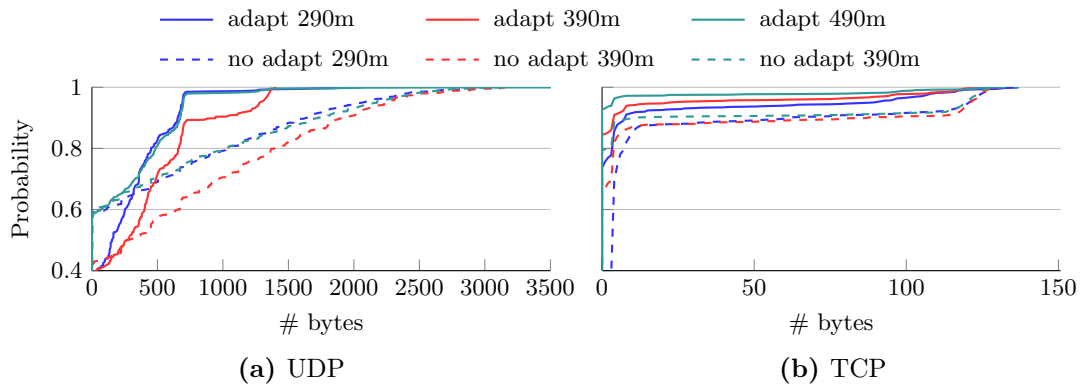
Figure 5.30: State of the RLC transmission buffer

5.2.2 Adaptive radio link content protection

This section describes the principles behind the aRLC-AM procedure, a modification of the legacy acknowledge mode of the RLC protocol [TS36322], that aims to reduce the EMF exposure without having a negative impact in the QoS. In brief, aRLC-AM is based on a dynamic ARQ scheme so that only critical data is protected against errors. This requires RLC protocol to be aware of the type of data each RLC PDU bears, thus leading to a cross-layer approach.

The implementation of the cross-layer solution to enable the aRLC-AM operation is based on the work presented in [Vie09]. In particular, a modified UDP header has been used to carry information regarding the importance of the data (i.e. critical or non-critical). Nevertheless, the main contribution of this work focuses on the concept of an adaptive RLC version and the evaluation of its performance. In this sense, the modified version of RLC acknowledged mode (RLC-AM) has been implemented within the *ns3* framework [Ns317]. In addition, we have also integrated a module so as to enable the cross-layer operation that would entail the RLC layer with the knowledge about the type of frame it is currently handling.

When configured as AM, legacy RLC establishes a maximum number of retransmissions ($MaxRetxThreshold$) for all PDCP PDUs. Differently, aRLC-AM adapts this value on a per packet basis (i.e. individually for each PDU). In this way, only those PDUs carrying critical information are protected, whilst the others are transmitted only once (i.e. they are never retransmitted).

The main idea behind aRLC-AM is to only protect critical information, thus assuming potential losses for those frames that do not bear such content. In this way, depending on the information type, the proposed scheme shows a behavior that resembles either AM or UM configurations. In our use case we consider a video stream comprising three different types of frames (intra coded frames-**I**, predicted frames-**P** and bi-directional predicted frames-**B**). While P-frames and B-frames improve the QoS, they require other frames that have been previously decoded. On the other hand, the first group (I-frames) are independently decoded and are in fact required to be able to decode other frames. In this sense, we configure the aRLC-AM procedure so as to protect I-frames, thus ensuring an acceptable video quality while losses of other groups are accepted.

In order to appropriately estimate the perceived QoS, we will use Eq. 5.19. This is a modification of the one presented in [Kha10], by expressing the BLER in terms of the Loss Rate (LR). In this sense, the probability of a PDU to be lost is calculated as the probability that it does not arrive at the destination after a maximum number of transmissions: $LR = BLER^{MaxData}$. The overall number of transmissions considers both the retransmissions and the initial transmission, and it can thus be expressed as $MaxData = MaxRetxThreshold + 1$, where $MaxRetxThreshold$ is an RLC protocol configuration parameter; note that, with RLC unacknowledged mode (RLC-UM), it equals 0.

$$MOS = a + \frac{b \cdot e^{FR} + c \cdot \log SBR + CT \cdot (d + e \cdot \log(SBR))}{1 + (f \cdot LR^{\frac{1}{MaxData}} + g \cdot LR^{\frac{2}{MaxData}}) \cdot h \cdot MBL} \tag{5.19}$$

Equation 5.19 allows us to estimate the Mean Opinion Score (MOS) as a function of the LR and the following traffic parameters: Frame Rate (FR), Source Bit Rate (SBR), Mean Burst Length (MBR) and Content Type (CT). In particular, we will focus on the impact of different $MaxRetxThreshold$ and the LR that is obtained with different aRLC-AM configurations, while the rest of parameters are considered alike in all cases.

Besides the impact over the QoS, the protection of less data would lead to lowering the number retransmissions, and this would be particularly relevant when link conditions worsen, as it would happen when the user is close to the cell edge. From the exposure point of view, this reduction would have a positive impact on the EMF exposure.

Experimental evaluation and validation

As mentioned before, aRLC-AM has been implemented in *ns3* and compared with legacy solutions under different configurations (i.e. different values of $MaxRetxThreshold$). Although the current implementation of *ns3* provides LTE access stratum protocols, including RLC in its different modes, it does not include an appropriate video service model, as was mentioned before. In this case, we have modeled the video application based on available statistical values [Dog08]. Hence, we distinguish three different types of frames within the same video flow, whose sizes follow the gamma distribution with the parameters that are shown in Table 5.7. Furthermore, we have defined a frame rate of 30 fps and the following Group of Pictures (GOP) pattern has been utilized: **I-P-B-P-B-P-B-P-B-P-B-P-B**.

The scenario has been configured as in Section 5.2.1, and main parameters detailed in Table 5.6 but for the RLC mode. It consists of one eNodeB and one UE, which is located at different distances from the eNodeB, in order to vary the channel conditions. In this sense, and based on the Adaptive Modulation and Coding (AMC) operation [Pir11], the amount of data (i.e. transmission opportunity) the UE can send decreases as it is moved away from the base station. As mentioned earlier, we use the RLC retransmissions as a proxy for the EMF exposure, taking as the baseline the RLC-UM, which uses no retransmissions

The validation procedure is carried out by establishing video sessions from the UE towards a remote device (i.e. assuming uplink communication). At each point (distance), we have carried out 20 independent simulations. The UE moves from a distance with good channel conditions to a point in which the sessions would be eventually interrupted due to link failures.

First, Figure 5.31 shows the throughput that was measured for the different RLC modes.

Table 5.7: Configuration of video service indicating the parameters of the gamma distribution, where the mean value $\mu = k_\gamma \theta_\gamma$

Resolution	SBR (Kbps)	Frame Type	μ (10^3)	θ_γ	k_γ (10^3)
	896	I	110.3	9.013	12.24
720×576		B	16.08	1.665	9.660
		P	33.40	2.812	11.88

For the proposed scheme, we can see that the higher the $MaxRetxThreshold$, the closer aRLC-AM performance is to that shown by legacy RLC-AM. On the other hand, when the maximum number of retransmissions decreases, aRLC-AM behavior gets closer to the one exhibited by the RLC-UM. Therefore, from the throughput viewpoint, the proposed scheme is able to offer a performance that stays within the bounds established by the two legacy RLC operations. The figure also shows that the performance of the proposed scheme is severely impacted by the maximum number of retransmissions, since a relevant decrease can be seen when this parameter increases from 1 to 3.

After assessing the correct operation of the proposed solution, the next aspect to be analyzed is the improvement brought about in terms of the EMF exposure. In this sense, Figure 5.32 depicts the number of retransmissions performed by the RLC modes. As expected, when the number of retransmissions that can be used increases, the behavior of the aRLC-AM gets closer to that shown by the RLC-AM. From the point of view of the exposure, the $MaxRetxThreshold$ value should be kept as low as possible.

One of the main goals of the proposed scheme is the differentiated data protection. In order to study this aspect, which might have a relevant impact over the QoS, Figure 5.33 shows the percentage of frames, of the three types, which are correctly received by the application. As can be seen, the losses start to be relevant when the UE is 360 m away from the eNodeB. The figure shows that the loss rate for the legacy RLC-AM is even higher, due to the fact that the simulation finishes before all frames are delivered. It is worth noting that the current LTE implementation does not define any bound for the number of retransmissions in AM mode, so that the UE may keep retransmitting indefinitely. Furthermore, the RLC-UM exhibits the worst performance, being particularly negative for the I-frames (see Figure 5.33b), due to their much larger length.

Furthermore, Figure 5.33 can be used to extract some additional conclusions regarding the modes that include retransmissions. They bring about some benefits, but, on the other hand, the maximum number of retransmissions they establish should not be too high, since that would increase the time spent in this procedure. In this sense, the configuration with 1 retransmission yields the best performance. As for aRLC-AM, we can observe that I-frames are indeed more protected, since the behavior is the same as for the other types of frames, despite the fact that the I-frames are much larger.
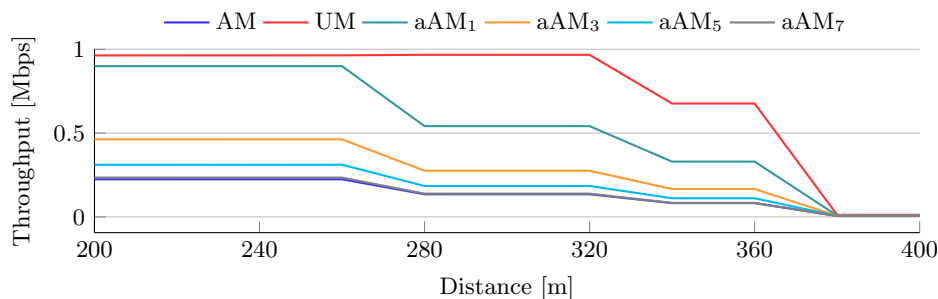
Figure 5.31: Average throughput as a function of the distance between the UE and the eNodeB. Adaptive RLC is represented as aAM$_x$ where x is the value of $MAxRetxThreshold$
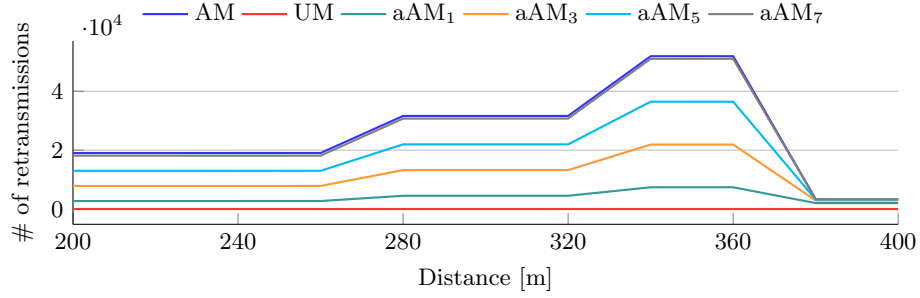
Figure 5.32: Number of retransmissions as a function of the distance between the UE and the eNodeB

(a) P-Frames
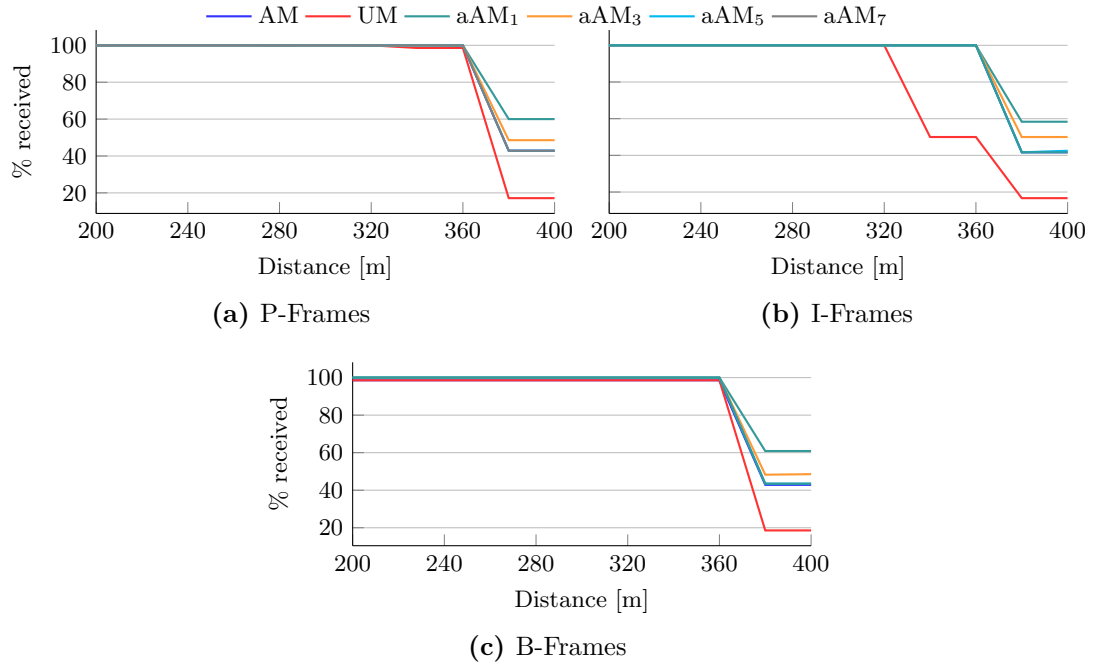
(b) I-Frames

(c) B-Frames

Figure 5.33: Percentage of received frames as a function of the distance between UE and eNodeB

To conclude the analysis, we have studied the impact of the different modes over the MOS. In order to use the MOS metric provided by Eq.5.19, we need to take into account that some of its parameters were tailored for the particular characteristics of the scenario used in [Kha10]. In this sense, to compare the different RLC modes, we can safely assume that the traffic parameters remains the same for all of them. Hence, the impact of the different RLC modes would be limited to the loss rate, while the remaining of the expression does not change, being constant. Thereby, we can study the impact over the MOS by

focusing on the expression provided on Eq. 5.20.

$$MOS_{approx} = \frac{1}{1 + LR^{\frac{1}{MaxData}} + LR^{\frac{2}{MaxData}}} \tag{5.20}$$

Figure 5.34 shows the values of the previous MOS approximation for the different aRLC-AM configurations, comparing them with the performance of RLC-UM. Since we have not established a limit for the maximum number of retransmissions to be used by RLC-AM, the previous expression cannot be used for this mode. We have estimated the LR by multiplying the LR of the different frames (see Figure 5.33) with the probability of each frame type, as given by the GOP pattern previously described. Furthermore, and in order to simplify the analysis, we have normalized the corresponding results.

