Quality of Service and MObility driven cognitive radio Systems

FP7-ICT-2009-4/248454

QoSMOS

D6.6

Spectrum management framework integration and implementation

Contractual Date of Delivery to the CEC: 31st December 2012

Actual Date of Delivery to the CEC: 24th December 2012

Editor: Keith Briggs (BT)

Internal Reviewer: Olasunkanmi Durowoju (UNIS)

External Reviewers: Shyamalie Thilakawardana (BBC), Ahmed Aldabbagh (Ofcom)

Authors: Keith Briggs (BT), Bernd Bochow (Fraunhofer), László Csurgai-Horváth (BME), Ramiro Samano Robles (IT), Pål Gronsund (TEL), Youngwook Ko, Miguel López-Benítez (UNIS), Ulrico Celentano (UOULU), Ingo Karla (Fraunhofer).

Workpackage:	WP6
Est. person months:	10PM
Security:	PU
Nature:	Public
Version:	Issue 1
Total number of pages:	106

Abstract:

This deliverable describes the functions and capabilities of the spectrum management framework used in the final implementation. The framework implemented is a subset of the spectrum manager framework defined in the deliverable D6.7, which is intended to be read in parallel with this report.

Keyword list: Cognitive spectrum management, Specification of cognitive functions, Opportunistic functions, Self-learning, CM-SM functions, interfaces and protocols.

Table of contents

List of figures

Executive Summary

This deliverable describes the functions and capabilities of the spectrum management framework as presented in the final implementation. The implemented framework is a large subset of the QoSMOS spectrum manager framework as defined in deliverable D6.7.

The role of the cognitive manager for spectrum management (CM-SM) is to manage spectrum portfolios and to provide them to the cognitive manager for resource management (CM-RM). In order to acquire the radio context, the CM-SM also interacts with the spectrum sensing (SS). Communication among remote functional blocks (except for the CM-RM/CM-SM interworking) happens through the adaptation layer (AL).

The global regulatory repository (GRGR) provides all the constraints and supporting information as required by the regulations. These, appropriately filtered and processed, are integrated by the CM-SM with radio context and the performance metrics received from the CM-RM to contribute to the common portfolio repository (CPFR). The information exchanged through the spectrum provider repository (SPRR) is exploited for spectrum trading and is also used for spectrum portfolio generation. Chapter 2 of this deliverable provides documentation on evaluation and capability on this CM-SM and CM-RM collaboration**.**

Chapter 3 is an evaluate spectrum repositories, including the Global Regulator Repository (GRG) interface and functions. In addition functional description and evaluation of the TV whitespace geo-location database is also presented in this chapter.

Chapter 4 assesses the capabilities of the common spectrum control (CSPC), with a detailed description of the specific portfolio optimization technique originated and developed in QoSMOS.

Chapter 5 defines the local spectrum control (LSPC) combines with a method for implementing the spectrum mal-usage detection in spectrum-management framework.

Finally Chapter 6 presents a detailed study of the spectrum analyser and selector (SSE).

Abbreviations

ACK	Acknowledge
AL	Adaptation Layer
AP	Access Point
BER	Bit Error Rate
BLER	BLock Error Rate
CCDF	Complementary Cumulative Distribution Function
CDF	Cumulative Distribution Function
CIR	Carrier-to-Interference Ratio
CM	Cognitive Manager
CM-RM	Cognitive Manager – Resource Manager
CN	Core Network
CNR	Carrier-to-Noise Ratio
CPFR	Common Portfolio Repository
CPOR	Common Policy Repository
CQI	Channel Quality Indicator
CSPC	Common Spectrum Control
DLC	Data Link Control
DTV	Digital TeleVision
DUR	Desired-to-Undesired power Ratio
FTP	File Transfer Protocol
GRGR	Global Regulator Repository
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
IEEE	The Institute of Electrical and Electronics Engineers
IPTV	Internet Protocol Television
IU	Incumbent User
LPFR	Local Portfolio Repository
LSPC	Local Spectrum Control
LTE	Long Term Evolution
MAC	Medium Access Control
MSC	Message Sequence Chart
MT	Mobile Terminal
NC	Network Cognition
OU	Opportunistic User
PDF	Probability Density Function
PHY	Physical Layer

PMF	Probability Mass Function
PMSE	Program Making and Special Events
QAM	Quadrature-Amplitude Modulation
QoS	Quality of Service
RAN	Radio Access Network
RAT	Radio Access Technology
RB	Resource Block
RM	Resource Manager
RRM	Radio Resource Management
SAN	Spectrum Analyser
SINR	Signal-to-Interference plus Noise Ratio
SLA	Service Level Agreement
SM	Spectrum Manager
SNR	Signal-to-Noise Ratio
SPI	Spectrum Efficiency Index
SPRR	Spectrum Provider Repository
SS	Spectrum Sensing
SSE	Spectrum Selector
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
TTI	Transmission Time Interval
TVWS	TV White Space
UE	User Equipment
VoIP	Voice over IP
WAP	Wireless Application Protocol
WLAN	Wireless Local Area Network

1 CM-SM Reference Model

1.1 CM-SM Reference Model

In a nutshell, the role of the CM-SM (cognitive manager for spectrum management) [D6.2] is to manage (generate, deploy, maintain and revoke) spectrum portfolios and to provide them to the CM-RM (cognitive manager for resource management) [CelEtal2011]. In order to acquire the radio context, the CM-SM also interacts with the spectrum sensing (SS). Communication among remote functional blocks (except for the CM-RM/CM-SM interworking) ensured via the adaptation layer (AL).

The global regulatory repository (GRGR) provides all the constraints and supporting information as required by the regulations. These, appropriately filtered and processed, are integrated by the CM-SM with radio context and the performance metrics (received from the CM-RM) to contribute to the common portfolio repository (CPFR). The information exchanged through the spectrum provider repository (SPRR) is exploited for spectrum trading and is also used for spectrum portfolio generation.

The QoSMOS architecture is designed to be flexibly adaptable to diverse scenarios [D2.4]. In order to ease the design reuse, the internal functional blocks are assigned to topological domains [CelEtal2011] instead of rigidly associate them with network nodes. See Figure 1 and Figure 2 for a diagram of this architecture.

The functions, which are described briefly below, are divided into three groups. The repository access (CM-SM COORD) resides at the coordination domain (e.g., the core network in the cellular case) and accesses the repositories at the coexistence domain (e.g., the Internet). The local spectrum management (CM-SM NET) and the spectrum selection (CM-SM END) are placed at the network domain (e.g., at the base station in the cellular example). The CM-SM NET is in charge of eliminating information not relevant locally and adding instead what is important at the networking domain, whereas the CM-SM END works in close interaction with the CM-RM.

The common spectrum control (CSPC) is in charge of accessing the repositories it manages and therefore of merging/splitting the relevant information. The common policy repository (CPOR) controls all spectrum usage constrains not covered at regulatory or spectrum provider side and that are needed at the radio access network in question. The CSPC also interfaces, through the AL, possible remote CM-SM's (or equivalent and compatible entities) in neighbouring networks.

The local spectrum control (LSPC) manages the local spectrum and is therefore accessing the local portfolio repository (LPFR), which, in short, caches the spectrum portfolio information and may trigger actions at the spectrum selector (SSE). Measurements from the CM-RM are gathered by the spectrum analyser at CM-SM END and used at LPFR for local decision and sent further up to the CM-SM COORD through the LSPC for exploitation at a broader level (coordination and coexistence domain). The LSPC has the duty of all (networking-coordination domain) interactions towards the CM-SM COORD.

1.2 CM-SM state machine description

Figure 1 shows the QoSMOS reference model, fully detailed in previous WP6 deliverables. Needless to say, in an actual implementation, different parts of the system would be managed by different organizations and commercial entities, such as MNOs, regulator(s), WiFi providers (public and private), etc. But it is outside the scope of this deliverable to specify these organizations, since this belongs to the field of business modelling which is covered in WP1 deliverables. The same comment applies to Figure 9 later.