As can be observed, the aRLC-AM configuration with one retransmission outperforms RLC-UM, since the MOS of the latter starts decreasing at a distance of 320 m (the same behavior was previously seen for the I-Frames, Figure 5.33b). For the rest of aRLC-AM configurations, the results show a rather unpredictable behavior, which can show low MOS values even for good link quality conditions.
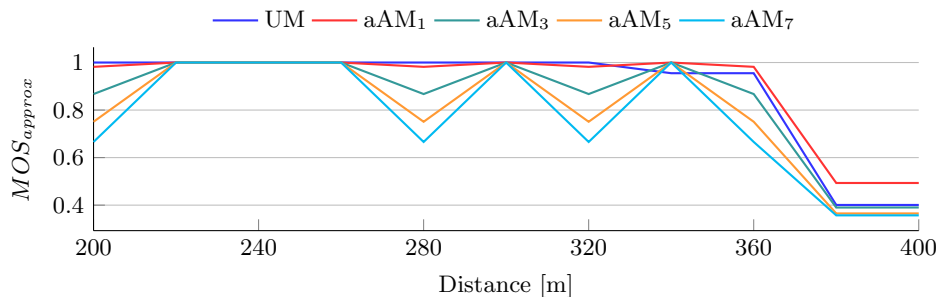
Figure 5.34: Normalized MOS approximation as a function of the distance between the UE and the eNodeB

5.3 Conclusion

Despite the growing concern from the general public regarding the EMF exposure induced by wireless communications, current network management techniques do not consider such exposure within the their KPIs. In this sense, considering the growing presence of wireless terminals, ranging from personal devices or wearables to Wireless Sensor Networks (WSN), it is necessary to integrate the exposure as a parameter of merit in such techniques.

We have tackled the study and modification of two management techniques that are foreseen to be of utter importance in the near future. On the one hand, we have studied the impact that routing decision making in wireless multi-hop networks may have over the population exposure. In fact, the use cases of wireless multi-hop and mesh networks have experienced an upturn in the last years. In this sense, the initial applications foreseen by the scientific community for emergency applications have been surpassed by topologies supporting new paradigms, such as device-to-device or machine-to-machine communications. In the same way, standardization bodies like 3GPP have also extended the application scenarios of such topologies for commercial applications, for instance proximity services, which will have a strong impact over the population exposure.

As for routing techniques, it has been shown that overall exposure reduction can be achieved by using green networking algorithms, and minimizing the transmitted power. Nevertheless, this may lead to quite unfair exposure distribution, so additional techniques, considering the accumulated exposure, are needed to foster a more balanced distribution. This twofold objective has been studied from different angles. First an EMF aware algorithm, considering both exposure minimization and balance, has been presented. In this analysis we have seen that the performance of a particular configuration may depend on the scenario where it is applied. Additionally, we have observed that configurations where the accumulated exposure is not considered after a time (applying an *aging* function) yield a better behavior.

Afterwards, the exposure metrics have been formalized by defining a theoretical model, based on the pace in exposure growing and accumulation. This analysis sets the basis to actually considering exposure from a routing perspective, and the obtained metrics serve as exposure proxies for routing solutions. Based on the previous analysis a routing protocol, namely EARP, has been specified and evaluated. It was designed to consider both the transmitted power and accumulated exposure in the routing decision process.

The assessment of the protocol by means of simulations showed the correct behavior of the proposed scheme, being able to reduce exposure peaks when compared with legacy solutions. Besides, the protocol has been analyzed in terms of the theoretical metrics, yielding results tightly coupled to the particular protocol configuration. As a final conclusion we can state that EARP performance, as was the case for the aforementioned algorithm, strongly depends on the particular network topology.

Along with growing presence of multi-hop wireless networks, data traffic is expected to experience a remarkable increase, in the coming years. Among the services that will lead this growth, video services stand out, since they are foreseen to carry almost 80% of mobile traffic. Hence, management techniques for video content delivery that consider the population exposure may lead to an important overall EMF exposure reduction in the near

future.

The techniques proposed to improve the video content delivery service, from an exposure perspective, have focused on the evaluation of cross-layer network adaptive solutions, where communication protocols tailor their behavior to the communication state and data carried. In particular, we have considered the reference 3GPP scenario for video streaming, embracing LTE as network access technologies and DASH as service solution. As for the protocols, the adaptive techniques have considered the RLC protocol, with its different operation modes, and how it interacts with upper layers.

First, we have evaluated the potential performance of an adaptive transport protocol selection (TCP or UDP), which uses RLC acknowledgments to estimate the communication bandwidth, so that the video server can adapt its bit rate. According to the results, we have observed that the adaptation of service bit rate has rather different impact, depending on the transport protocol: TCP, due to its own congestion control, does not take advantage of the bit rate adaptation, while UDP is greatly favored by it. Besides, and for those cases where rate adaptation is used, we have also observed that, depending on the communications state, which boils down to wireless channel conditions in our scenario, UDP can provide significant QoS improvements when compared to TCP. In fact, the results yield that, in adverse channel conditions and using the adaptive bit rate, UDP provides higher application throughout and byte delivery ratio.

Then, aRLC-AM, an adaptive version of the AM mode of the RLC protocol, has been proposed and evaluated. This new scheme aims to dynamically adapt the ARQ protection to the particular content it carries. Its evaluation, by means of simulations, has shown that configuring RLC with rather low level of protection (i.e. maximum number of retransmissions) strongly reduces the number of packets sent without a relevant impact on the QoE. This behavior is specially remarkable for those video frames that carry less relevant information.

Altogether, we can conclude that the usage of dynamic and adaptive network management solutions can bring significant benefits from a EMF exposure viewpoint, without jeopardizing the QoS. In some cases these techniques might even improve the QoS perceived by end-users.

CHAPTER 6

Summary and Outlook

Information and communications technologies are becoming a key asset in almost every aspect of our lives. Apart from the ever growing usage of communications technologies for human interaction, public and private entities are exploiting ICT to improve the efficiency of their services. This pervasive presence of ICT makes it necessary to consider both technical and societal aspects in the definition of future technologies and their exploitation. Amongst other aspects, the concern about EMF population exposure has recently become a hot topic in many countries, and it may actually hinder the acceptance of new communications systems. Bearing this in mind, this Thesis has addressed some of the most relevant aspects for the design and analysis of EMF-aware networks.

First, two analysis tools have been designed and developed with different objectives. The first one is an EMF sensing tool able to measure the exposure induced by real networks over large areas. Based on a thorough analysis of the scenario and deployed devices, a complete methodology has been defined to exploit the samples provided by the sensing tool, and to use them to compute complex metrics.

Although the sensing tool provides EMF exposure values from real networks, the analysis of new solutions requires different approaches. The second tool is thus a simulation framework, GWNSyM, specially conceived fo study large and complex scenarios. It adopts the main ideas of system-level simulation and provides a set of abstractions to simplify the development of new network models and to enable code re-usability.

Later on, GWNSyM has been exploited to study the impact that different solutions may have over the EMF exposure. In particular, we have focused on techniques that model the performance and usage of radio resources, access selection and resource management, since they have a strong impact on the population exposure. First, we have defined and evaluated two advanced models for radio resources, both for downlink and uplink, which take into account the system dynamics. The evaluation of these models has evinced that typical assumptions, such as constant interference, can yield rather inaccurate results, which would in fact depend on the specific topology. These models have been exploited for the analysis of different access selection techniques. Afterwards, we have proposed a scheduling algorithm for multi-RAT and multi-connectivity scenarios, whose performance was evaluated by means of the simulation methodology fostered by GWNSyM.

Besides the usage of radio resources, we have proposed solutions for other network techniques that may as well impact the population exposure. In this sense, we have tackled the addition of EMF exposure in the routing decision process for multi-hop wireless networks. Considering different aspects of the exposure as far routing is concerned, a

theoretical framework has been developed, and has been used to design and evaluate a routing protocol, EARP. On the other hand, taking into account the notable relevance of video services in the data traffic volume, we have proposed two cross-layer solutions for video content delivery. In particular, the proposed techniques seek to optimize the interaction between transport and link layers protocols, in order to reduce the number of transmissions, while keeping the QoS.

Below, we highlight the main contributions of this Thesis. Afterwards, we depict the research lines that remain open, either as extensions or additional aspects derived from such contributions.

6.1 Contributions

In the following we summarize the main outcomes of this dissertation:

- Design and development of an EMF sensing tool able to continuously monitor the EMF exposure over a large area (Santander downtown). The tool consists of a distributed sensing system, which is integrated within a smart-city testbed, thus leveraging the management functionalities and data precessing facilities provided by it. In order to optimize the sensor deployment, a thorough analysis of the particular scenario has been performed. Afterwards, a set of uncertainty sources have been identified, accounting for different factors (installation, technical characteristics of the sensors, etc.). All in all, a complete measurement methodology has been defined: (1) to improve the accuracy of the samples obtained by the sensors, and (2) to compute different components of the LEXNET EI metric. Based on such methodology, the EMF exposure over the area during one year has been analyzed to characterize the EMF behavior and its potential spatio-temporal relationships. The framework also features a web portal from which the measurements provided by the sensing tool can be retrieved. The portal is publicly available and can be thus used by the scientific community.

- Design and implementation of GWNSyM, a simulation framework that features systematic simulation-based analysis of large systems. Based on the system-level simulation paradigm, the framework provides the core procedures and generic semantics. It is designed to ease the addition of new network solutions and to perform architectural changes based on the scenario configuration. Furthermore, it has been designed to simplify code re-usability, by means of a minimal, on demand, binding between the framework and network models. In this sense, it simplifies the addition of legacy simulation models, as well as the reuse, in other tools, of GWNSyM modules.

- To validate GWNSyM, an optimum multi-criteria access selection policy has been defined. We have posed a binary linear optimization problem that considers both operator preferences and QoS. The access selection policy can be tailored, according to the particular scenario, and different configurations have been analyzed over large heterogeneous scenarios.

- In order to appropriately model the EMF exposure on simulation environments, advanced models for radio resources have been defined, for both downlink and uplink.

In general, both models consider the effect that access selection decisions have over radio resource efficiency, capturing system dynamics. The models are translated to optimization problems (geometric and linear), with tractable solutions. The evaluation of these models reveals that assuming constant interference, may yield rather inaccurate performance.

- Based on the aforementioned models, we have assessed the impact that traffic offloading, from cellular to WiFi networks, has over the population exposure. In this case, the results obtained by GWNSyM are part of a larger evaluation methodology, which also involves radio planning tools that accurately model radio aspects. The results obtained by GWNSyM, mainly focused on QoS evaluation, show that traffic offloading remarkably reduces population exposure while keeping the QoS. However, depending on the operator policies, which may prevent some services from exploiting the WiFi capacity, the quality of service perceived by users may become quite unbalanced.

- Afterwards, different access selection techniques, for dense and heterogeneous cellular networks, have been studied, focused on the uplink communication, and exploiting the radio resource models. It was seen that the decoupling of downlink and uplink would bring about remarkable benefits, both in terms of transmitted power and QoS. On the other hand, the use of DUDe causes a higher number of handovers, which may jeopardize the QoS.

- Resource management in heterogeneous networks has been also studied by defining a scheduling algorithm for multi-connectivity scenarios, LaSR. This work has been filed as a provisional European patent, in collaboration with NEC Laboratories Europe, in March, 2017. The results evince that the proposed algorithm is able to fulfill the requirements of multi-connectivity scenarios. One of the most relevant features of the proposed design is that it is able to consider practical constraints, while ensuring the system stability.

- A framework to include EMF exposure into the routing decision process. First, an algorithm that considers both instantaneous and accumulated exposure was designed. Afterwards, a theoretical analysis was carried out to derive metrics that can be used to assess the performance of routing protocols in terms of exposure. Finally, based on such algorithm, we developed a routing protocol, EARP. It was evaluated, in a simulation environment, based on the aforementioned metrics. According to the results, the protocol is able to adapt its behavior according to past situation and events, in contrast to static protocols such as AODV. Furthermore, the results evince that the use of *aging* functions, which strengthen near-past decisions, brings about benefits in terms of exposure.

- Finally, we have proposed two cross-layer solutions to reduce the use of radio resources for video services in cellular networks. In particular, the techniques seek to optimize the interaction between transport and link layer (RLC) protocols as well as the characteristics of video content. The results show that a dual stack in which either TCP or UDP can be used as transport protocols, would provide benefits in terms

of both exposure and QoS. Furthermore, the analysis evinces that unequal error protection in the RLC, based on the particular packet content, may also provide benefits. In fact, we have shown that losses might be acceptable in the least relevant packets without a remarkable impact on the QoS.

The research carried out during this Thesis has been the basis for a number of publications, which cover all the contributions enumerated above, as can be seen in Annex D.

6.2 Future lines of research

In the following we enumerate some research lines that remain open after the work conducted during this Thesis:

- Exploitation of the sensing tool. Further characterization of EMF exposure in Santander would allow us to identify spatio-temporal dependencies, and its relationship with social events or periods of the year.

- Exploit the statistical exposure levels measured by the sensing tool in simulation models. In order to fill the gap between exposure levels (working with E-field values) and simulation (resources) other tools would need to be employed. At the time of writing we are using SDR solutions to measure the usage of resources by base stations in different scenarios. These measurements, obtained for each individual cell, can be used to extrapolate the relationship between resource utilization and EMF exposure.

- The simulation framework has been made available for the community. Hence, it will be further exploited to perform large simulation analysis on different types of systems. Besides the items discussed below, additional analysis, focused on the front-haul and back-haul of cellular systems, in particular assuming C-RAN architecture, will be carried out.

- As for the advance resource modeling, we have so far assumed that each user can be connected to just one access element. However, as has been mentioned, multi-connectivity is a promising technique in future scenarios. Hence, the resource models will be revised in order to identify the required modifications to consider multi-connectivity.

- Exploiting the resource models, further analysis will be performed for access selection strategies. It is an ongoing work, where techniques as DUDe and CRE, will be further studied over a variety of scenarios.

- Although the analysis of access selection techniques based on simulation can provide valuable information, it is usually limited to the particular scenario under study. On the other hand, it might be interesting to provide performance bounds for each access selection policy, regardless of the particular scenario, to rank the different policies in terms of specific performance parameters.

- The scheduling algorithm will be further analyzed. Although the results are already interesting and validate algorithm behavior, its performance in realistic scenarios is still ongoing. In this sense, the implementation of the algorithm in more complex

simulation environments is being studied. In particular, we are investigating how to include it within the *ns3* simulator, equipped with LTE modules. However, the LTE implementation does not support carrier aggregation enabling cross-carrier scheduling yet, nor implementation of mmW communications. Hence, it will be necessary to clearly identify the modifications to undertake in the simulator. Other future work related to the multi-connectivity scheduling is the usage of SDR to assess the algorithm over real systems.

- The performance of EARP will be compared with other protocols fostering load balancing. In this sense, the theoretical framework will allow us to evaluate the different routing alternatives in terms of EMF exposure.

- Following our research on EMF aware techniques for video content delivery, it would be interesting to provide dual solutions, where either legacy or new techniques are used, according to the particular circumstances, exploiting cognitive and cross-layer techniques. Both the dynamic transport protocol selection and the adaptive radio link scheme could be jointly studied, in order to better understand the interactions between transport and link layer protocols in cellular networks.

A Example of GWNSyM scenario definition

This annex shows the usage of GWNSyM with a practical example that can be found in the code repository[1], so as to complement the simulation and design principles provided in Section 3.1. We will focus on the definition of *NetTypes* and their interaction with *NetActions*, leaving aside the implementation of actual models.

The scenario consists of a set of users and macro base stations, which will comprise a number of cells. All scenarios require a basic configuration that establishes the number of snapshots to be performed during the system-level simulation, as well as the number of entities that define the first level of the dependencies tree, see Figure 3.2. Figure A.1 shows the example basic configuration, in one of the currently accepted formats (so far *INI* and *XML* formats are implemented). Is it worth noting that the tool can be customized with other configuration formats, if the appropriate parser is created. The first tag of the configuration file defines the total simulation time, and the duration of a snapshot. In the example, the simulation takes 10 time units, while each snapshot is 2 units, so that 5 snapshots will be executed. Furthermore, the configuration file indicates that the scenario will include 2 instances of both *NetTypes MACRO* and *USER*.

Next, as depicted in Figure A.2, we define the *C++* classes that define both *NetTypes* and the *NetActions*. As can be observed, GWNSyM allows using plain classes (or structs) and does not require any hierarchy. Opposed to that, the user needs to be aware of the interface that will be used if implemented. In particular, Figure A.2a shows the definition of both models and configurations for the three *NetTypes* in the scenario. In the case of the models (classes *ExampleUser*, *ExampleMacro* and *ExampleCell*) we have only added the interface that links the instances with their configuration (*SetConfiguration*), which gets called when the instances are created, so that the user can keep a reference to read the configuration when needed. On the other hand, classes implementing the configuration

```
1  [TIME]
2  TOTAL=10
3  SNAPSHOT=2
4
5  [MACRO_INST]
6  NUMBER=2
7
8  [USER_INST]
9  NUMBER=2
```

Figure A.1: Example of configuration file, using INI format.

---

1 https://github.com/ldiez/GWNSyM

```
1  // Definition of NetType for cells
2  struct ExampleCellConf { };
3
4  struct ExampleCell {
5      void SetConfiguration(ExampleCellConf const& conf) { }
6  };
7
8  // Definition of NetType for macro base stations owning cells
9  struct ExampleMacroConf {
10     InnerItemsCounterList_t ReadInnerConf(void) const {
11         return {{"CELL", 3}};
12     }
13 };
14
15 struct ExampleMacro {
16     void SetConfiguration(ExampleMacroConf const& conf) { }
17 };
18
19 // Definition of NetType for users
20 struct ExampleUserConf {};
21
22 struct ExampleUser {
23     void SetConfiguration(ExampleUserConf const& conf) { }
24 };
```

(a) *NetTypes* to implement macro base stations, cells and users

```
1  struct ExampleAction {
2      void operator()(Ptr_t<ExampleCell> m, Ptr_t<ExampleUser> u) {
3          std::cout << "One-to-one" << std::endl;
4      }
5  };
6
7  struct ExampleAction2 {
8      void operator()(Vec_t<ExampleCell> m, Ptr_t<ExampleUser> u) {
9          std::cout << "Many-to-one" << std::endl;
10     }
11 };
12
13 struct ExampleAction3 {
14     void operator()(Vec_t<ExampleCell> m, Vec_t<ExampleUser> u) {
15         std::cout << "Many-to-many" << std::endl;
16     }
17 };
```

(b) *NetTypes* to implement macro base stations, cells and users

Figure A.2: Minimal working example of *NetTypes* and *NetActions* definition

are left empty, but for the macro base stations. In this case, the configuration defines the set of elements belonging to macro base stations, lower level in the dependency tree.

Besides, Figure A.2b illustrates the definition of *NetActions*, which are consumed by users and cells in different ways according to their interface: one-to-one, many-to-one and many-to-many semantics.

Based on the implemented classes, Figure A.3 shows the actual implementation of the scenario. After creating an instance of the system at line 3, the user needs to define the *NetTypes*, indicating model and configuration instance for each of them, as well as the name for the *NetType*, lines 5, 6 and 7. Next, in lines 9 and 10, we indicate GWNSyM the *NetTypes* that will define the base of the dependency tree. For instance, line 9 establishes that the number of *MACRO* instances are defined in the configuration file path *MACRO_INST→NUMBER*, see Figure A.1. Then, we specify the parser for the configuration file and deploy the scenario in line 13. At this point, the dependency tree is created and the elements of the different *NetTypes* are instantiated according to both the basic configuration (i.e. the file) and the configuration of the different types. Although in this example the classes that define the *NetTypes* configuration are very simple, they would, in general, use some method to read the configuration from external sources (e.g. local or remote files). Before defining the *NetActions*, line 14 prints out the tree of instances (instantiation of the dependency tree based on the configuration).

Figure A.4 depicts the resulting tree of instances, whose first level consists of the elements defined as the base of the tree (lines 9 and 10 of Figure A.3). Then, GWNSyM recursively reads the configuration of each *NetType*, and creates the corresponding instances, resulting

```
1  #define GLOB_CONF "./config-files/EXAMPLE_CONF.cfg"
2  int main() {
3      System network;
4      // Definition of NetTypes
5      network.Type<ExampleUser, ExampleUserConf>("USER")
6              .Type<ExampleCell, ExampleCellConf>("CELL")
7              .Type<ExampleMacro, ExampleMacroConf>("MACRO")
8      // Setting global configuration
9              .SetTreeBase("MACRO",{"MACRO_INST", "NUMBER"})
10             .SetTreeBase("USER",{"USER_INST", "NUMBER"})
11             .SetConfig<Configurer>(GLOB_CONF)
12     // Deploy network
13             .Deploy()
14             .PrintTreeInstances();
15     // Apply actions
16     network.Action<ExampleAction>({"MACRO::*::CELL", "USER"})
17             .Action<ExampleAction2>({"MACRO::*::CELL", "USER"})
18             .Action<ExampleAction3>({"MACRO::*::CELL", "USER"})
19     // Start simulation
20             .Run();
21     return 0;
22  }
```

Figure A.3: Scenario description using the *NetTypes* and *NetAction* defined.

```
1  ============= INSTANCIATED TREE =============
2  |-USER [2]
3  |-MACRO [2]
4      |-MACRO::CELL [3]
5  ============================================
```

Figure A.4: Tree of instances as a consequence of the scenario in defined in Figure A.3

in the inner branches of the tree. In the example, a set of 3 elements of the type *CELL* is created for each element of the type *MACRO*, as indicated in Figure A.2a, line 11.

Finally, we indicate which *NetActions* will be executed during the simulation, as shown in Figure A.3 from line 16 onwards. As can be observed, we make use of the *set resolution operator* to indicate the sets of elements that need to be passed to the actions. In particular, the three defined actions accept the same two sets: six elements of type *CELL* (three instances from each of the two *MACRO*) and two elements of type *USER*.

Then, the system internally carries out the inference process described in Section 3.1.2, so that each action is called accordingly. As a consequence, in each snapshot the message *One-to-one* from action *ExampleAction* is printed 12 times, corresponding to the iteration over the elements of both sets (6 and 2 elements respectively). Likewise, the message *Many-to-one* from action *ExampleAction2* is printed twice, and the third action (*ExampleAction3*) is called just once per snapshot.

B Cycle canceling algorithm

In order to better understand the implementation of the EMF aware routing algorithm, we present here a practical example of its core operation (CCA), by using a simple graph. As depicted in Figure B.1, the graph consists of 4 nodes, where 2 of them are traffic sources and the other 2 are potential destinations.

The links are represented by arcs with an associated cost (ω) that is proportional to the flow traversing the arc. Besides, arcs are also given a capacity (in number of flow units). In order not to limit the algorithm, this capacity is set to a value greater than the total number of unit flows, so that it is never saturated. In the following, every arc connecting two nodes i, j, is assigned a pair of numbers $(c_{i,j}, \omega_{i,j})$, where $c_{i,j}$ and $\omega_{i,j}$ represent the capacity and cost of such arc respectively, as depicted in Figure B.1.

First, as commented in Section 5.1.1 the initial graph is transformed into the one shown in Figure B.2. The steps to transform the graph are as follows:

- Virtual sources and destination are added to the initial graph. As shown in Figure B.2, the links leaving the virtual source (S) have no cost and capacity 1, while the links entering the virtual destination D have no cost an a capacity greater than the total number of active flows (2 in our examples). Besides, we denote the traffic generation of a node N as $b(N)$. Hence, traffic demand is assigned to both virtual source and destination to ensure the actual traffic is accommodated in the graph.

- In order to allocate the accumulated exposure to links rather than nodes, each initial node N is split into two virtual nodes (N' and N'') connected by a link without capacity limitations and a cost that equals the cumulated exposure (ω_{acc}).
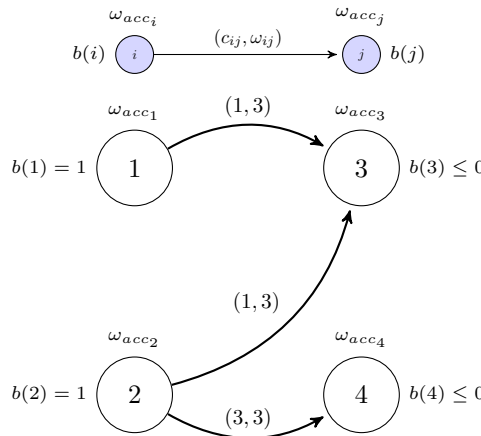
Figure B.1: Actual graph
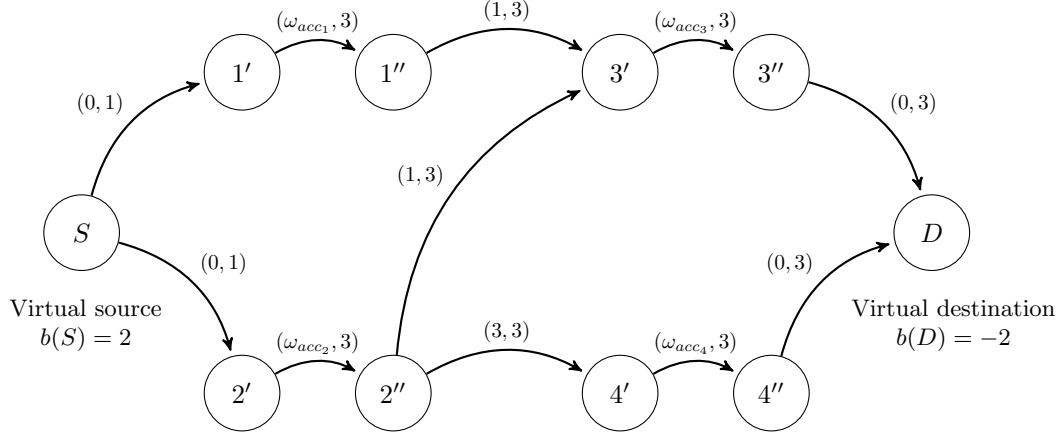
Figure B.2: Transformed graph
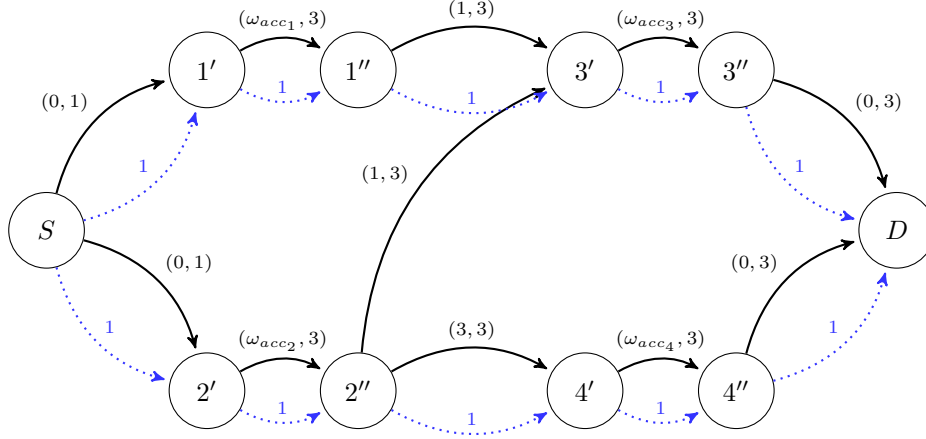
Figure B.3: Arbitrary flow assignment

Once the graph is transformed, the algorithm is applied as indicated in Section 5.1.1. First, a feasible solution is chosen to allocate the flows. In this sense, Figure B.3 depicts a possible initial allocation, showing the amount of flow assigned to each link, while the cost associated to this allocation, Ω, is calculated as follows:

$$\Omega = \left(0 + 1 + 0 + 0 + 3 + 0\right) \cdot 1 + \omega_{acc_1} + \omega_{acc_3} + \omega_{acc_2} + \omega_{acc_4} \tag{B.1}$$

Iterating the algorithm described in Section 5.1.1, we obtain a residual graph and locate negative cycles. As shown in Figure B.4, for this example the cycle $4' \to 2'' \to 3' \to 3'' \to D \to 4'' \to 4'$ has been chosen, with an associated cost $\left(\left(-3 + 1 + \omega_{acc_3} + 0 + 0 - \omega_{acc_4}\right)\right)$ equal to -2, since $\omega_{acc} = 0$ for all links (no accumulated exposure).

Then, the minimum capacity of the cycle is sought, so as to accommodate the flow in the cycle. In the example, as can be seen in Figure B.4, the minimum capacity is 1.
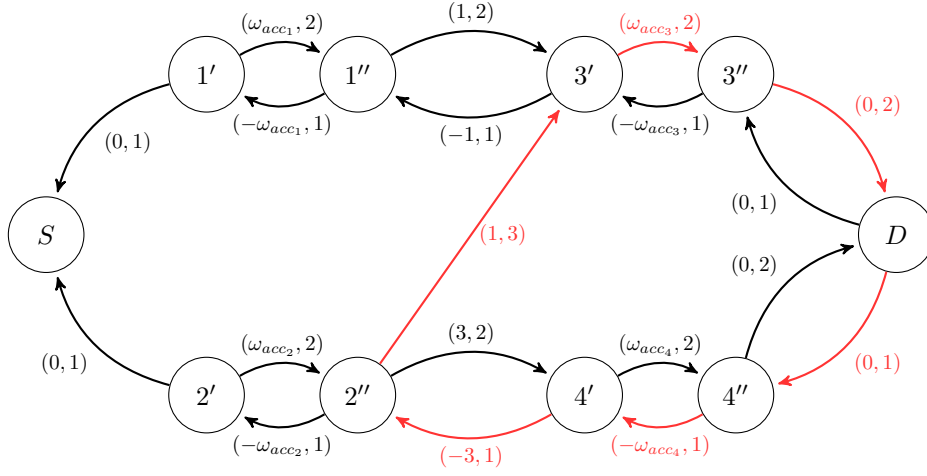
Figure B.4: First residual graph

Next, the residual graph is again computed considering the flow introduced in the negative cycle. This modification is depicted in Figure B.5. At this point, and according to the CCA description, we check whether more negative cycles exist. In our example, since the residual graph has no more negative cycles, the algorithm finishes.