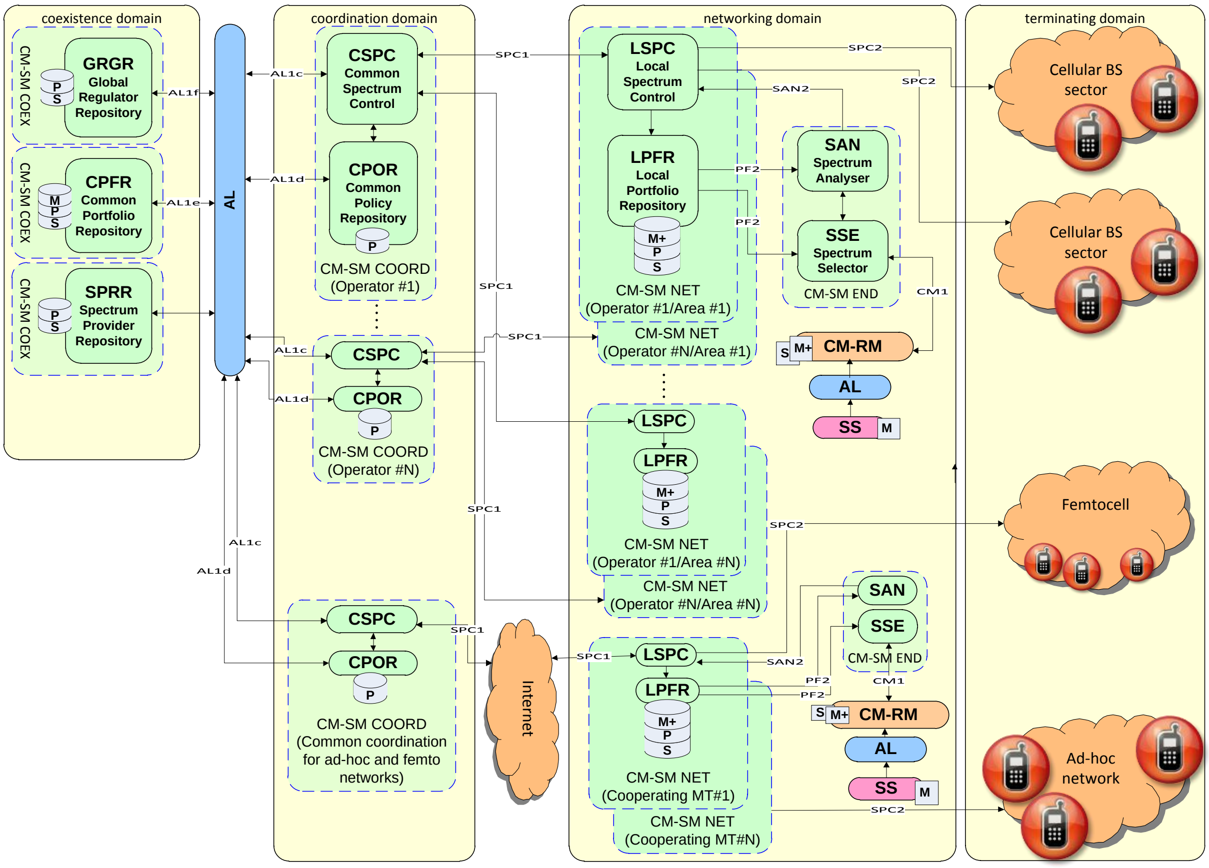


Figure 1 QoSMOS reference model of the CM-SM

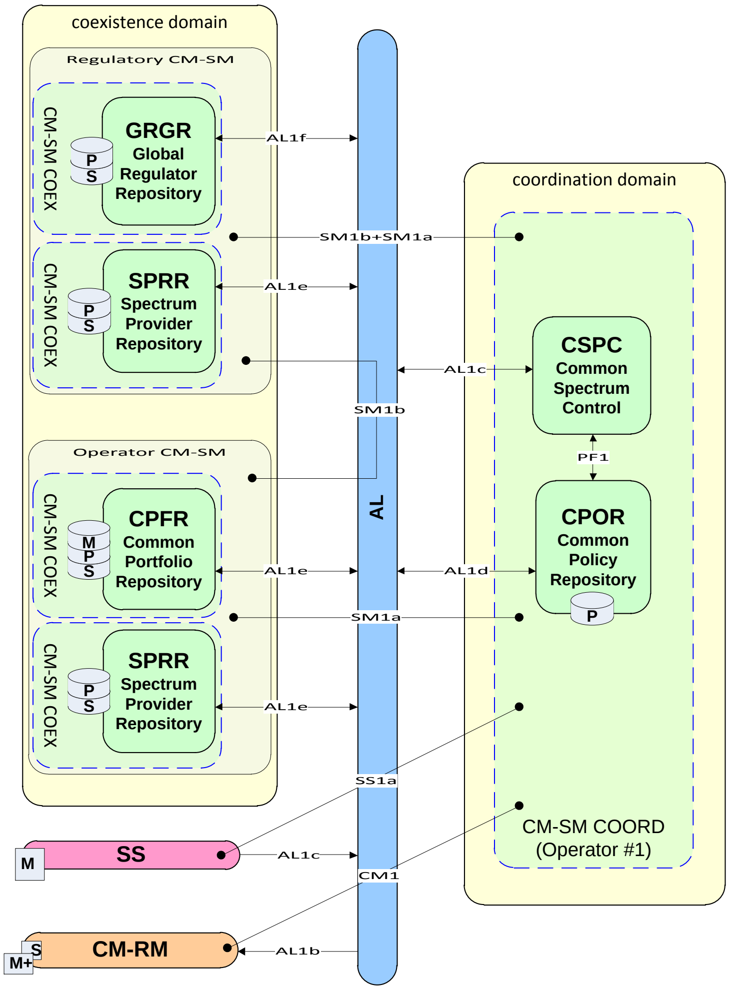


Figure 2 Interfaces between coexistence and coordination domain entities over adaptation layer

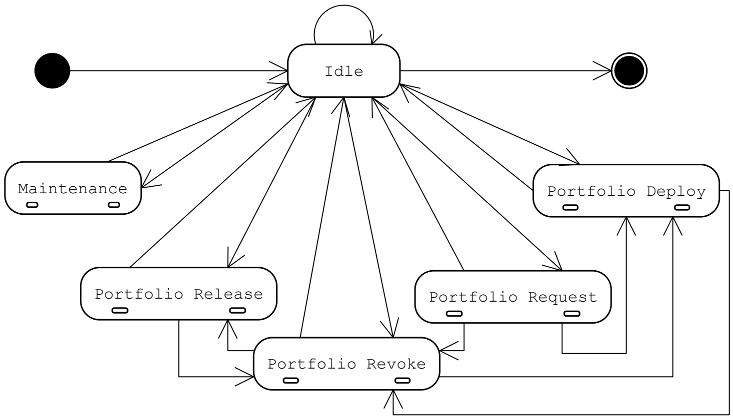
A CM-SM entity consists of multiple concurrent threads implementing state machines each responsible for handling a number of clients in parallel. The discussion of internal states of a CM-SM following refers to the sum of those. A particular thread may not implement all of details but only those of relevance for its current task and role.

Throughout the discussion of CM-SM states the following roles are assumed (see [D6.5] for naming details):

- `SM` – the spectrum manager. This role implements the full functionality of the CM-SM. It can deploy and revoke spectrum portfolios, and it can request a spectrum portfolio from a spectrum provider such as the `SP` or `GDB`. The closest realization of an `SM` role is in the CM-SM of the networking or coordination domains.
- `SP` – a spectrum provider. This role implements the capacity to deploy and revoke spectrum portfolio. In contrast to the `SM` it has no `RM` clients but only `SM` clients. The closest realization of an `SP` role is in the CM-SM of the coordination or coexistence domains. The communication between `SM` and `SP` is realized by the SM1a interface
- `GDB` – a Geolocation database. This role implements the capacity to respond to a portfolio request by an `SM`. It provides spectrum portfolios representing underutilized spectrum in a certain geographical area. The closest realization of a `GDB` role is in the Regulatory CM-SM of the coexistence domain encapsulating an authority's database. The communication between

GDB and SM is realized by the SM1a/b interfaces. Note that a Geolocation database client may obtain a time-limited lease for a portfolio, which implies a portfolio revocation when the lease expires. In that case, the Geolocation database is better represented by the SP role.

- RM – a resource manager. This role implements the capacity to request and release portfolios. It may provide portfolios generated from local sensing. Note that this not an SP role since the SM involved here is not considering that spectrum as a resource provided but as context information. The closest realization of an RM is in the CM-RM of the networking domain. The communication between RM and SM is realized by the CM1, SPC1 and SPC2 interfaces.
- SS – a spectrum sensing sub-system. This role implements the capacity to provide spectrum portfolios generated from local sensing. The closest realization of an SS is in the SS entities of the networking domain. The communication between SS and SM is realized by the CM1 and SS1b interfaces and often involves the RM as a proxy.



State	Valid for Role
Maintenance	SM,SP,GDB,RM,SS
Portfolio Release	SM,SP,GDB
Portfolio Revoke	SM,RM
Portfolio Request	SM,SP,GDB,SS
Portfolio Deploy	SM,RM

Figure 3 CM-SM top-level state diagram

Figure 1 depicts the top-level state diagram of the CM-SM. All states or sub-states shown can be realized as distinct concurrent processes or threads in an implementation. State transitions thus coincide with threat creation or termination. In the following CM-SM states are detailed further.