From the residual graph, it is simple to identify the routes assigned to the flows in the initial graph. This is shown in Figure B.6, where the final flow allocation has the associated cost calculated in Eq. B.2, which, as expected, is lower (or equal) than the one initially assigned (Eq. B.1).
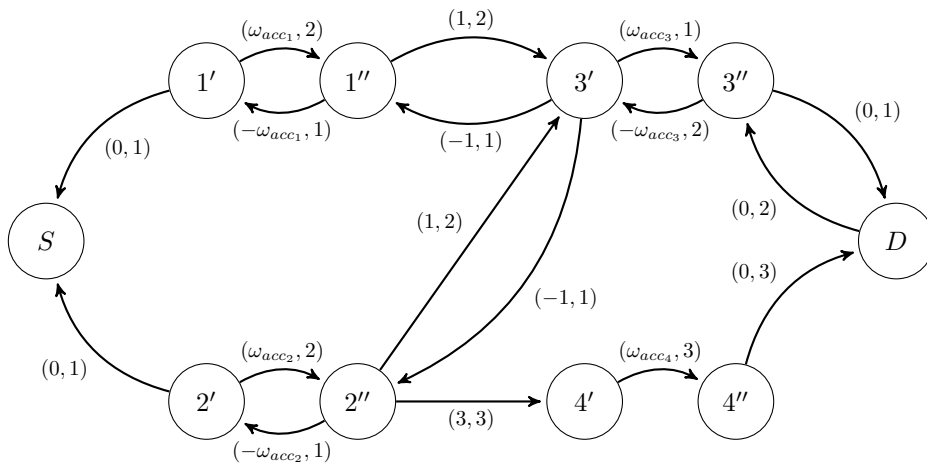
Figure B.5: Second residual graph
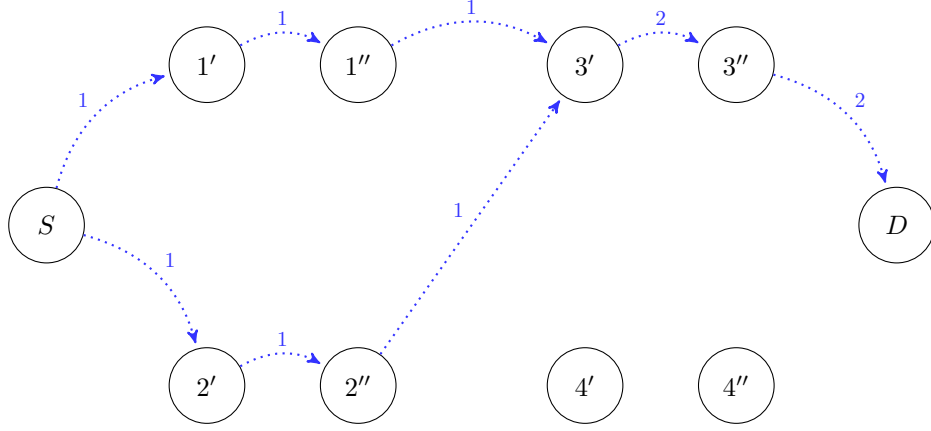
Figure B.6: Final assignment in the virtual graph

$$
\begin{aligned}
\Omega &= (0 + 1 + 0 + 1) \cdot 1 + (0) \cdot 2 + \omega_{acc_1} + \omega_{acc_2} + \omega_{acc_3} \Rightarrow (k_3 = 0, \text{actual destination}) \\
&= 2 + \omega_{acc_1} + \omega_{acc_2} \Rightarrow (\omega_{acc_1} \& \omega_{acc_2}, \text{constants}) \\
&= 2 + \text{Constant}
\end{aligned} \tag{B.2}
$$

Finally, the initial transformation applied to the graph is reverted, so that we obtain the actual routing decision depicted in FigureB.7.
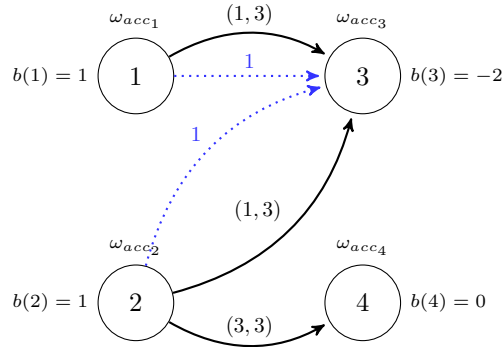
Figure B.7: Final assignment in the actual graph

C Further EARP specifications

C.1 Dynamic power adjustment

In this section we explain the implementation of the dynamic power management used by EARP, covering both the calculation of the transmission power based on the next hop information as well as how this information is sent to the routing protocol through a cross-layer solution.

As was mentioned in Section 5.1.3, we have used *ns3* as simulation environment to implement and evaluate EARP. The characterization has been carried out with a simple setup of 2 nodes equipped with *802.11b* interfaces configured as shown in Table C.1. We establish a traffic flow between the nodes, saturating the channel, and we measure the received power and the throughput at the receiver during 5 seconds. The measurements are obtained for different transmission power levels and distances. In particular, the transmission power is within the range $[-1, 16]$ dBm and the distance varies from 0 to 55 m. Once all the measurements are obtained, the power level selected for each distance is the minimum one that guarantees that the average throughput is 90% of the maximum.

We thus obtain the distance for each received power level, as well as the minimum transmission power that has to be used for each distance. Figure C.1 shows the mapping functions that have been obtained with the aforementioned methodology. It is worth highlighting that, while the particular values depend on the propagation and loss models, the methodology is independent and could be applied indistinctly.

Based on these results, and by means of lookup tables, when a node gets a *HELLO* message from a neighbor, the received power is mapped to distance and then the transmission power needed to send data to that neighbor is stored in the corresponding routing entry, whose structure is described in Section C.2.

Table C.1: Configuration of the characterization scenario

Technology	*802.11b*
Traffic flow	UDP at 11 Mbps
Packet size	1472 bytes
Max supported thput.	≈ 6 Mbps
Min allowed thput.	5.4 Mbps

(a) Rx. power Vs. distance   **(b)** Distance Vs. Tx. power
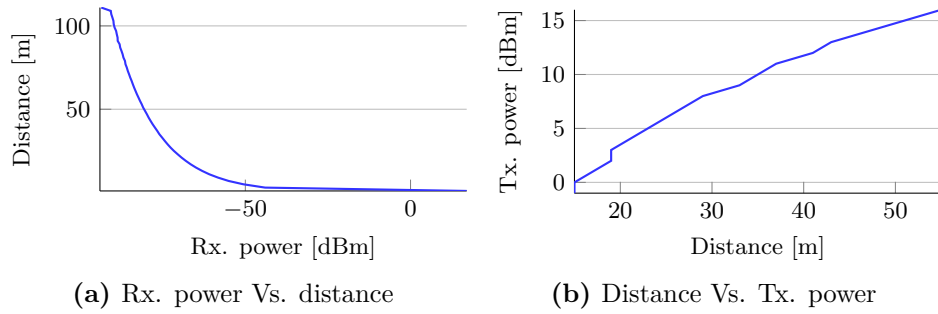
Figure C.1: Mapping between received power and distance, and between distance and power, to assure 90% of the maximum capacity

C.2 EARP data structures

Here we briefly describe the data structure used by EARP to keep the list of neighbors and store information regarding the routes. Note that, as mentioned in Section 5.1.3, EARP basic functionality is based on AODV, so the data structures are similar.

- Routes table: each node stores a table with all routes established so far. Each entry has the following fields:
  - *Destination*: IP address of the destination node.
  - *Cost*: overall cost of the route entry. It is independent of how the cost is calculated.
  - *Number of hops*: number of hops to reach the destination. It may be used to establish limits for the route length.
  - *Sequence number*: it is used to refresh routes and avoid cycles.
  - *Time stamp*: it stores the last time the route was refreshed. It is used to discard entries not used during a given time. Its value is configurable.
  - *Gateway*: IP address of the next hop towards the destination.
  - *State*: it indicates the route state and is used to update the cost. It can take values *InUse*, *Active*, *Active_InUse* or *Valid*, as described in Section 5.1.3 .
  - *Transmission power*: power level (in *dBm* and integer value) that will be used to send a packet to the next hop indicated by *Gateway*.

- Inverse route table: it is a route table whose entries are created during the routing discovery process by intermediate nodes. A timer is linked to each entry, so that if the timer expires the entry is removed. Conversely, if the intermediate node has to forward an acknowledgment of a route discovery associated to an entry, that entry is moved to the routes table. This way, all nodes would keep the corresponding routing information at the end of the route discovery. The main difference of this table with the previous one is the time elapsed until the entry is removed.

- Neighbors table: this table stores information related to the nodes within the coverage range of a given node, so that they can be reached in one hop. The table is updated when receiving *HELLO* messages, and each entry has the following fields:

  - *Id*: IP address of the neighbor.
  - *Cost*: cost to reach the neighbor.
  - *Sequence number*: it stores the last sequence number seen from the neighbor.
  - *Time stamp*: it is used to control the temporal validity of the entry.
  - *Transmission power*: transmission power to be used to reach the neighbor.

C.3 EARP logic

This section describes the logic behind the most relevant routing procedures: neighbor discovery, route discovery and cost change. The last one was introduced in Section 5.1.3 with an illustrative example.

The explanation will be complemented with flow diagrams, using the acronyms defined in Table C.2. First, we will describe the neighbor discovery procedure, which exploits the aforementioned transmission power estimation process to obtain route cost metrics. Afterwards, the route discovery procedure will be described discussing how the discovery is triggered, as well as the procedures initiated upon receiving both *ROUTE_DISC* and *DISC_ACK*. Finally, we will describe the logic implemented to perform the cost change illustrated in Section 5.1.3.

Neighbor discovery

Messages of type *HELLO* are periodically sent to keep updated information in the neighbor table. In the following we will describe the procedure triggered upon the reception of a *HELLO* message, leaving aside the case where it is sent to update route cost; this is, excluding the cases with flag $M = 1$.

The procedure is described in Figure C.2 with a flow diagram. Upon the reception of the *HELLO*, the received power is used to estimate the required transmission power.

Table C.2: Simplified nomenclature used in the flow diagrams of the EARP logic. Besides, elements in tables will be indicated using square brackets (i.e. $Ro_t[D]$ corresponds to the routing table entry towards D)

Acronym	Description
RD	message *ROUTE_DISC*: $RD(S, D)$ means route discovery from S to discover D
ACK_{RD}	message *DISC_ACK*: $ACK_{RD}(S, D)$ acknowledgment to route discovery from S to discover D
Ne_t	Neighbors table
Re_t	Reverse routing table
Ro_t	Routing table
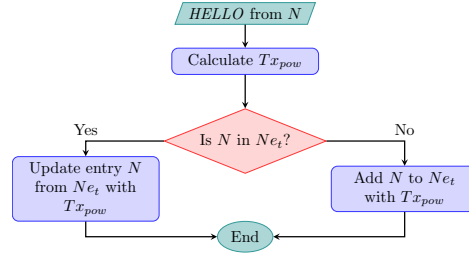
Figure C.2: Flow diagram for *HELLO* reception

Besides, the estimated value will be also stored in the neighbor table and afterwards used to calculate the routing cost. Later on, it checks whether an entry exist for the sender of the *HELLO* in the neighbors table, updating it accordingly.

C.3.1 Route discovery

The route discovery procedure is triggered when a node has a packet to be sent, as depicted in Figure C.3. Upon this event, the node checks whether a route exists for the destination. In such case, it updates the time-stamp of such route to keep it alive and it ensures that it has an *Active* state. Otherwise, a new route discovery procedure is triggered, using an updated sequence number. Furthermore, a timer is started at the same time the *ROUTE_DISC* is sent, to control that it is not kept waiting for the *ROUTE_ACK*. Later on, two events can happen: the timer expires of the *ROUTE_ACK* is received. In the former case, it is assumed that the destination is unreachable, and the packet is dropped (and upper layers are notified accordingly). Otherwise, the timer is removed and the packet is sent, according to the updated routing table.

Following the previous description, Figure C.4 depicts the logic when a *ROUTE_DISC* is received. First, the receiver checks whether it is the target node. In such case, it creates a new entry (or updates if it existed) in its routing table towards the source of the route
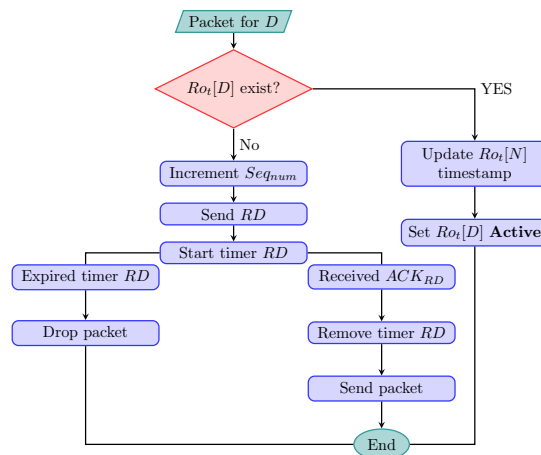
Figure C.3: Flow diagram to start route discovery

discovery, and answers it with the *DISC_ACK*. On the other hand, if the receiver is not the destination of the discovery, a new entry is created in the reverse table, or an existing one is updated, and the *ROUTE_DISC* is forwarded with new information.

When a node receives an *ACK_DISC* message, it has also to check whether it is the node that launched the discovery procedure, as depicted in Figure C.5. During the way forward of the route discovery, intermediate nodes created an entry in the reverse table, corresponding to the source of the discovery, as explained in Figure C.4. Then, if an intermediate node receives an *ACK_DISC*, it means that it will take part in the new route. Hence, it has to: forward the message towards the source (note that *ACK_DISC* is a unicast message), and move the corresponding entry from the reverse to the routing table. Besides, intermediate nodes and the source of the route discovery need to create an entry in the routing table corresponding to the destination. This way, both the source and the intermediate nodes get aware of the required routes, forwards and backwards, to transport packets.

C.3.2 Cost change

As mentioned in Section 5.1.3, the procedures to keep the table of neighbors as well as to perform route discovery is mostly taken from AODV. In addition, EARP implements route dynamic monitoring and cost of change, in order to balance the network exposure over time.