Idle state – Upon start-up a CM-SM is entering the Idle state. In this state it responds to requests from clients or it reacts on timed events. In consequence it enters one of the Maintenance, Portfolio Release, Portfolio Revoke, Portfolio Request or Portfolio Deploy states. It may remain in the Idle state in case no pending actions from earlier requests remain or in case of certain error conditions. Upon shut-down of the CM-SM, the Maintenance state is first reached for issuing requests to revoke or release portfolios earlier deployed or obtained, respectively, prior to the finally terminating the CM-SM.

Maintenance state – The Maintenance state is entered upon request of an internal timer event, or during start-up and shutdown procedures. In the Maintenance state the CM-SM is performing sanity checks on its internal threads and data structures, is checking current status and authorization of its clients and is executing internal performance monitoring tasks. A major task credited to the Maintenance state is in detecting dead or disconnected clients and to initiate the release of spectrum portfolios deployed earlier to such client (potentially revoking portfolios from clients if possible). The Maintenance state is left before the next internal timer event. Tasks initiated while in Maintenance state are processed in parallel (by a dedicated thread) or by re-entering the Idle state and then another suitable state depending on the task scheduled.

Portfolio Release state – The Portfolio Release state is entered when a CM-SM (in its SM role) receives a request to release a spectrum portfolio from local use or from client use (i.e. RM or GDB) thus freeing up resources associated with this portfolio. This may involve communication with another CM-SM (in its SP role) when the SM is requesting to release portfolios in turn. The Portfolio Release state is left after freeing up the portfolios referenced and confirming to the RM or GDB if required. Then Idle state is reached. The Portfolio Release state is entered also in consequence of a task initiated in Maintenance state, potentially subsequent to leaving the Portfolio Revoke state, which in general happens during regular CM-SM shutdown.

Portfolio Revoke state – The Portfolio Revoke state is entered by an SM in consequence of the request of a CM-SM in its SP role, or by an RM in consequence of the request of a CM-SM in its SM role. The Portfolio Revoke state may be included as a sub-state in several procedures such as programmed portfolio updates (initiated by the SM or by the RM). The Portfolio Revoke state is left after freeing up the portfolios referenced and confirming to the SM or to the SP. The Portfolio Revoke state also may be entered as a sub-state of the Portfolio Release state in case a programmed decision of the CM-SM demands to free up certain spectral resources during regular operations.

Portfolio Request state – The Portfolio Request state is entered by an SP upon request of an SM, by an SM upon request of an RM, or by a GDB upon request of an SM. The Portfolio Request state is left after deploying the requested portfolio to the requesting client, which includes issuing a request to the client (that in turn reaches the Portfolio Deploy state) before re-entering the Idle state. The Portfolio Request state in general is not a sub-state to any other state since it is initiated only by a protocol-driven request. The Idle state is reached from the Portfolio Request state immediately only if an error occurs and no portfolio can be deployed.

Portfolio Deploy state – The Portfolio Deploy state is entered by an SM or by an RM in consequence of receiving a portfolio from an SP, SM, SS, or GDB respectively in in consequence of an earlier request or revoke procedure. The Portfolio Deploy state may be reached also upon a local programmed decision when populating a local spectrum portfolio repository. The Portfolio Deploy state thus mostly appears as a sub-state of other states. The Portfolio Deploy state is left after issuing a confirmation to the portfolio provider. The Idle state is reached from a Portfolio Deploy state only in the most basic procedures, usually used during initialization of a client when the client is populated with a default set of portfolios.

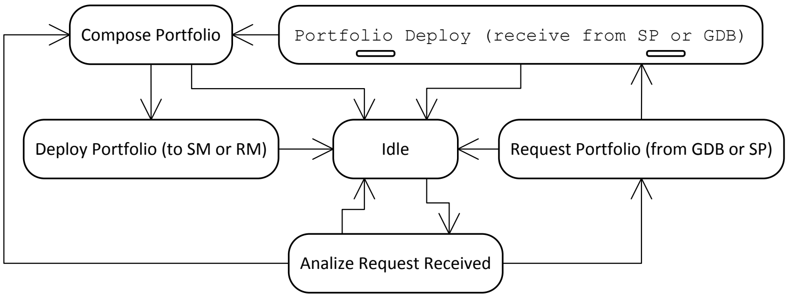


Figure 4 Portfolio Request state diagram

The Portfolio Request state is reached from the Idle state upon reception of a client request for a spectrum portfolio. A first check is done on the request, to identify if this request can be satisfied from the local spectrum portfolio repository, or if additional spectrum must be acquired to compose a portfolio for deployment towards the requesting client. This is outlined by Figure 4. The Portfolio Deploy state is entered in order to wait for a portfolio request response and for integrating a portfolio received into the local portfolio repository.

The Idle state is reached upon all error conditions (i.e. request failures, portfolio compose failures or upon a portfolio request failure response).

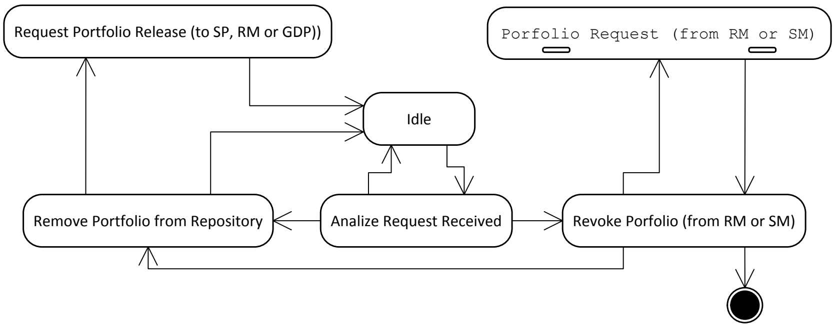


Figure 5 Portfolio Release state diagram

Figure 5 further details the Portfolio Release state. It is entered upon receiving a request from a client that wants to free up spectrum resources allocated earlier. If portfolios to be released are only of local relevance, the request can be satisfied by removing the portfolios from the local repository of the CM-SM. Local relevance is given if a portfolio has not been obtained earlier from an SP, SM or GDB or has not been deployed towards an SM or RM (or is part of a deployed portfolio).

If a portfolio has been obtained earlier, a release request may be issued towards the portfolio provider, or the portfolio may be retained for further use until revoked by the provider.

If a portfolio has been deployed earlier (or the portfolio released is part of a portfolio deployed earlier), a revocation request needs to be issued towards the client. Such revocation request may involve a portfolio update procedure, replacing the portfolio revoked by a different portfolio prior to removing to ensure seamless spectrum handover, or it may cause a portfolio request by the client detecting that it has insufficient spectrum resources to continue operation.

Note that under certain conditions a revocation may fail leaving the CM-SM in a deadlock where it cannot release spectrum any more. In such cases the only option may be to terminate CM-SM operation. In practice, the CM-SM then enters Maintenance state.

The Idle state is reached upon all error conditions except that it must not be reached through sub-states Portfolio Deploy and Portfolio Request since any error handling must be included with the previous sub-state attained before entering these states (see the discussion of the Portfolio Request state above).

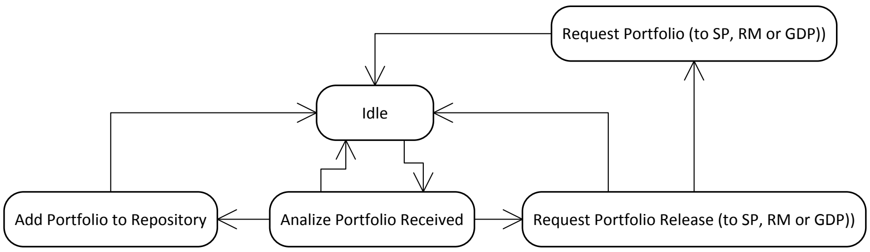


Figure 6 Portfolio Deploy state diagram

The Portfolio Deploy state is outlined by Figure 6. It is reached from the Idle state upon reception of a deployable portfolio from an SP, SM, or GDB possibly in response to an earlier portfolio request. In general, the portfolio received is added to a local repository and subsequently utilized as a spectrum resource.

It may happen that the portfolio received does not match the requested portfolio characteristics (e.g. if the portfolio provider was short on spectrum resources) or does not satisfy the current demand (e.g.

because the demand has changed between issuing a request and receiving the portfolio). In that case the portfolio may be released immediately and the request is re-issued (potentially with parameters having changed), or an additional complementing portfolio may be requested. The `Idle` state is reached next including upon all error conditions.

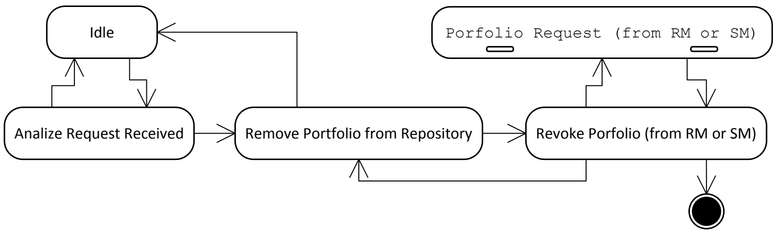


Figure 7 Portfolio Revoke state diagram

The `Portfolio Revoke` state as detailed by Figure 7 is reached upon reception of a portfolio revocation request from an `SM` or `SP`. Since portfolios up to be removed may have been deployed earlier to clients of this CM-SM, a subsequent portfolio revocation request may be needed in turn, potentially causing subsequent portfolio requests from those clients (see the discussion of the `Portfolio Release` state above) before reaching the `Idle` state again. The `Idle` state is reached upon all error conditions except that it must not be reached through sub-states `Portfolio Deploy` and `Portfolio Request` since any error handling must be included with the previous sub-state attained before entering these states (see the discussion of the `Portfolio Request` state above).

Note that under certain conditions a revocation may fail leaving the CM-SM in a deadlock where it cannot release spectrum any more. In such cases the only option may be to terminate CM-SM operation. In practice, the CM-SM then enters `Maintenance` state.

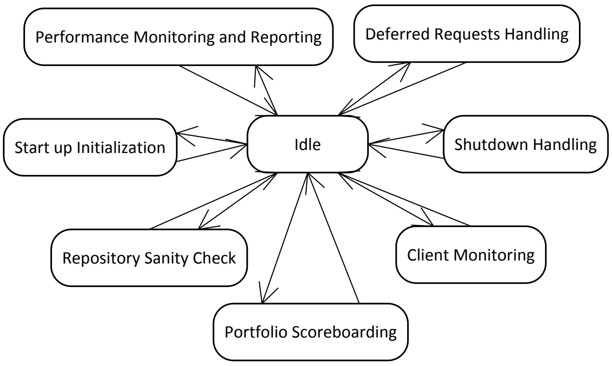


Figure 8 Maintenance state diagram

Details of the `Maintenance` state as given by Figure 8 have been already discussed in the scope of other CM-SM states. Most of the procedures executed in that state are considered management and administration processes that increase robustness and ensure resilience of the CM-SM. These are of no great relevance for the discussions made throughout this document and may change in future realizations.

In addition, the `Maintenance` state provides some performance optimization procedures. As an example Figure 8 references a state denoted as `Portfolio Scoreboarding`. Under this term a number of methods is summarized that evaluate the utilization of spectrum portfolios in the local repository according to certain performance metrics such as “use count”, “area coverage”, “utilization history”, “interference vulnerability” and similar. Those metrics are continuously updated and are utilized to increase the access time to the portfolio repository when selecting one out of many suitable portfolios for deployment in response to a client request.

2 Evaluation and capacity of CM-SM and CM-RM collaboration

2.1 Introduction

The QoSMOS reference model and its functional blocks were presented and described in Chapter 3 of [D2.2] and refined in [D2.3]. At the core of the reference model lie the two Cognitive Managers (CMs): the Cognitive Manager for Resource Management (CM-RM), which is developed in WP5, and the Cognitive Manager for Spectrum Management (CM-SM), which is developed in WP6. Both CMs are decision-making entities operating at different time scales and frequency granularities and need to communicate in order to coordinate their decision processes. The interworking between these two entities is therefore of particular importance and a clarification of the associated interactions is helpful. Previous deliverables from WP2, WP5 and WP6 have partially addressed several aspects of the interactions between the CM-SM and the CM-RM functional blocks from certain, partial points of view: from the system architecture viewpoint in WP2, from the CM-RM viewpoint in WP5 and from the CM-SM viewpoint in WP6. A previous study reported in [D6.5] provided a unified high-level clarification of interactions and the information exchanged, which was further elaborated in [D5.3]. In this context, this section provides a more detailed description, gathering and structuring some of the previous existing materials and providing a detailed and precise description of the functional blocks and their interface as well as the procedures involving interactions and the corresponding messages and exchanged data (see [D6.7] for the specification of the interface). This chapter provides a detailed presentation of the CM-RM/CM-SM interworking, with a twofold purpose. On the one hand, this section clarifies and helps to understand the diverse aspects of the interactions between the QoSMOS CMs, which were presented in a scattered manner across previous deliverables. On the other hand, the description presented in this chapter and the accompanying examples illustrate the capacity of the CM-RM and CM-SM collaborations and the benefits derived therefrom.

Section 2.2 describes the functionalities of the CM-RM and the CM-SM to provide a clear distinction of the roles played by each CM. Section 2.3 then presents a categorisation of the basic procedures requiring interworking of the CM-RM and CM-SM entities. Afterwards, Section 2.4 provides a discussion of selected examples requiring interworking and collaboration between the CM-RM and CM-SM entities in some practical situations in QoSMOS scenarios, which illustrates how some of these procedures are put into practice. A formal description of the CM1 interface between the CMs, including the Message Sequence Charts (MSCs) and involved primitives for the procedures highlighted in Section 2.3 can be found in [D6.7].

2.2 Functional description of the cognitive managers

This section describes the functionalities of the CM-RM and the CM-SM in order to provide a clear distinction of the roles and responsibilities of each functional block.

The CM-SM is responsible for the management of the frequency spectrum allocated to the QoSMOS entities for dedicated use. To this end, the CM-SM first acquires the relevant context information, such as environment information obtained from spectrum sensing results provided by the Spectrum Sensing (SS) block and performance reports of current spectrum usage provided by the CM-RM. Based on this information and on external constraints such as regulatory policies, operator policies and operator frequency planning, the CM-SM then builds up the spectrum portfolio. The spectrum portfolio condenses all the above information for further use by other QoSMOS entities, containing spectrum usage information and spectrum usage policies and putting constraints on the decisions that can be taken by other entities of the QoSMOS system. A spectrum portfolio constitutes a data structure describing arbitrary frequency bands and related information that summarises all relevant information on contiguous or non-contiguous chunks of frequency spectrum. A more detailed discussion of spectrum portfolios, including a formal definition of the concept, properties, operations and examples, can be found in Sections 5.1.4 and 7.1.3 of [D6.1].

The functionalities of the CM-SM are studied in detail in WP6 and include [D2.2]:

1. Management of portfolio sensing (i.e., sensing performed in portfolio channels[1]).
2. Exploitation of spectrum usage and performance reports provided by the CM-RM.
3. Management of spectrum portfolios (definition of the portfolio content, portfolio selection and splitting/merging of spectrum portfolios).
4. Provide spectrum portfolio to the CM-RM (deployment and revocation of portfolios).

The CM-RM is responsible for providing service to the application layer according to an agreed level of Quality of Service (QoS). To this end, the CM-RM allocates radio resources to the end users by making use of the spectrum opportunities and rules indicated by the spectrum portfolio. The CM-RM is therefore the main user of the spectrum portfolio generated by the CM-SM. The CM-RM is also responsible for the management of the user mobility as well as the implementation of functionalities needed to protect the incumbent users, relying in a close cooperation with the CM-SM, which in turn implements incumbent protection on a spectrum management level.

The functionalities of the CM-RM are studied in detail in WP5 and include [D2.2], [D5.2]:

1. Allocation of radio resources to the end users based on the opportunities and rules indicated by the spectrum portfolio and taking into account QoS requirements, in order to provide service to the upper layers (application layer).
2. Management of user mobility.
3. Ensure protection of incumbent users by managing the sensing of operating and reserve channels and using the spectrum sensing information along with the received portfolios to select appropriate node configurations that avoid harmful interference.
4. Configuration of the physical (PHY) and Medium Access Control (MAC) layers
5. Provide spectrum usage and performance reports to the CM-SM.