In this sense, every node in the multi-hop network periodically monitors the accumulated level of exposure as described in Figure C.6a. If the exposure has changed in the interval, the routing table is updated, and a *HELLO* message with $M = 1$ is sent, starting the cost change mechanism. Otherwise, the node keeps waiting. As can be observed, EARP does not describe how the cost balance has to be implemented, but it provides the mechanisms to take it into account. As was discussed in Section 5.1.3 different criteria can be used for the exposure change, such as threshold values, aging functions, etc.

Then, when neighbors receive the *HELLO* with the flag $M = 1$, they update the corresponding entry in neighbor and route tables. Besides, the change in the cost would
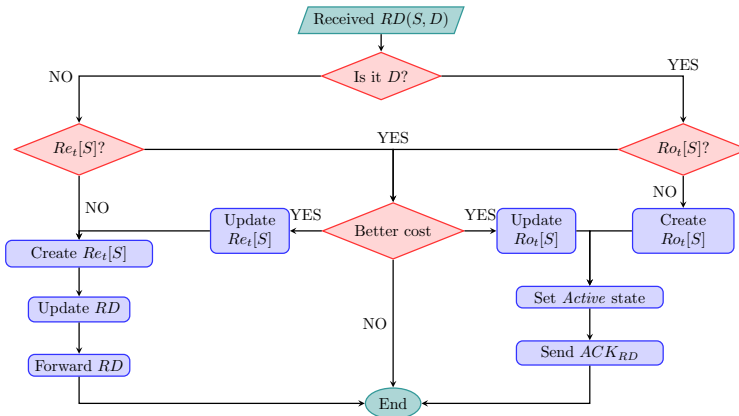
Figure C.4: Flow diagram upon reception of *ROUTEDISC* message
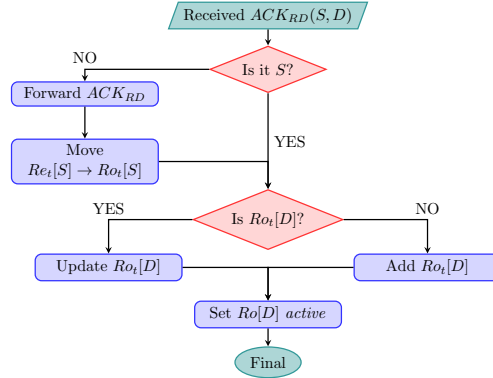
Figure C.5: Flow diagram upon reception of *DISCACK* message

(a) Exposure monitoring loop

(b) Reception of *HELLO* for cost change
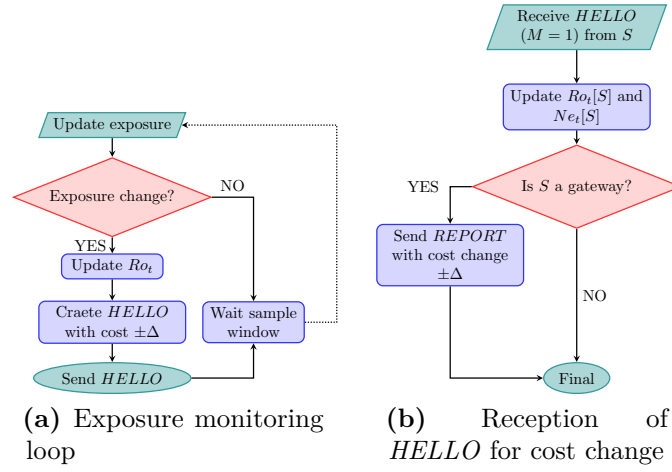
Figure C.6: Flow diagrams to start and propagate a cost change
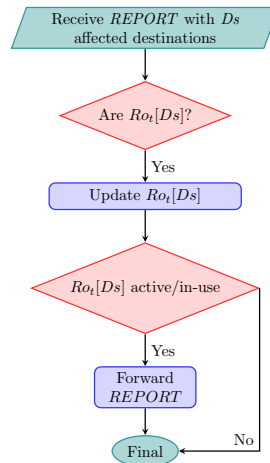
Figure C.7: Flow diagram upon reception of *REPORT* message

also affect all the routes that have the source of the *HELLO* as gateway, so that the cost change needs to be propagated, as was explained in Section 5.1.3.

Finally, when a node receives a *REPORT* message, it checks whether any of the current entries in the route table are affected, updating them accordingly, following the diagram in Figure C.7. Besides, the receiver would forward the cost report provided any of the affected entries have state *Active*, *InUse* or both. This way the cost change is propagated as needed, according to the information stored in the routing tables.

D Publications

We provide here a list of publications that were made based on the results obtained during the research carried out within this Thesis.

International journals

1. Choque, J., L. Diez, A. E. García, R. Agüero, and L. Muñoz: 'Optimum Multi-service Access Selection over Heterogeneous Wireless Networks'. *Trans. Emerg. Telecommun. Technol.* (Mar. 2016), vol. 27(3): pp. 328–338.

2. Diez, L., R. Agüero, and L. Muñoz: 'Electromagnetic Field Assessment as a Smart City Service: The SmartSantander Use-Case'. *Sensors* (2017), vol. 17(6).

3. Diez, L., G. P. Popescu, and R. Agüero: 'A Geometric Programming Solution for the Mutual-Interference Model in HetNets'. *IEEE Communications Letters* (Sept. 2016), vol. 20(9): pp. 1876–1879.

4. Domenico, A. de, L. Diez, R. Agüero, D. Kténas, and V. Savin: 'EMF-Aware Cell Selection in Heterogeneous Cellular Networks'. *IEEE Communications Letters* (Feb. 2015), vol. 19(2): pp. 271–274.

International conferences

1. Agüero, R., J. Choque, L. Diez, and L. Muñoz: 'Binary linear programming formulation of the access selection problem in multi-service and heterogeneous network environments'. *2013 IFIP Wireless Days (WD)*. Nov. 2013: pp. 1–4.

2. Brau, M., J. Stéphan, L. Diez, Y. Corre, Y. Lostanlen, and R. Agüero: 'Assessing the WiFi offloading benefit on both service performance and EMF exposure in urban areas'. *2016 Wireless Days (WD)*. Mar. 2016: pp. 1–4.

3. Choque, J., L. Diez, R. Agüero, and L. Muñoz: 'Providing differentiated QoS with dynamic pricing policies for wireless access networks'. *2013 IFIP Wireless Days (WD)*. Nov. 2013: pp. 1–4.

4. Diez, L., R. Agüero, and J. Penhoat: 'Reducing the Electromagnetic Exposure over LTE Networks by Means of an Adaptive Retransmission Scheme: A Use Case Based on a Video Service'. *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*. May 2015: pp. 1–5.

5.   Diez, L., R. Agüero, J. Penhoat, E. Stéphan, and F. Heliot: 'Exploiting the DASH Framework to Promote a Dynamic Transport Selection for Video Streaming over LTE Networks'. *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing.* Oct. 2015: pp. 37–43.

6.   Diez, L., S. M. Anwar, L. Rodríguez de Lope, M. L. Hennaff, Y. Toutain, and R. Agüero: 'Design and integration of a low-complexity dosimeter into the smart city for EMF assessment'. *Networks and Communications (EuCNC), 2014 European Conference on.* June 2014: pp. 1–5.

7.   Diez, L., J. Arroyo, and R. Agüero: 'On the use of emulation techniques over the ns-3 platform to analyze multimedia service QoS'. *LNICST.* Vol. 158. 2015.

8.   Diez, L., Y. Corre, B. Radier, M. Brau, and R. Agüero: 'Complementing radio-planning tools to analyze EMF-aware access selection'. *2015 International Symposium on Wireless Communication Systems (ISWCS).* Aug. 2015: pp. 431–435.

9.   Diez, L., J. Igareda, V. Iancu, E. Slusanschi, and R. Agüero: 'Routing algorithm to fairly distribute the exposure to electromagnetic fields over wireless mesh networks'. *2015 IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob).* Oct. 2015: pp. 468–473.

10.   Diez, L., L. Rodríguez de Lope, R. Agüero, Y. Corre, J. Stéphan, M. Brau, S. Aerts, G. Vermeeren, L. Martens, and W. Joseph: 'Optimal dosimeter deployment into a smart city IoT platform for wideband EMF exposure assessment'. *Networks and Communications (EuCNC), 2015 European Conference on.* June 2015: pp. 528–532.

11.   Diez, L., S. Sergio Izúel, and R. Agüero: 'Generic Wireless Network System Modeler: Fostering the Analysis of Complex LTE Deployments'. *LNICST.* Vol. 191. 2017.

12.   Iancu, V., L. Diez, L. Rodríguez de Lope, E. Slusanschi, and R. Agüero: 'A reward-based routing protocol to reduce the EMF exposure over wireless mesh networks'. *2014 IFIP Wireless Days (WD).* Nov. 2014: pp. 1–4.

13.   Rubio-Loyola, J., L. Gonzalez-Hernandez, L. Diez, R. Agüero, and J. Serrat: 'An energy-oriented optimization algorithm for solving the cell assignment problem in 4G-LTE communication networks'. *2014 IFIP Wireless Days (WD).* Nov. 2014: pp. 1–4.

14. Sarrebourse, T., L. Rodríguez de Lope, A. Hadjem, L. Diez, S. M. Anwar, R. Agüero, Y. Toulain, and J. Wian: 'Towards EMF exposure assessment over real cellular networks: An experimental study based on complementary tools'. *2014 11th International Symposium on Wireless Communications Systems (ISWCS)*. Aug. 2014: pp. 786–790.

National conferences

1. Choque, J., L. Diez, and R. Agüero: 'Formulación del problema de selección de acceso en entornos heterogéneos y multi-servicio como un problema de programación lineal binaria'. *XI Jornadas de Ingeniería Telemática: JITEL*. Oct. 2013.

2. Igareda, J., L. Diez, V. Iancu, E. Slusanschi, and R. Agüero: 'Propuesta de un algoritmo de encaminamiento para distribuir la exposición a los campos electromagnéticos en redes inalámbricas malladas'. *XII Jornadas de Ingeniería Telemática: JITEL*. Oct. 2015.

3. Rodríguez, P., P. Sarasúa, L. Diez, and R. Agüero: 'Simulación genérica a nivel de sistema para soluciones avanzadas de gestión de recursos'. *XIII Jornadas de Ingeniería Telemática: JITEL*. Sept. 2017.

Others

1. Corre, Y. et al.: *Deliverable D6.2 Report on Validation, Part-A*. Tech. rep. LEXNET, 2015.

2. Corre, Y. et al.: *Deliverable D6.2 Report on Validation, Part-B*. Tech. rep. LEXNET, 2015.

3. Heliot, F. et al.: *D5.2 Smart low-EMF architectures: results and recommendations*. Tech. rep. LEXNET, 2015.

4. Popović, M. et al.: *D5.1 Smart low-EMF architectures: novel technologies overview*. Tech. rep. LEXNET, 2014.

5. Vermeeren, G et al.: *D3.3 Exposure Index assessment v2*. Tech. rep. LEXNET, 2014.

Summary of publications and their relationship with Thesis contributions. This table does not include deliverables from LEXNET project

	Publication	Sensing tool	Simulation framework - GWNSYM	Access selection & Resource management	EMF-aware routing & video delivery
International Journals	[Dom15]			✓	
	[Cho16]		✓		
	[Die16]			✓	
	[Die17a]	✓			
International Conferences	[Cho13b]		✓		
	[Agü13]		✓		
	[Die14]	✓			
	[Sar14]	✓			
	[Ian14]				✓
	[Rub14]			✓	
	[Die15c]				✓
	[Die15a]				✓
	[Die15f]	✓			
	[Die15d]			✓	
	[Die15e]				✓
	[Die15b]				✓
	[Die17b]		✓		
	[Bra16]			✓	
National Confs.	[Cho13a]			✓	
	[Iga15]				✓
	[Rod17]		✓		

Bibliography

[Abd15]   ABDELHADI, A. and C. CLANCY: 'An Optimal Resource Allocation with Joint Carrier Aggregation in 4G-LTE'. *Computing, Networking and Communications (ICNC), 2015 International Conference on*. Feb. 2015: pp. 138–142 (cit. on p. 95).

[Adh12]   ADHIKARI, V. K., Y. GUO, F. HAO, M. VARVELLO, V. HILT, M. STEINER, and Z. L. ZHANG: 'Unreeling netflix: Understanding and improving multi-CDN movie delivery'. *2012 Proceedings IEEE INFOCOM*. Mar. 2012: pp. 1620–1628 (cit. on p. 126).

[Agi16]   AGIWAL, M., A. ROY, and N. SAXENA: 'Next Generation 5G Wireless Networks: A Comprehensive Survey'. *IEEE Communications Surveys Tutorials* (2016), vol. PP(99): pp. 1–1 (cit. on p. 41).

[Agü13]   AGÜERO, R., J. CHOQUE, L. DIEZ, and L. MUÑOZ: 'Binary linear programming formulation of the access selection problem in multi-service and heterogeneous network environments'. *2013 IFIP Wireless Days (WD)*. Nov. 2013: pp. 1–4 (cit. on p. 192).

[Ahu93]   AHUJA, R. K., T. L. MAGNANTI, and J. B. ORLIN: *Network Flows: Theory, Algorithms, and Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993 (cit. on pp. 128, 129).

[Alt04]   ALTALHI, A. H. and G. G. RICHARD: 'Load-balanced routing through virtual paths: highly adaptive and efficient routing scheme for ad hoc wireless networks'. *Performance, Computing, and Communications, 2004 IEEE International Conference on*. 2004: pp. 407–413 (cit. on p. 126).

[And14]   ANDREWS, J. G., S. BUZZI, W. CHOI, S. V. HANLY, A. LOZANO, A. C. K. SOONG, and J. C. ZHANG: 'What Will 5G Be?' *IEEE Journal on Selected Areas in Communications* (June 2014), vol. 32(6): pp. 1065–1082 (cit. on p. 67).

[And07]   ANDREWS, M.: 'A Survey of Scheduling Theory in Wireless Data Networks'. *Wireless Communications*. Springer, 2007: pp. 1–17 (cit. on p. 94).

[Aue11]   AUER, G., V. GIANNINI, C. DESSET, I. GODOR, P. SKILLERMARK, M. OLSSON, M. A. IMRAN, D. SABELLA, M. J. GONZALEZ, O. BLUME, and A. FEHSKE: 'How much energy is needed to run a wireless network?' *Wireless Communications, IEEE* (Oct. 2011), vol. 18(5): pp. 40–49 (cit. on p. 126).

[Ban14]   BANGERTER, B., S. TALWAR, R. AREFI, and STEWART K.: 'Networks and devices for the 5G era'. *IEEE Communications Magazine* (Feb. 2014), vol. 52(2): pp. 90–96 (cit. on p. 67).

[Bar02]   BARRETT, C. L., M. V. MARATHE, D. C. ENGELHART, and A. SIVASUB-
           RAMANIAM: 'Analyzing the Short-term Fairness of IEEE 802.11 in Wireless
           Multi-hop Radio Networks'. *Proceedings. 10th IEEE International Symposium
           on Modeling, Analysis and Simulation of Computer and Telecommunications
           Systems.* 2002: pp. 137–144 (cit. on p. 116).

[Ber14]   BERGER, S., B. ALMEROTH, V. SURYAPRAKASH, P. ZANIER, I. VIERING, and
           G. FETTWEIS: 'Dynamic Range-Aware Uplink Transmit Power Control in LTE
           Networks: Establishing an Operational Range for LTE's Open-Loop Trans-
           mit Power Control Parameters ($alpha, P_0$)'. *IEEE Wireless Communications
           Letters* (Oct. 2014), vol. 3(5): pp. 521–524 (cit. on p. 78).

[Bhu14]   BHUSHAN, N., J. LI, D. MALLADI, R. GILMORE, D. BRENNER, A. DAMN-
           JANOVIC, R. SUKHAVASI, C. PATEL, and S. GEIRHOFER: 'Network densifica-
           tion: the dominant theme for wireless evolution into 5G'. *IEEE Communica-
           tions Magazine* (Feb. 2014), vol. 52(2): pp. 82–89 (cit. on pp. 41, 67).

[Boc16]   BOCCARDI, F., J. ANDREWS, H. ELSHAER, M. DOHLER, S. PARKVALL, P.
           POPOVSKI, and S. SINGH: 'Why to decouple the uplink and downlink in
           cellular networks and how to do it'. *IEEE Communications Magazine* (Mar.
           2016), vol. 54(3): pp. 110–117 (cit. on p. 90).

[Boc14]   BOCCARDI, F., R. W. HEATH, A. LOZANO, T. L. MARZETTA, and P.
           POPOVSKI: 'Five disruptive technology directions for 5G'. *IEEE Commu-
           nications Magazine* (Feb. 2014), vol. 52(2): pp. 74–80 (cit. on pp. 41, 68).

[Boc11]   BOCHE, H., S. NAIK, and T. ALPCAN: 'Characterization of Convex and Con-
           cave Resource Allocation Problems in Interference Coupled Wireless Systems'.
           Signal Processing, IEEE Transactions on (May 2011), vol. 59(5): pp. 2382–
           2394 (cit. on p. 69).

[Boy07]   BOYD, S., S-J. KIM, L. VANDENBERGHE, and A. HASSIBI: 'A tutorial on
           geometric programming'. *Optimization and Engineering* (2007), vol. 8(1):
           pp. 67–127 (cit. on p. 73).

[Boy04]   BOYD, S. and L. VANDENBERGHE: *Convex Optimization.* New York, NY, USA:
           Cambridge University Press, 2004 (cit. on pp. 73, 100, 101, 104, 107).

[Bra16]   BRAU, M., J. STÉPHAN, L. DIEZ, Y. CORRE, Y. LOSTANLEN, and R. AGÜERO:
           'Assessing the WiFi offloading benefit on both service performance and EMF
           exposure in urban areas'. *2016 Wireless Days (WD).* Mar. 2016: pp. 1–4
           (cit. on pp. 82, 192).

[Bur15]   BURNELL, E. and W. HOBURG: *GPkit software for geometric programming.*
           https://github.com/hoburg/gpkit. Version 0.4.0. 2015 (cit. on p. 76).

[Cab11]   CABAN, SEBASTIAN, CHRISTIAN MEHLFÜHRER, MARKUS RUPP, and MARTIN
           WRULICH: 'LTE Link- and System-Level Simulation'. *Evaluation of HSDPA
           and LTE.* John Wiley & Sons, Ltd, 2011: pp. 243–270 (cit. on p. 44).

[Cap13]   Capozzi, F., G. Piro, L. A. Grieco, G. Boggia, and P. Camarda:
           'Downlink Packet Scheduling in LTE Cellular Networks: Key Design Issues and
           a Survey'. *IEEE Communications Surveys Tutorials* (Feb. 2013), vol. 15(2):
           pp. 678–700 (cit. on p. 94).

[Cas07]   Castaneda, M., M. T. Ivrlac, J. A. Nossek, I. Viering, and A. Klein:
           'On Downlink Intercell Interference in a Cellular System'. *Personal, Indoor and
           Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International
           Symposium on.* Sept. 2007: pp. 1–5 (cit. on p. 68).

[Che17]   Chen, B., J. Chen, Y. Gao, and J. Zhang: 'Coexistence of LTE-LAA and
           Wi-Fi on 5 GHz With Corresponding Deployment Scenarios: A Survey'. *IEEE
           Communications Surveys Tutorials* (July 2017), vol. 19(1): pp. 7–32 (cit. on
           p. 68).

[Che13]   Cheng, X., G. Gupta, and P. Mohapatra: 'Joint Carrier Aggregation
           and Packet Scheduling in LTE-advanced Networks'. *2013 IEEE International
           Conference on Sensing, Communications and Networking (SECON).* June
           2013: pp. 469–477 (cit. on p. 108).

[Cho13a]   Choque, J., L. Diez, and R. Agüero: 'Formulación del problema de selección
           de acceso en entornos heterogéneos y multi-servicio como un problema de
           programación lineal binaria'. *XI Jornadas de Ingeniería Telemática: JITEL.*
           Oct. 2013 (cit. on p. 192).

[Cho13b]   Choque, J., L. Diez, R. Agüero, and L. Muñoz: 'Providing differentiated
           QoS with dynamic pricing policies for wireless access networks'. *2013 IFIP
           Wireless Days (WD).* Nov. 2013: pp. 1–4 (cit. on p. 192).

[Cho16]   Choque, J., L. Diez, A. E. García, R. Agüero, and L. Muñoz: 'Optimum
           Multi-service Access Selection over Heterogeneous Wireless Networks'. *Trans.
           Emerg. Telecommun. Technol.* (Mar. 2016), vol. 27(3): pp. 328–338 (cit. on
           pp. 43, 192).

[Chu13]   Chu, X., D. Lopez-Perez, Y. Yang, and F. Gunnarsson: *Heterogeneous
           Cellular Networks: Theory, Simulation and Deployment.* New York, NY, USA:
           Cambridge University Press, 2013 (cit. on p. 45).

[Cis17]   CISCO: *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast
           Update, 2016–2021 White Paper.* 2017. URL: https://www.cisco.com/c/
           en/us/solutions/collateral/service-provider/visual-networking-
           index-vni/mobile-white-paper-c11-520862.html (visited on 02/12/2018)
           (cit. on pp. 41, 67, 126).

[Col13]   Coll-Perales, B., J. Gozalvez, and J. Sanchez-Soriano: 'Empirical
           Performance Models for P2P and Two Hops Multi-hop Cellular Networks
           with Mobile Relays'. *Proceedings of the 8th ACM Workshop on Performance
           Monitoring and Measurement of Heterogeneous Wireless and Wired Networks.*
           PM2HW2N '13. Barcelona, Spain: ACM, 2013: pp. 21–28 (cit. on p. 83).

[Cor09]   CORRE, Y. and Y. LOSTANLEN: 'Three-Dimensional Urban EM Wave Prop-
            agation Model for Radio Network Planning and Optimization Over Large
            Areas'. *IEEE Transactions on Vehicular Technology* (Sept. 2009), vol. 58(7):
            pp. 3112–3123 (cit. on p. 17).

[Cor15]   CORRE, Y. et al.: *Deliverable D6.2 Report on Validation, Part-A*. Tech. rep.
            LEXNET, 2015 (cit. on pp. 16, 18, 24, 25, 37).

[Cor10]   CORREIA, L. M., D. ZELLER, O. BLUME, D. FERLING, Y. JADING, I. GODOR,
            G. AUER, and L. VAN DER PERRE: 'Challenges and enabling technologies for
            energy aware mobile radio networks'. *Communications Magazine, IEEE* (Nov.
            2010), vol. 48(11): pp. 66–72 (cit. on p. 126).

[Cou05]   COUTO, D. S. J. D., D. AGUAYO, J. BICKET, and R. MORRIS: 'A High-
            throughput Path Metric for Multi-hop Wireless Routing'. *Wirel. Netw.* (July
            2005), vol. 11(4): pp. 419–434 (cit. on p. 125).

[Dam11]   DAMNJANOVIC, A., J. MONTOJO, Y. WEI, T. JI, T. LUO, M. VAJAPEYAM,
            T. YOO, O. SONG, and D. MALLADI: 'A survey on 3GPP heterogeneous
            networks'. *IEEE Wireless Communications* (June 2011), vol. 18(3): pp. 10–21
            (cit. on pp. 68, 90).

[Dav15]   DAVIDOVIC, M., S. BJEKOVIC, and I. TOMIC: 'On the impact of network load
            on LTE network downlink performance'. *Proceedings of Abstracts and Program
            of the 2nd International Conference on Electrical, Electronic and Computing
            Engineering, IcETRAN 2015*. Nov. 2015: pp. 60–60 (cit. on p. 68).

[Dec13]   DECHENE, D. J. and A. SHAMI: 'Energy efficient QoS constrained scheduler
            for SC-FDMA uplink'. *Physical Communication* (2013), vol. 8: pp. 81–90
            (cit. on p. 95).

[Die17a]   DIEZ, L., R. AGÜERO, and L. MUÑOZ: 'Electromagnetic Field Assessment
            as a Smart City Service: The SmartSantander Use-Case'. *Sensors* (2017), vol.
            17(6) (cit. on pp. 7, 192).

[Die15a]   DIEZ, L., R. AGÜERO, and J. PENHOAT: 'Reducing the Electromagnetic
            Exposure over LTE Networks by Means of an Adaptive Retransmission Scheme:
            A Use Case Based on a Video Service'. *2015 IEEE 81st Vehicular Technology
            Conference (VTC Spring)*. May 2015: pp. 1–5 (cit. on pp. 152, 192).

[Die15b]   DIEZ, L., R. AGÜERO, J. PENHOAT, E. STÉPHAN, and F. HELIOT: 'Exploit-
            ing the DASH Framework to Promote a Dynamic Transport Selection for
            Video Streaming over LTE Networks'. *2015 IEEE International Conference on
            Computer and Information Technology; Ubiquitous Computing and Communi-
            cations; Dependable, Autonomic and Secure Computing; Pervasive Intelligence
            and Computing*. Oct. 2015: pp. 37–43 (cit. on p. 192).

[Die14]   DIEZ, L., S. M. ANWAR, L. RODRÍGUEZ DE LOPE, M. L. HENNAFF, Y.
            TOUTAIN, and R. AGÜERO: 'Design and integration of a low-complexity
            dosimeter into the smart city for EMF assessment'. *Networks and Communi-
            cations (EuCNC), 2014 European Conference on.* June 2014: pp. 1–5 (cit. on
            p. 192).

[Die15c]   DIEZ, L., J. ARROYO, and R. AGÜERO: 'On the use of emulation techniques
            over the ns-3 platform to analyze multimedia service QoS'. *LNICST.* Vol. 158.
            2015 (cit. on p. 192).

[Die15d]   DIEZ, L., Y. CORRE, B. RADIER, M. BRAU, and R. AGÜERO: 'Comple-
            menting radio-planning tools to analyze EMF-aware access selection'. *2015
            International Symposium on Wireless Communication Systems (ISWCS).* Aug.
            2015: pp. 431–435 (cit. on pp. 82, 192).

[Die15e]   DIEZ, L., J. IGAREDA, V. IANCU, E. SLUSANSCHI, and R. AGÜERO: 'Routing
            algorithm to fairly distribute the exposure to electromagnetic fields over
            wireless mesh networks'. *2015 IEEE 11th International Conference on Wireless
            and Mobile Computing, Networking and Communications (WiMob).* Oct. 2015:
            pp. 468–473 (cit. on p. 192).

[Die16]   DIEZ, L., G. P. POPESCU, and R. AGÜERO: 'A Geometric Programming
            Solution for the Mutual-Interference Model in HetNets'. *IEEE Communications
            Letters* (Sept. 2016), vol. 20(9): pp. 1876–1879 (cit. on p. 192).

[Die15f]   DIEZ, L., L. RODRÍGUEZ DE LOPE, R. AGÜERO, Y. CORRE, J. STÉPHAN, M.
            BRAU, S. AERTS, G. VERMEEREN, L. MARTENS, and W. JOSEPH: 'Optimal
            dosimeter deployment into a smart city IoT platform for wideband EMF ex-
            posure assessment'. *Networks and Communications (EuCNC), 2015 European
            Conference on.* June 2015: pp. 528–532 (cit. on pp. 18, 192).

[Die17b]   DIEZ, L., S. SERGIO IZÚEL, and R. AGÜERO: 'Generic Wireless Network Sys-
            tem Modeler: Fostering the Analysis of Complex LTE Deployments'. *LNICST.*
            Vol. 191. 2017 (cit. on p. 192).

[Dju16]   DJURIC, N., J. BJELICA, D. KLJAJIC, M. MILUTINOV, K. KASAS-LAZETIC,
            and D. ANTIC: 'The SEMONT continuous monitoring and exposure assessment
            for the low-frequency EMF'. *2016 IEEE International Conference on Emerging
            Technologies and Innovative Business Practices for the Transformation of
            Societies (EmergiTech).* Aug. 2016: pp. 50–55 (cit. on p. 8).

[Dog08]   DOGGEN, JEROEN and FILIP VAN DER SCHUEREN: 'Design and Simulation
            of a H.264 AVC Video Streaming Model'. *European Conference on the Use
            of Modern Information and Communication Technologies, ECUMICT* (Mar.
            2008), vol. (cit. on p. 160).

[Dom13]   DOMBROWSKI, C., N. PETRESKA, S. GÖRTZEN, A. SCHMEINK, and J. GROSS: 'Energy-efficient multi-hop transmission for machine-to-machine communications'. *2013 11th International Symposium and Workshops on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt).* May 2013: pp. 341–348 (cit. on p. 125).

[Dom15]   DOMENICO, A. de, L. DIEZ, R. AGÜERO, D. KTÉNAS, and V. SAVIN: 'EMF-Aware Cell Selection in Heterogeneous Cellular Networks'. *IEEE Communications Letters* (Feb. 2015), vol. 19(2): pp. 271–274 (cit. on p. 192).

[Dra04]   DRAVES, R., J. PADHYE, and B. ZILL: 'Routing in Multi-radio, Multi-hop Wireless Mesh Networks'. *Proceedings of the 10th Annual International Conference on Mobile Computing and Networking.* MobiCom '04. Philadelphia, PA, USA: ACM, 2004: pp. 114–128 (cit. on p. 125).