Both CM-RM and CM-SM are decision-making entities that select and allocate spectrum to spectrum user entities. However, there are important differences. The CM-RM and the CM-SM differ from each other in their timescale of operation and the frequency granularity. In terms of time scale, the CM-RM is responsible for short term decisions (down to the scheduling of data transmission at the slot level) while the CM-SM targets longer-term decisions. The CM-RM therefore operates on short time-scales in the order of milliseconds, typically in the range from a few to up to 100 ms approximately. As a result, spectrum usage and performance reporting procedures occur over relatively shorter time-scales compared to other interworking procedures. On the other hand, the CM-SM operates with knowledge obtained from larger amounts of spectrum, involving several CM-RMs, and as a dynamic spectrum management system it operates over much longer time-scales that could typically lie in the range from a few seconds to several hours (a more detailed discussion on the time-scales of operation is provided in Section 3.1 of [D6.5]). Therefore, interworking procedures for spectrum portfolio and spectrum

[1] An *operating channel* is a channel in active use at a given opportunistic communication device. A *reserve channel* is a channel that may turn into operating state in a short period, upon unavailability of the current operating channel (e.g., appearance of an incumbent). A *portfolio channel* is a channel an opportunistic device may provide sensing information about. Portfolio channels are for example those a device may be interested in for identification of alternative reserve channels (e.g., they are interesting due to their radio propagation or usage statistical properties but they are currently in use by an incumbent) or they may also be channels a device might not have any interest about. The portfolio is defined as the superset containing operating channels, reserve channels as well as inactive portfolio channels. More details on the categorization of the channels for sensing can be found in Section 9.1 of [D2.2].

policy management occur over longer time-scales. In terms of spectrum granularity, the CM-SM handles spectrum bands or blocks of spectrum (portfolios) that are allocated to CM-RM entities while the CM-RM allocates individual channels from the received spectrum portfolio to individual users. Therefore, the CM-RM is aware of the particular needs of every individual user while the CM-SM aggregates and represents in an abstract way the needs of all the users of the system. These complimentary time-scales of operation and frequency granularities add flexibility and efficiency to the management, selection and allocation of spectral and radio resources to the QoSMOS entities.

Another important difference is the network elements where the functional blocks are implemented. The mapping of a functional block (i.e., each CM) to a network element depends on the considered scenario. For example, in a cellular scenario, the CM-RM is mostly terminal-centric, with some component in the network, while the CM-SM is mostly network-centric, distributed in the Radio Access Network (RAN) and Core Network (CN), but may have an extension in the Mobile Terminal (MT) as well. The distribution and mapping of the functional blocks of the QoSMOS reference model into network nodes for the QoSMOS scenarios is out of the scope of this section. A detailed discussion can be found in Chapter 4 of [D2.2] as well as in chapter 3 of [D2.3] for different topologies' combinations and architecture options.

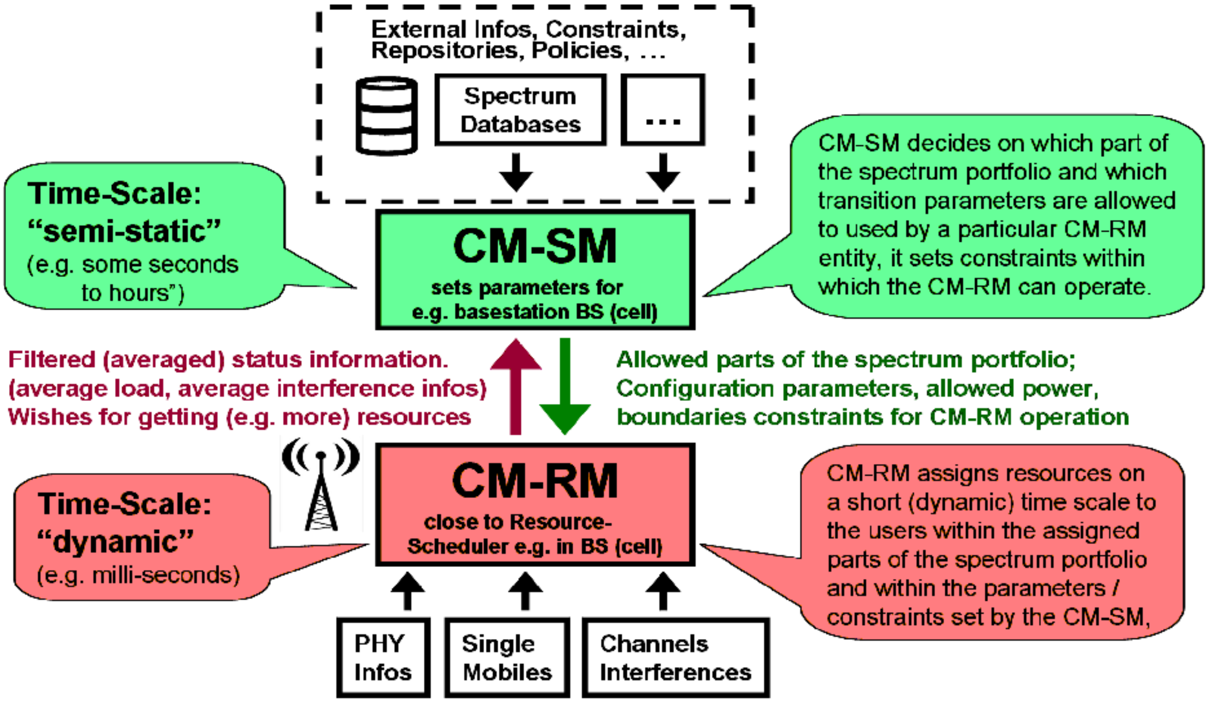


Figure 9 Overview of CM-SM and CM-RM tasks, functions and responsibilities [D6.5]

Figure 9 illustrates schematically the interworking of the CM-SM and CM-RM as well as their differences in tasks, responsibilities, time scales and available information as described in [D6.5]. The details of this interworking are discussed in the following sections.

2.3 Procedures requiring interactions

Several elementary procedures showing the interactions among the functional blocks of the QoSMOS reference model were identified in Section 6.1 of [D2.2]. The corresponding MSCs were provided in [D2.2] as well. Some of these procedures require certain interaction between the CM-RM and the CM-SM functional blocks in the form of basic communication operations, and can be summarised as follows:

- *Spectrum usage and performance reporting*. This procedure is related to the provisioning of reports on spectrum usage and performance from the CM-RM to the CM-SM. These reports are used by the CM-SM as an input to elaborate and keep up-to-date the spectrum portfolios.
- *Spectrum portfolio management*. This set of procedures is related to the management of the spectrum portfolio by the CM-SM, especially the gathering of the information required for building-up the spectrum portfolio and the provision of the portfolio to the CM-RM. The interaction between the CM-SM and the CM-RM for spectrum portfolio management is described by a number of procedures:
    - *Portfolio deployment*. The purpose of this procedure is to deploy a spectrum portfolio either upon request of an associated CM-RM or upon request of a next-layer CM-SM entity. Deployment communication flow is always down-level, i.e. from next-layer CM-SM to CM-SM to CM-RM.
    - *Portfolio revocation*. The purpose of this procedure is to revoke a spectrum portfolio either upon request of an associated CM-RM or upon request of a CM-SM or next-layer CM-SM entity. A spectrum portfolio may be revoked because it is not used any longer or it is for some reason reallocated (deployed) to another entity.
    - *Portfolio update or modification*. The purpose of this procedure is the modification of a spectrum portfolio either upon request of an associated CM-RM or upon request of a CM-SM or the next-layer CM-SM entity. The modification of portfolios is an alternative to replacing portfolios. It is applicable to CM-RMs when extending or reducing their portfolios in use (e.g., when requesting backup channels). It is also applicable to CM-SMs managing portfolio repositories more efficiently. Basic modification operations on portfolios are split and merge.
- *Spectrum policy management procedures.* This set of procedures is related to the management of the spectrum usage policies and constraints associated to the deployed spectrum portfolios.
    - *Policy set.* The purpose of this procedure is to set up a policy for a corresponding spectrum portfolio. The policy determines the set of rules that must be followed and the conditions under which a deployed spectrum portfolio can be exploited by a QoSMOS entity. The final set of rules and conditions set for a particular spectrum portfolio is determined by the regulatory policies of a frequency regulation agency and the operator’s own policies. The setting of a new policy is initiated by the entity in charge of deploying the spectrum portfolios (i.e., the CM-SM).
    - *Policy revocation.* The purpose of this procedure is to revoke a previously set spectrum policy for an already deployed spectrum portfolio. A policy revocation request is initiated by a CM-SM and results in the revocation of the corresponding policy in the requested CM-RM. After a policy revocation, the associated set of rules and conditions do not apply any longer.
    - *Policy update or modification*. The purpose of this procedure is the modification of a previously set spectrum policy for an already deployed spectrum portfolio. A policy update or modification request is initiated by a CM-SM and results in the modification of the corresponding policy in the requested CM-RM. After a policy modification, the new set of rules and conditions for the exploitation of the spectrum portfolio applies.
- *Cognitive managers’ operation.* These procedures are related to the management of CMs that start/cease their operation or reconfigure their operation parameters.
    - *Entity registration*. The purpose of this procedure is to register a CM-RM with the closest CM-SM and requesting an initial spectrum portfolio. If the closest CM-SM cannot resolve the portfolio request from its local portfolio repository, then the “next

layer" CM-SM is contacted in order to resolve the request, though CM-RM interacts directly only with its closest CM-SM.