[Efh14]   EFHRAN: *: Report on the level of exposure (frequency patterns and modulation) in the European Union – Part 1: Radiofrequency (RF) radiation.* Tech. rep. (cit. on p. 12).

[Ess13]   ESSAILI, A. e., D. SCHROEDER, D. STAEHLE, M. SHEHADA, W. KELLERER, and E. STEINBACH: 'Quality-of-experience driven adaptive HTTP media delivery'. *Communications (ICC), 2013 IEEE International Conference on.* June 2013: pp. 2480–2485 (cit. on p. 151).

[Miw41]   EU FP7 MiWEBA, D4.1: *Radio Resource Management for mm-wave Overlay HetNets. D4.1: System Level Simulator Specification.* Tech. rep. Dec. 2014 (cit. on p. 109).

[Fal10]   FALOWO, O. E., S. ZEADALLY, and H. A. CHAN: 'Dynamic pricing for load-balancing in user-centric joint call admission control of next-generation wireless networks'. *International Journal of Communication Systems* (2010), vol. 23(3): pp. 335–368 (cit. on p. 43).

[Faz16]   FAZLIU, Z. L., C. F. CHIASSERINI, and G. M. DELL'AERA: 'Downlink Transmit Power Setting in LTE HetNets with Carrier Aggregation'. *2016 IEEE 17th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM).* June 2016: pp. 1–9 (cit. on p. 95).

[Fuj03]   FUJIMOTO, R. M., K. PERUMALLA, A. PARK, H. WU, M. H. AMMAR, and G. F. RILEY: 'Large-scale network simulation: how big? how fast?' *Modeling, Analysis and Simulation of Computer Telecommunications Systems, 2003. MASCOTS 2003. 11th IEEE/ACM International Symposium on.* Oct. 2003: pp. 116–123 (cit. on p. 44).

[Gat10]   GATI, A., E. CONIL, M. F. WONG, and J. WIART: 'Duality Between Uplink Local and Downlink Whole-Body Exposures in Operating Networks'. *IEEE Transactions on Electromagnetic Compatibility* (Nov. 2010), vol. 52(4): pp. 829–836 (cit. on pp. 7, 20).

[Gnu17]   GNU: *Linear Programming Kit.* Version 4.32. 2017 (cit. on pp. 58, 85).

[Got13]   GOTSIS, A. G., A. S. LIOUMPAS, and A. ALEXIOU: 'Analytical Modelling
           and Performance Evaluation of Realistic Time-controlled M2M Scheduling
           over LTE Cellular Networks'. *Transactions on Emerging Telecommunications
           Technologies* (2013), vol. 24(4): pp. 378–388 (cit. on p. 95).

[Gu09]   GU, C., Y. ZHANG, W. MA, N. LIU, and Y. MAN: 'Universal Modeling and
           Optimization for Multi-Radio Access Selection'. *Wireless Communications,
           Networking and Mobile Computing, 2009. WiCom '09. 5th International
           Conference on.* Sept. 2009: pp. 1–4 (cit. on p. 43).

[Gue07]   GUERIN, J., M. PORTMANN, and A. PIRZADA: 'Routing metrics for multi-
           radio wireless mesh networks'. *Telecommunication Networks and Applications
           Conference, 2007. ATNAC 2007. Australasian.* Dec. 2007: pp. 343–348 (cit. on
           p. 125).

[Has01]   HASSANEIN, H. and A. ZHOU: 'Routing with Load Balancing in Wireless
           Ad Hoc Networks'. *Proceedings of the 4th ACM International Workshop on
           Modeling, Analysis and Simulation of Wireless and Mobile Systems.* MSWIM
           '01. Rome, Italy, 2001: pp. 89–96 (cit. on p. 126).

[Hen03]   HENDERSON, D., S. H. JACOBSON, and A. W. JOHNSON: 'Handbook of Meta-
           heuristics'. Ed. by GLOVER, FRED and GARY A. KOCHENBERGER. Boston,
           MA: Springer US, 2003. Chap. The Theory and Practice of Simulated Anneal-
           ing: pp. 287–319 (cit. on p. 71).

[Her17]   HERMES: *Project for Systematic Measurements of the Electromagnetic Radi-
           ation.* 2017. URL: http://hermes.physics.auth.gr/en/elearn (visited on
           02/12/2018) (cit. on p. 8).

[Hir10]   HIRATA, A., O. FUJIWARA, T. NAGAOKA, and S. WATANABE: 'Estimation of
           Whole-Body Average SAR in Human Models Due to Plane-Wave Exposure at
           Resonance Frequency'. *IEEE Transactions on Electromagnetic Compatibility*
           (Feb. 2010), vol. 52(1): pp. 41–48 (cit. on p. 30).

[Hor10]   HORE, A. and D. ZIOU: 'Image Quality Metrics: PSNR vs. SSIM'. *Pat-
           tern Recognition (ICPR), 2010 20th International Conference on.* Aug. 2010:
           pp. 2366–2369 (cit. on p. 153).

[Hua13]   HUANG, J., F. QIAN, Y. GUO, Y. ZHOU, Q. XU, Z. M. MAO, S. SEN, and
           O. SPATSCHECK: 'An In-depth Study of LTE: Effect of Network Protocol and
           Application Behavior on Performance'. *SIGCOMM Comput. Commun. Rev.*
           (Aug. 2013), vol. 43(4): pp. 363–374 (cit. on p. 67).

[Ian14]   IANCU, V., L. DIEZ, L. RODRÍGUEZ DE LOPE, E. SLUSANSCHI, and R.
           AGÜERO: 'A reward-based routing protocol to reduce the EMF exposure over
           wireless mesh networks'. *2014 IFIP Wireless Days (WD).* Nov. 2014: pp. 1–4
           (cit. on p. 192).

[Iga15]   Igareda, J., L. Diez, V. Iancu, E. Slusanschi, and R. Agüero: 'Propuesta de un algoritmo de encaminamiento para distribuir la exposición a los campos electromagnéticos en redes inalámbricas malladas'. *XII Jornadas de Ingeniería Telemática: JITEL*. Oct. 2015 (cit. on p. 192).

[Iqb14]   Iqbal, F., M. Y. Javed, and M. J. Iqbal: 'Diversity based review of Multi-path Routing Metrics of Wireless Mesh Networks'. *Multi-Topic Conference (INMIC), 2014 IEEE 17th International*. Dec. 2014: pp. 320–325 (cit. on p. 125).

[Jai84]   Jain, R. K., D. W. Chiu, and W. R. Hawe: *A Quantitative Measure Of Fairness And Discrimination For Resource Allocation In Shared Computer Systems*. Tech. rep. Digital Equipment Corporation, Sept. 1984 (cit. on p. 116).

[Jia07]   Jiang, W., S. Liu, Y. Zhu, and Z. Zhang: 'Optimizing Routing Metrics for Large-Scale Multi-Radio Mesh Networks'. *2007 International Conference on Wireless Communications, Networking and Mobile Computing*. Sept. 2007: pp. 1550–1553 (cit. on p. 125).

[Kal15]   Kalil, M., A. Shami, and A. Al-Dweik: 'QoS-aware Power-efficient Scheduler for LTE Uplink'. *IEEE Transactions on Mobile Computing* (Aug. 2015), vol. 14(8): pp. 1672–1685 (cit. on p. 95).

[Kar09]   Karray, M. K.: 'Study of a key factor for performance evaluation of wireless cellular networks: The f-factor'. *Wireless Days (WD), 2009 2nd IFIP*. Dec. 2009: pp. 1–6 (cit. on p. 68).

[Kha10]   Khan, A., L. Sun, E. Ifeachor, O. J. Fajardo, F. Liberal, and H. Koumaras: 'Video Quality Prediction Models Based on Video Content Dynamics for H.264 Video over UMTS Networks'. *International Journal of Digital Multimedia Broadcasting* (2010), vol. 2010 (cit. on pp. 159, 162).

[Kha14]   Khan, S., D. Schroeder, A. e. Essaili, and E. Steinbach: 'Energy-efficient and QoE-driven adaptive HTTP streaming over LTE'. *Wireless Communications and Networking Conference (WCNC), 2014 IEEE*. Apr. 2014: pp. 2354–2359 (cit. on p. 151).

[Kra11]   Krasniqi, B. and C. F. Mecklenbrauker: 'Maximization of the Minimum Rate by Geometric Programming for Multiple Users in Partial Frequency Reuse Cellular Networks'. *Vehicular Technology Conference (VTC Fall), 2011 IEEE*. Sept. 2011: pp. 1–5 (cit. on p. 69).

[Ku15]   Ku, G. and J. M. Walsh: 'Resource Allocation and Link Adaptation in LTE and LTE Advanced: A Tutorial'. *IEEE Communications Surveys Tutorials* (Dec. 2015), vol. 17(3): pp. 1605–1633 (cit. on p. 94).

[Kwa09]   Kwan, R., C. Leung, and J. Zhang: 'Proportional Fair Multiuser Scheduling in LTE'. *IEEE Signal Processing Letters* (June 2009), vol. 16(6): pp. 461–464 (cit. on p. 94).

[Lee05]   LEE, J. W., R. R. MAZUMDAR, and N. B. SHROFF: 'Downlink Power Alloca-
           tion for Multi-class Wireless Systems'. *IEEE/ACM Trans. Netw.* (Aug. 2005),
           vol. 13(4): pp. 854–867 (cit. on p. 106).

[Li12]   LI, X., N. MITTON, A. NAYAK, and I. STOJMENOVIC: 'Localized load-aware
           geographic routing in wireless ad hoc networks'. *2012 IEEE International
           Conference on Communications (ICC)*. June 2012: pp. 7478–5482 (cit. on
           p. 126).

[Liu15]   LIU, J., N. KATO, J. MA, and N. KADOWAKI: 'Device-to-Device Communica-
           tion in LTE-Advanced Networks: A Survey'. *IEEE Communications Surveys
           Tutorials* (Dec. 2015), vol. 17(4): pp. 1923–1940 (cit. on p. 125).

[Liu14]   LIU, J., J. WANG, G. SHEN, and Y. YANG: 'IoT control based on uplink
           inter-cell power control in LTE system'. *2014 IEEE 25th Annual International
           Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC)*.
           Sept. 2014: pp. 717–721 (cit. on p. 78).

[Luc12]   LUCAS-ESTAÑ, M. C., J. GOZÁLVEZ, and J. SANCHEZ-SORIANO: 'Integer
           linear programming optimization of joint RRM policies for heterogeneous
           wireless systems'. *Computer Networks* (2012), vol. 56(1): pp. 112–126 (cit. on
           pp. 43, 52).

[Mar12]   MARQUES, A. G., L. M. LOPEZ-RAMOS, G. B. GIANNAKIS, J. RAMOS,
           and A. J. CAAMAÑO: 'Optimal Cross-Layer Resource Allocation in Cellular
           Networks Using Channel- and Queue-State Information'. *IEEE Transactions
           on Vehicular Technology* (July 2012), vol. 61(6): pp. 2789–2807 (cit. on p. 96).

[Mar16]   MARTIKAINENM, H., I. VIERING, and B. WEGMANN: 'Dynamic range aware
           LTE uplink P0 optimization in HetNet'. *2016 IEEE International Conference
           on Communications (ICC)*. May 2016: pp. 1–6 (cit. on p. 80).

[Mas01]   MASCOLO, S., C. CASETTI, M. GERLA, M.Y. SANADIDI, and R. WANG: 'TCP
           Westwood: Bandwidth Estimation for Enhanced Transport over Wireless Links'.
           *Proceedings of the 7th Annual International Conference on Mobile Computing
           and Networking*. MobiCom '01. Rome, Italy: ACM, 2001: pp. 287–297 (cit. on
           p. 153).

[Meh11]   MEHLFÜHRER, C., J. COLOM IKUNO, M. ŠIMKO, S. SCHWARZ, M. WRULICH,
           and M. RUPP: 'The Vienna LTE simulators - Enabling reproducibility in
           wireless communications research'. *EURASIP Journal on Advances in Signal
           Processing* (2011), vol. 2011(1): pp. 1–14 (cit. on pp. 42, 43).

[Mil12]   MILLER, K., E. QUACCHIO, G. GENNARI, and A. WOLISZ: 'Adaptation
           algorithm for adaptive streaming over HTTP'. *Packet Video Workshop (PV),
           2012 19th International*. May 2012: pp. 173–178 (cit. on p. 151).

[Mog07]   MOGENSEN, P., W. NA, I. Z. KOVACS, F. FREDERIKSEN, A. POKHARIYAL,
           K. I. PEDERSEN, T. KOLDING, K. HUGL, and M. KUUSELA: 'LTE Capacity
           Compared to the Shannon Bound'. *Vehicular Technology Conference, 2007.
           VTC2007-Spring. IEEE 65th*. Apr. 2007: pp. 1234–1238 (cit. on pp. 69, 82).

[Muh10]   MUHAMMAD, B. and A. MOHAMMED: 'Uplink closed loop power control for LTE system'. *2010 6th International Conference on Emerging Technologies (ICET)*. Oct. 2010: pp. 88–93 (cit. on p. 78).

[Nas16]   NASRI, R. and A. JAZIRI: 'On the Analytical Tractability of Hexagonal Network Model with Random User Location'. *Wireless Communications, IEEE Transactions on* (2016), vol. PP(99): pp. 1–1 (cit. on p. 68).

[Nas04]   NASSER, N. and H. HASSANEIN: 'Dynamic threshold-based call admission framework for prioritized multimedia traffic in wireless cellular networks'. *Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE*. Vol. 2. Nov. 2004: 644–649 Vol.2 (cit. on p. 52).

[Ned08]   NEDIC, A. and A. OZDAGLAR: 'Subgradient Methods in Network Resource Allocation: Rate Analysis'. *Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference on*. Mar. 2008: pp. 1189–1194 (cit. on p. 96).

[Nee13]   NEELY, M. J. and S. SUPITTAYAPORNPONG: 'Dynamic Markov Decision Policies for Delay Constrained Wireless Scheduling'. *IEEE Transactions on Automatic Control* (Aug. 2013), vol. 58(8): pp. 1948–1961 (cit. on p. 95).

[Ns317]   *Network Simulator 3*. 2017. URL: http://www.nsnam.org/ (visited on 02/12/2018) (cit. on pp. 42, 127, 143, 154, 159).

[Neu05]   NEUBAUER, G., M. RÖÖSLI, M. FEYCHTING, Y. HAMNERIUS, L. KHEIFETS, N. KUSTER, I. RUIZ, J. SCHÜZ, R. ÜBERBACHER, and J. WIART: *Study on the Feasibility of Epidemiological Studies on Health Effects of Mobile Telephone Base Stations*. Tech. rep. 2005 (cit. on p. 12).

[Niy12]   NIYATO, D., P. WANG, E. HOSSAIN, W. SAAD, and Z. HAN: 'Game theoretic modeling of cooperation among service providers in mobile cloud computing environments'. *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*. Apr. 2012: pp. 3128–3133 (cit. on p. 43).