- *Entity reconfiguration*. The purpose of this procedure is to trigger the reconfiguration of a CM or inform another CM that a reconfiguration has been performed. The reconfiguration of a CM-SM can be triggered internally or by a next-layer CM-SM, but not by a CM-RM. When this occurs, the CM-SM informs the associated CM-RMs that a reconfiguration has been performed. On the other hand, the reconfiguration of a CM-RM can be decided and initiated either internally by the CM-RM or by the CM-SM. In both cases the reconfiguration is informed/requested and confirmed. Reconfiguration procedures may occur for example in terms of operation distance from the primary system, transmission powers and load levels/number of users operating over a certain spectrum band.
- *Entity deregistration*. The purpose of this procedure is to deregister a CM-RM or CM-SM entity before a controlled shutdown. When an entity needs to shut down (i.e., an instance of a CM-RM or CM-SM), the associated CM-SM(s) or CM-RM(s) need to be informed about its intention. A CM-SM then must revoke or invalidate deployed spectrum portfolios before the requesting CM-RM or CM-SM is allowed to shut down. Usually, some safety measures have to be taken in order to avoid inconsistencies in the CM-SM repository in case a CM-RM or CM-SM disconnects unexpectedly. This case and its associated restart procedure applied after an unsolicited disconnect is considered out of scope for this procedure description.

It is worth noting that there are other more specific procedures not included in the categorisation above that may require interactions between the CMs, as it the case, for example, of the detection of the Radio Access Technologies (RATs) available in the vicinity of a QoSMOS system or the update of the transmission parameters. Such specific procedures are out of the scope of the level of detail of this chapter, and due to the nature of this chapter, they are therefore not treated here. A detailed description of such procedures along with the corresponding MSCs can be found in [D2.2] and [D2.3].

2.4 Practical examples of CM-RM/CM-SM interworking capacity

This section provides a discussion of selected examples requiring interworking and collaboration between the CM-RM and CM-SM entities in some practical situations in QoSMOS scenarios.

2.4.1 Cellular QoSMOS example use case: CM-SM assigns a spectrum portfolio to the CM-RM

For the QoSMOS cellular scenario, a typical CM-SM/CM-RM interworking use case is the procedure where a CM-RM entity is assigned a new or an updated spectrum portfolio and/or new configuration parameters. There are several possible variations in the sequence of messages, some messages depend on who is triggering and the involved messages do also depend on which/how recent information is considered in the spectrum-portfolio configuration decision. The kind of message sequences is illustrated in Figure 10. The procedure could be logically separated into the following steps:

1) *Triggers to initiate the procedure*

   Different entities can trigger the procedure for re-evaluation and for possibly re-assignment of a new spectrum portfolio. These possible triggers are:

   a. The CM-RM judges whether a different spectrum portfolio may be better suited for itself and sends a trigger message to the CM-SM, asking to re-evaluate its spectrum portfolio. This trigger could also be indirectly given by the CM-RM by sending its current CM-RM status which then induces that the CM-SM will judge that there is the need to trigger this procedure.

b. The CM-SM judges that there is the possibility that a different spectrum portfolio may be better suited for the CM-RM and initiates itself the evaluation procedure.

c. External entities may also send information messages result in that the CM-SM evaluation procedure is started. For example, the CM-SM gets informed from a spectrum repository about new spectrum constraints, which the CM-SM then has to apply for its spectrum portfolio.

2) *Information collection*

The CM-SM bases its decisions on its available information. These can either be previously obtained and stored information as well as it is possible that the CM-RM requests current information directly prior to starting its evaluation procedure. For example the CM-SM could provide to the CM-SM some information about its load status and about how well a certain part of the spectrum could be utilised. Information exchanged with other CM-SMs may also be considered.

3) *Spectrum portfolio / configuration assignment*

The CM-SM then evaluates the situation and could come to the decision that a different spectrum portfolio and or different configuration would be better suited for the CM-RM. Then the CM-SM assigns this new spectrum portfolio and/or other configurations to the CM-RM.

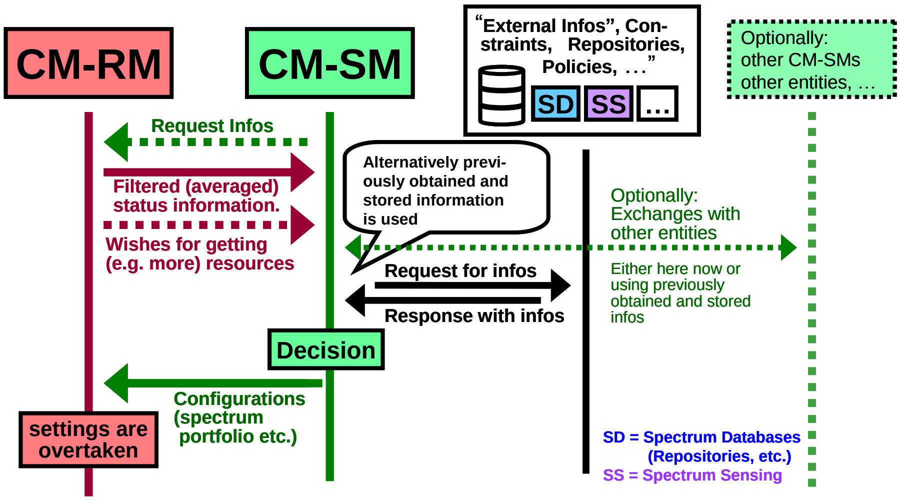


Figure 10. Illustration of signalling message sequences

2.4.2 Handover procedure in QoSMOS cellular scenarios

In cellular scenarios one of the typical challenges faced by the network is how to manage mobility from different users in a seamless way so as to maintain the QoS levels fixed in the Service Level Agreement (SLA). The mobility in cognitive radio systems is seen from two perspectives:

- The first one is physical mobility, meaning that a User Element (UE) connected into a certain cell is going out of its coverage area, so it is needed to reconnect it to the Access Point (AP)/eNodeB of a new cell.
- The second case considers mobility in spectrum, since the entrance of an Incumbent User (IU) in a cell could lead to the users operating in opportunistic spectrum to be moved to another spectrum location.

In both cases a user is changing their connectivity parameters and a reconfiguration of the UE is needed. Since the cellular case is centralised in terms of control, the AP should provide the details of the new connection and so as to do so different interactions between CM-RM and CM-SM take place.

CM-RM is in charge of carrying out resource allocation and admission control duties. In order to perform these tasks it is mandatory to be aware of the current spectrum usage so as to make the better decision, thus reducing as much as possible the impact over UE QoS. The entity updating the status of the spectrum for a given zone is the CM-SM, so this implies the interaction between the two entities.

Figure 11 illustrates a scenario where an IU comes into a cell forcing an opportunistic user (OU) to be moved from the opportunistic resources it was using. In this single scenario both handovers have to be addressed, as this covers not only physical mobility of the IU from cell A to cell B, but also spectrum mobility as the OU in cell B is not allowed to continue using these resources anymore since the IU is going into the cell. As a result of the spectrum handover of the OU, the CM-RM entity in cell B needs an updated version of the portfolio to be exploited in order to support the new OU. Notifying this situation, along with the new needs resulting from the presence of the new OU, and retrieving and updated spectrum portfolio requires a close interworking and collaboration of the CM-RM in cell B with its associated CM-SM. Once the CM-RM in cell B gets the information needed and the updated spectrum portfolio, it is then able to make the most convenient decision to support the new OU as well the rest of OUs.

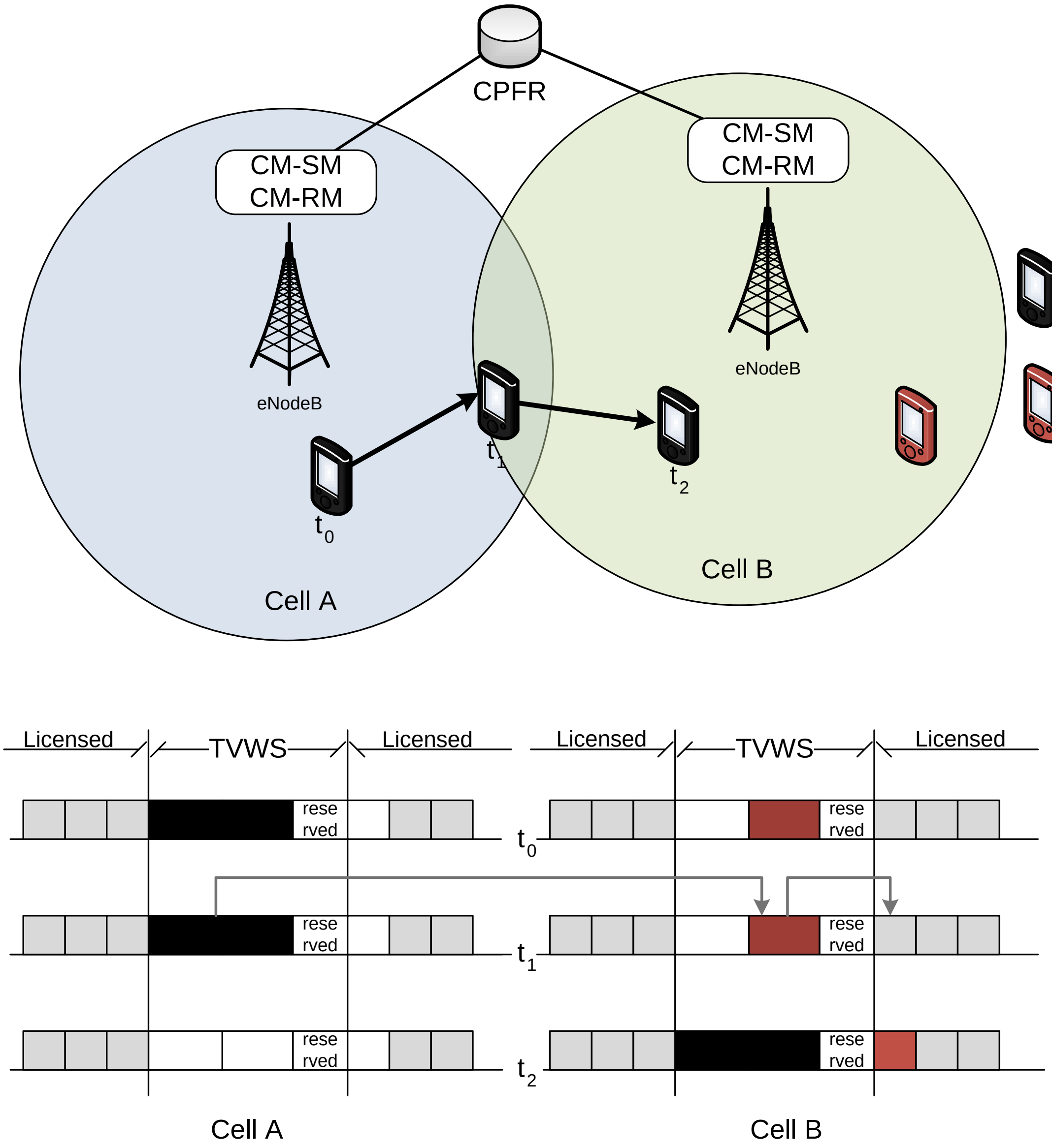


Figure 11. Handover in cellular scenarios.

3 Evaluation and capacity of Repositories

3.1 Global Regulator Repository (GRGR) – interface and functions

3.1.1 TV whitespace geo-location database functional description and evaluation

One of the key functions to enable opportunistic spectrum access is efficient automated spectrum management. One spectrum management method is a geo-location database approach, which has been adopted in the recent rules of Federal Communications Commission in the US to allow opportunistic spectrum access to use TV whitespace. In the FCC scenarios, the devices of the opportunistic systems identify their locations by using a geo-location capability and then query the database to determine which TV channels they can use at their locations. To verify that the reuse of the TV channel does not cause harmful interference to the incumbent receiver, the database checks whether the signal-to-interference ratio of the incumbent receiver can be kept at a required level. A transmit power which just satisfies the signal-to-interference ratio requirement corresponds to an allowable transmit power of the opportunistic system. Thus, the opportunistic system (or the database) has to estimate the signal-to-interference ratio of the incumbent receiver to determine the allowable transmit power. Pathloss prediction using propagation models is one signal-to-interference ratio estimation method. However, the propagation model inevitably includes prediction error in actual radio environments, resulting in signal-to-interference ratio estimation error. To absorb this error, a margin of the allowable transmit power has to be considered to keep the required signal-to-interference ratio. Consequently, the allowable transmit power has to be limited, and thus white space utilization efficiency might be reduced. Therefore, it is important to improve the signal-to-interference ratio estimation accuracy for expanding the white space opportunities.

Within QoSMOS, a prototype TV whitespace database (TVWSDB) has been implemented for demonstration purposes. At present it is populated only with the TV transmission data in the UK; nevertheless, the database is capable of adding data related to any country. The primary client communication method is via XML-RPC (http://xml-rpc.org), which is a remote procedure call standard capable of accepting clients in almost any computer language in an easy and efficient manner. A demonstration web-browser interface based on the above is available at http://www.ict-qosmos.eu/project/demos.html. The demonstration uses php scripts to translate the user requests into XML-RPC. In addition spectrum management entity written in C++ language (developed within QoSMOS) is able to communicate with this TVWSDB. A detailed description of this method is presented in Section 3.1.2 and deliverable D6.7.

The purpose of the remainder of this section is to specify the functional design of the TVWSDB, and its interface.

3.1.2 Spectrum Databases – Whitespace Estimator

There are two fundamental methods for assessing availability of spectrum: sensing, and the use of a pre-computed database. It is likely that many white space secondary devices will have to combine information from both for optimal performance. We here show how the database is generated and used. There are two aspects: a pre-computed database must be generated, and a server and associated protocol must be created and defined. Though there is no novelty in the individual components of this construction, the new field of cognitive radio has created a new need for a good integration of these components, and the QoSMOS project is thus contributing to this integration process.

We thus first describe methods for spectrum availability estimation, with examples for the United Kingdom. The assessment uses Huffman's irregular terrain propagation model (ITM; see [Huffman]), which is a variant of the Longley-Rice model. This model incorporates empirical fitting to measured signal data, especially regarding diffraction effects over hills. Future systems will most likely use more standardized propagation models, such as ITU-R P.546-11 and P.1812-1, though these typically slow the computation considerably.

The database construction starts by pixellating the country to be covered, typically in 100 metre squares. This process considers every pixel, despite the fact that the UK TV planning model only considers populated pixels. The obligation from Ofcom is to protect DTT coverage of 98.5% of households in the UK. The modelling assumes that if there is no line of sight to the transmitter from a given pixel (taking account of terrain), the pixel can be taken to be in a whitespace area. This simplifies the calculation, and is found to be true in practice. Otherwise, we compute the field strength at the pixel, using terrain data and list of TV transmitter information provided by the national regulator. This field strength is recorded in the database. Figure 12 gives a rough general impression of spectrum occupancy in Great Britain (blue=low, red=high, with intermediate values according to spectral colours) computed in this way.

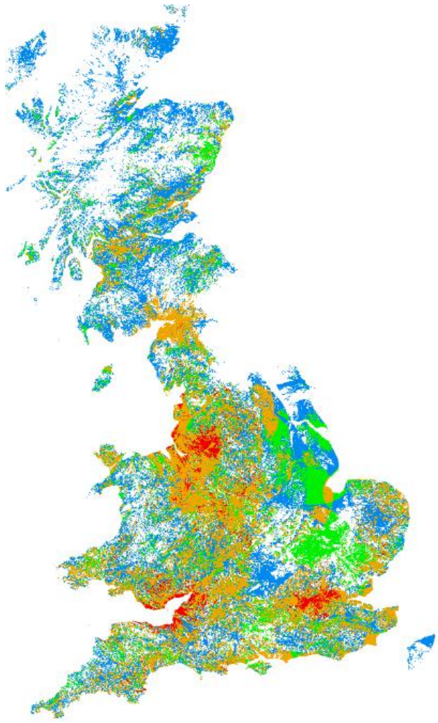


Figure 12 TV spectrum occupancy in Great Britain (blue=low, red=high).

Another requirement for practical application of TVWS is to know the pathloss from a proposed base-station site. For this purpose, we use again the Huffman propagation model and plot contours of equal signal strength. Figure 13 shows typical results, which are based on the BT research facility in Martlesham Heath, UK.