[Oli07]   OLIVEIRA, C., D. SEBASTIÃO, G. CARPINTEIRO, L. M. CORREIA, C. A. FERNANDES, A. SERRALHA, and N. MARQUES: 'The moniT Project: Electromagnetic Radiation Exposure Assessment in Mobile Communications'. *IEEE Antennas and Propagation Magazine* (Feb. 2007), vol. 49(1): pp. 44–53 (cit. on p. 8).

[Pel11]   PELKEY, J. and G. RILEY: 'Distributed Simulation with MPI in Ns-3'. *Proceedings of the 4th International ICST Conference on Simulation Tools and Techniques*. SIMUTools '11. Barcelona, Spain: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2011: pp. 410–414 (cit. on p. 44).

[Pen15]   PENG, M., C. WANG, V. LAU, and H. V. POOR: 'Fronthaul-constrained cloud radio access networks: insights and challenges'. *IEEE Wireless Communications* (Apr. 2015), vol. 22(2): pp. 152–160 (cit. on p. 68).

[Per03]   PERKINS, C., E. BELDING-ROYER, and S. DAS: *Ad Hoc On-Demand Distance Vector (AODV) Routing*. United States, 2003 (cit. on p. 139).

[Pir11]   Piro, G., N. Baldo, and M. Miozzo: 'An LTE Module for the Ns-3 Network Simulator'. *Proceedings of the 4th International ICST Conference on Simulation Tools and Techniques*. SIMUTools '11. Barcelona, Spain: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2011: pp. 415–422 (cit. on pp. 42, 160).

[Pra14]   Prasad, N., H. Zhang, H. Zhu, and S. Rangarajan: 'Multiuser Scheduling in the 3GPP LTE Cellular Uplink'. *IEEE Transactions on Mobile Computing* (Jan. 2014), vol. 13(1): pp. 130–145 (cit. on p. 94).

[Qua16]   Qualcomm: 'Exploring 5G New Radio: Use Cases, Capabilities & Timeline'. (Sept. 2016), vol. (cit. on pp. 67, 93).

[Qua13]   Qualcomm: *The 1000x data challenge*. Nov. 2013. URL: %7Bhttps://www.qualcomm.com/invention/1000x%7D (cit. on p. 67).

[Rav16]   Ravanshid, A., P. Rost, D. S. Michalopoulos, V. V. Phan, H. Bakker, D. Aziz, S. Tayade, H. D. Schotten, S. Wong, and O. Holland: 'Multi-connectivity Functional Architectures in 5G'. *2016 IEEE International Conference on Communications Workshops (ICC)*. May 2016: pp. 187–192 (cit. on p. 67).

[Riz16]   Rizk, A. and M. Fidler: 'Queue-aware Uplink Scheduling with Stochastic Guarantees'. *Computer Communications* (2016), vol. 84: pp. 63–72 (cit. on p. 95).

[Rod17]   Rodríguez, P., P. Sarasúa, L. Diez, and R. Agüero: 'Simulación genérica a nivel de sistema para soluciones avanzadas de gestión de recursos'. *XIII Jornadas de Ingeniería Telemática: JITEL*. Sept. 2017 (cit. on p. 192).

[Rub14]   Rubio-Loyola, J., L. Gonzalez-Hernandez, L. Diez, R. Agüero, and J. Serrat: 'An energy-oriented optimization algorithm for solving the cell assignment problem in 4G-LTE communication networks'. *2014 IFIP Wireless Days (WD)*. Nov. 2014: pp. 1–4 (cit. on p. 192).

[San14]   Sanchez, L., L. Muñoz, J. A. Galache, P. Sotres, J. R. Santana, V. Gutiérrez, R. Ramdhany, A. Gluhak, S. Krco, E. Theodoridis, and D. Pfisterer: 'SmartSantander: IoT Experimentation over a Smart City Testbed'. *Comput. Netw.* (2014), vol. 61: pp. 217–238 (cit. on pp. 7, 10, 11).

[Sar14]   Sarrebourse, T., L. Rodríguez de Lope, A. Hadjem, L. Diez, S. M. Anwar, R. Agüero, Y. Toulain, and J. Wian: 'Towards EMF exposure assessment over real cellular networks: An experimental study based on complementary tools'. *2014 11th International Symposium on Wireless Communications Systems (ISWCS)*. Aug. 2014: pp. 786–790 (cit. on p. 192).

[Sch12]   Schneider, C. and R. S. Thomä: 'Evaluation of LTE link-level performance with closed loop spatial multiplexing in a realistic urban macro environment'. *Antennas and Propagation (EUCAP), 2012 6th European Conference on*. Mar. 2012: pp. 2725–2729 (cit. on p. 44).

[Sha15]   SHAJAIAH, H., A. ABDEL-HADI, and C. CLANCY: 'An Efficient Multi-carrier
          Resource Allocation with User Discrimination Framework for 5G Wireless
          Systems'. *International Journal of Wireless Information Networks* (2015), vol.
          22(4): pp. 345–356 (cit. on pp. 95, 106, 108).

[Sha13]   SHAJAIAH, H., A. ABDEL-HADI, and C. CLANCY: 'Utility Proportional Fair-
          ness Resource Allocation with Carrier Aggregation in 4G-LTE'. *MILCOM
          2013 - 2013 IEEE Military Communications Conference.* Nov. 2013: pp. 412–
          417 (cit. on p. 95).

[She12]   SHEN, Z., A. PAPASAKELLARIOU, J. MONTOJO, D. GERSTENBERGER, and F.
          XU: 'Overview of 3GPP LTE-advanced Carrier Aggregation for 4G Wireless
          Communications'. *IEEE Communications Magazine* (Feb. 2012), vol. 50(2):
          pp. 122–130 (cit. on p. 67).

[Sim13]   SIMONE, F. D. and F. DUFAUX: 'Comparison of DASH adaptation strategies
          based on bitrate and quality signalling'. *Multimedia Signal Processing (MMSP),
          2013 IEEE 15th International Workshop on.* Sept. 2013: pp. 087–092 (cit. on
          pp. 151, 153).

[Sim08]   SIMONSSON, A. and A. FURUSKAR: 'Uplink Power Control in LTE - Overview
          and Performance, Subtitle: Principles and Benefits of Utilizing rather than
          Compensating for SINR Variations'. *2008 IEEE 68th Vehicular Technology
          Conference.* Sept. 2008: pp. 1–5 (cit. on pp. 77, 80).

[Vol17]   SIRADEL: *Volcano Propagation Model.* URL: https://www.siradel.com/
          software/connectivity/volcano-software/ (visited on 02/12/2018) (cit.
          on pp. 17, 82).

[Sir12]   SIRIS, V. A., E. Z. TRAGOS, and N. E. PETROULAKIS: 'Experiences with
          a metropolitan multiradio wireless mesh network: design, performance, and
          application'. *IEEE Communications Magazine* (July 2012), vol. 50(7): pp. 128–
          136 (cit. on p. 8).

[Sod11]   SODAGAR, I.: 'The MPEG-DASH Standard for Multimedia Streaming Over
          the Internet'. *MultiMedia, IEEE* (Apr. 2011), vol. 18(4): pp. 62–67 (cit. on
          pp. 126, 149).

[Sub06]   SUBRAMANIAN, A. P., M. M. BUDDHIKOT, and S. MILLER: 'Interference
          aware routing in multi-radio wireless mesh networks'. *Wireless Mesh Networks,
          2006. WiMesh 2006. 2nd IEEE Workshop on.* Sept. 2006: pp. 55–63 (cit. on
          p. 125).

[Sum06]   SUMAN, B. and P. KUMAR: 'A survey of simulated annealing as a tool for
          single and multiobjective optimization'. *Journal of the Operational Research
          Society* (Oct. 2006), vol. 57(10): pp. 1143–1160 (cit. on p. 71).

[Tar15]   TARANETZ, M., T. BLAZEK, T. KROPFREITER, M. K. MÜLLER, S. SCHWARZ,
          and M. RUPP: 'Runtime Precoding: Enabling Multipoint Transmission in LTE-
          Advanced System-Level Simulations'. *IEEE Access* (2015), vol. 3: pp. 725–736
          (cit. on p. 44).

[TS36331]   *Technical Specification Group Radio Access Network; Evolved Universal Ter-restrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification.* TS 36.331. Version V14.4.0. 3rd Generation Partnership Project (3GPP), 2017 (cit. on pp. 93, 109).

[TR36814]   *Technical Specification Group Radio Access Network;Evolved Universal Ter-restrial Radio Access (E-UTRA);Further advancements for E-UTRA physical layer aspects.* TR 36.814. Version V9.2.0. 3rd Generation Partnership Project (3GPP), 2017 (cit. on p. 76).

[TS36213]   *Technical Specification Group Radio Access Network;Evolved Universal Ter-restrial Radio Access (E-UTRA);Physical layer procedures.* TS 36.213. Ver-sion V14.4.0. 3rd Generation Partnership Project (3GPP), 2017 (cit. on pp. 109, 114, 119).

[TS36322]   *Technical Specification Group Radio Access Network;Evolved Universal Terres-trial Radio Access (E-UTRA);Radio Link Control (RLC) protocol specification.* TS 36.322. Version V14.1.0. 3rd Generation Partnership Project (3GPP), 2017 (cit. on p. 159).

[TS25322]   *Technical Specification Group Radio Access Network;Radio Link Control (RLC) protocol specification.* TS 25.322. Version V14.0.0. 3rd Generation Partnership Project (3GPP), 2017 (cit. on p. 151).

[TS26244]   *Technical Specification Group Services and System Aspects; Transparent end-to-end packet switched streaming service (PSS);3GPP file format (3GP).* TS 26.244. Version V14.0.0. 3rd Generation Partnership Project (3GPP), 2017 (cit. on p. 126).

[TS26234]   *Technical Specification Group Services and System Aspects;Transparent end-to-end Packet-switched Streaming Service (PSS);Protocols and codecs.* TS 26.234. Version V14.0.0. 3rd Generation Partnership Project (3GPP), 2017 (cit. on p. 126).

[Tes14]   TESANOVIC, M., E. CONIL, A. DE DOMENICO, R. AGÜERO, F. FREUDEN-STEIN, L. M. CORREIA, S. BORIES, L. MARTENS, WIEDEMANN P. M., and J. WIART: 'The LEXNET Project: Wireless Networks and EMF: Paving the Way for Low-EMF Networks of the Future'. *IEEE Vehicular Technology Magazine* (June 2014), vol. 9(2): pp. 20–28 (cit. on p. 7).

[Toh09]   TOH, C. K., A. N. LE, and Y. Z. CHO: 'Load balanced routing protocols for ad hoc mobile wireless networks'. *IEEE Communications Magazine* (Aug. 2009), vol. 47(8): pp. 78–84 (cit. on p. 125).

[Tos12]   TOSEEF, U., Y. ZAKI, A. TIMM-GIEL, and C. GÖRG: 'Uplink QoS Aware Multi-homing in Integrated 3GPP and non-3GPP Future Networks '. *Pro-ceedings of the 4thd ICST International Conference on Mobile Networks & Management, MONAMI'12.* Sept. 2012 (cit. on p. 43).

[Val17]   VALLS, V. and D. J. LEITH: *A Convex Optimization Approach to Discrete Optimal Control.* Tech. rep. 2017 (cit. on pp. 96, 101–103, 123).

[Val16]   VALLS, V. and D. J. LEITH: 'Max-Weight Revisited: Sequences of Nonconvex Optimizations Solving Convex Optimizations'. *IEEE/ACM Transactions on Networking* (Oct. 2016), vol. 24(5): pp. 2676–2689 (cit. on p. 96).

[Var15a]   VARSIER, N., D. PLETS, Y. CORRE, G. VERMEEREN, W. JOSEPH, S. AERTS, L. MARTENS, and J. WIART: 'A novel method to assess human population exposure induced by a wireless cellular network'. *Bioelectromagnetics* (2015), vol. 36(6): pp. 451–463 (cit. on p. 15).

[Var15b]   VARSIER, N. et al.: *D2.8 Global Wireless Exposure Metric Definition*. Tech. rep. LEXNET, 2015 (cit. on pp. 16, 17, 29, 30, 32).

[Vie09]   VIEL, J.-F., E. CARDIS, M. MOISSONNIER, R. DE SEZE, and M. HOURS: 'Radiofrequency exposure in the French general population: Band, time, location and activity variability'. (Nov. 2009), vol. 8: pp. 1150–1154 (cit. on pp. 12, 159).

[Wan15]   WANG, H., F. XU, Y. LI, P. ZHANG, and D. JIN: 'Understanding Mobile Traffic Patterns of Large Scale Cellular Towers in Urban Environment'. *Proceedings of the 2015 ACM Conference on Internet Measurement Conference*. IMC '15. Tokyo, Japan: ACM, 2015: pp. 225–238 (cit. on p. 67).

[Wu15]   WU, J., Y. ZHANG, M. ZUKERMAN, and E. K. N. YUNG: 'Energy-Efficient Base-Stations Sleep-Mode Techniques in Green Cellular Networks: A Survey'. *IEEE Communications Surveys Tutorials* (Feb. 2015), vol. 17(2): pp. 803–826 (cit. on p. 67).

[Xue12]   XUE, P., P. GONG, J. H. PARK, D. PARK, and D. K. KIM: 'Radio Resource Management with Proportional Rate Constraint in the Heterogeneous Networks'. *Wireless Communications, IEEE Transactions on* (Mar. 2012), vol. 11(3): pp. 1066–1075 (cit. on p. 43).

[Yam12]   YAMAMOTO, R., T. MIYOSHI, and Y. TANAKA: 'Neighbour Traffic-Aware Load Balancing Method in Ad Hoc Networks'. *Intelligent Networking and Collaborative Systems (INCoS), 2012 4th International Conference on*. Sept. 2012: pp. 193–197 (cit. on p. 126).

[Yan05]   YANG, Y., J. WANG, and R. H. KRAVETS: 'Interference-aware Load Balancing for Multihop Wireless Networks'. (2005), vol. (UIUCDCS-R-2005-2526) (cit. on p. 125).

[You14]   YOUSSEF, M., M. IBRAHIM, M. ABDELATIF, L. CHEN, and A. V. VASILAKOS: 'Routing Metrics of Cognitive Radio Networks: A Survey'. *IEEE Communications Surveys Tutorials* (Jan. 2014), vol. 16(1): pp. 92–109 (cit. on p. 125).

[Yu15]   YU, G., Q. CHEN, R. YIN, H. ZHANG, and G. Y. LI: 'Joint Downlink and Uplink Resource Allocation for Energy-Efficient Carrier Aggregation'. *IEEE Transactions on Wireless Communications* (June 2015), vol. 14(6): pp. 3207–3218 (cit. on p. 95).