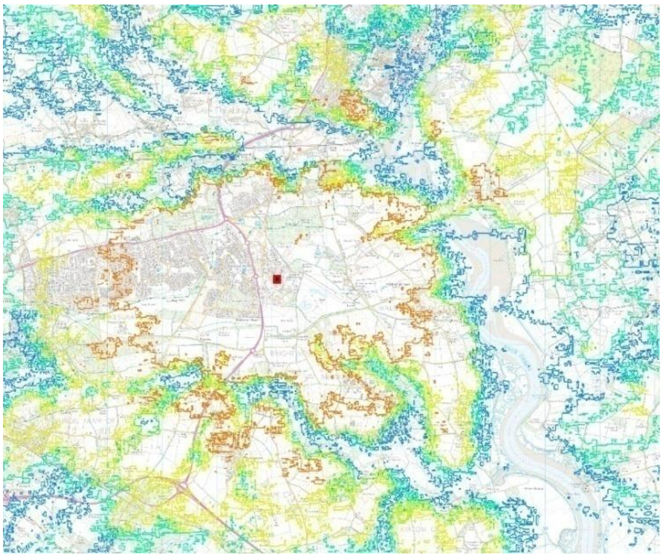


Figure 13 Estimated field-strength contours from a base-station at Martlesham Heath (UK)

The pathloss along a specific point-to-point path is even more important for a deployment. Figure 14 shows clearly how the terrain (green) affects the pathloss (red). Note that diffraction effects cause a sharp increase in pathloss just behind the lip of a valley, but the signal strength can recover for locations within the valley. In practice a pathloss of 140dB is about the maximum that a working deployment can tolerate.

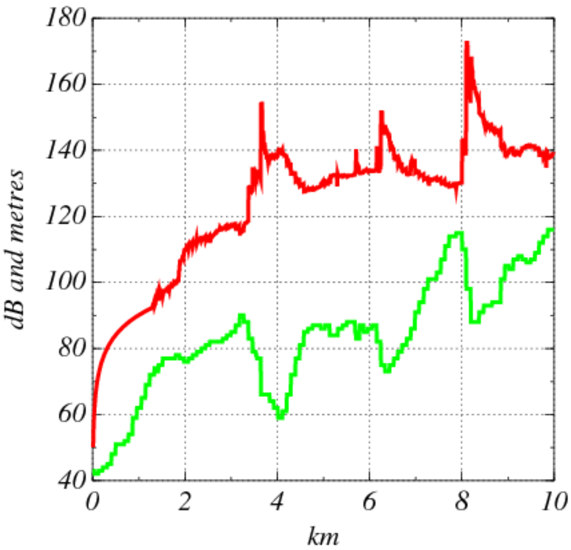


Figure 14 Typical pathloss dependence (dB, red) on terrain (green, metres)

The results are most useful if they are weighted with respect to population density, rather than per unit area. This is not considered here, due to the difficulty of getting data on population density. For this purpose the study uses the output of another internal BT project on housing density, which analyzed maps at the scale of a single square metre to locate houses in Britain. In Figure 15, the y-axis shows the fraction of population which is located in an area having *x* free channels (thus considered to be whitespace; each channel is 8MHz wide). This is based on a simplistic criterion – that a channel is free if the TV signal strength is below -100dBm. This ignores effects such as interference beyond the adjacent channels, and so the results are approximate and for rough guidance only. It can be seen, for example, that about 80% of the population have 17MHz or free spectrum available. There is a need for such calculations to be extended to other countries, both with higher and lower population density than Great Britain.

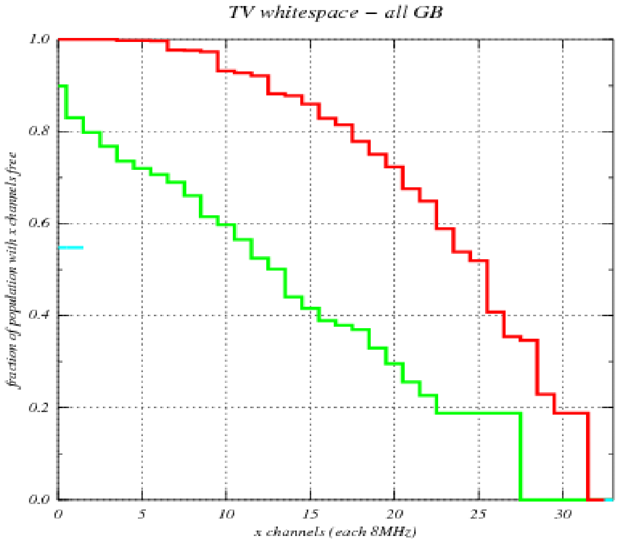


Figure 15 Cumulative population-weighted spectrum availability (red) in UK. Green=both adjacent channels also required to be free.

3.1.3 Database server design

This section describes the design of the database server which can be accessed by the clients. A typical block structure, used for the QoSMOS prototype database, is shown in Figure 16. The client (a TVWS secondary device) talks to the server with a RPC (Remote Procedure Call) type of interface, XML-RPC. The RPC server authenticates the client, and then looks up the static pre-computed database for the requested location. This is then returned to the client. Production-grade database systems will need further sophistication, such as a system to record which spectrum is already issued to TVWS devices in which locations. Furthermore, if competing databases are to be run by different service providers, some means of sharing information will be needed.

A TVWS device may thus query this database over a network connection; it sends its location, and TV channel occupancy at that location is returned by the database. From the data returned the TVWS device may choose a channel which will give minimal probability of interfering with TV reception, and also may set its own transmit power.

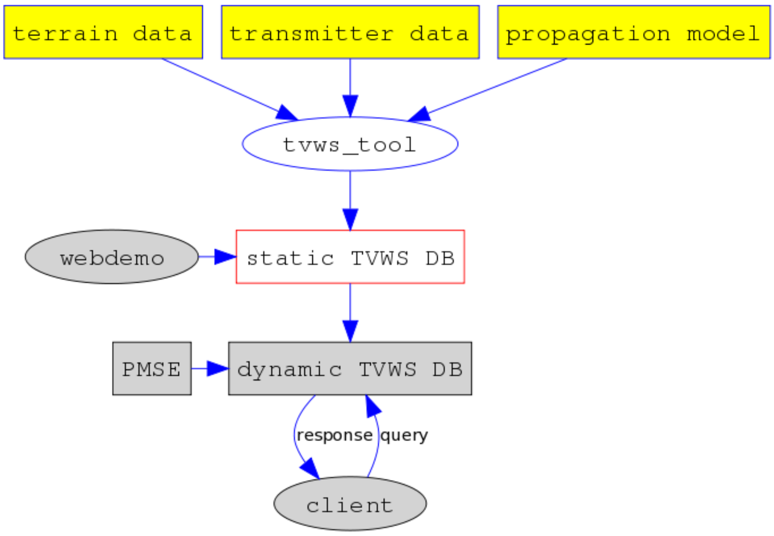


Figure 16 Database system architecture

3.1.4 Web-browser demonstrator

A prototype database of this type has been developed by BT as part of QoSMOS, and the link with the figure below provides a web-browser interface to this database. The system architecture is shown in Figure 16. The user should move the map marker to a location in Great Britain. A query is then sent to the database, and the data returned is displayed as a bar graph below the map. The bar graph gives the estimated TV channel occupancy at the selected map location.

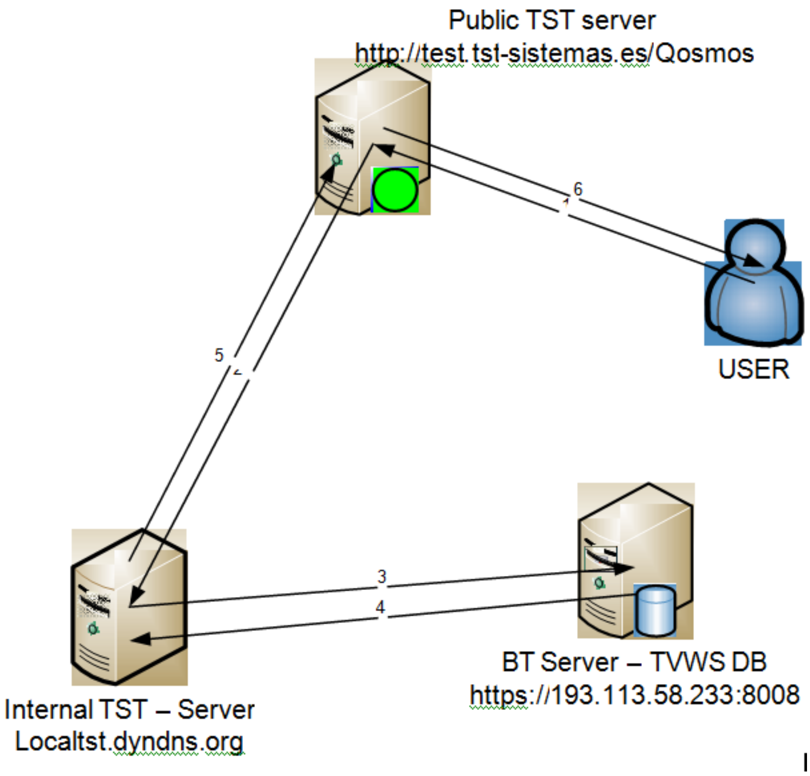


Figure 17 Web-demo system architecture

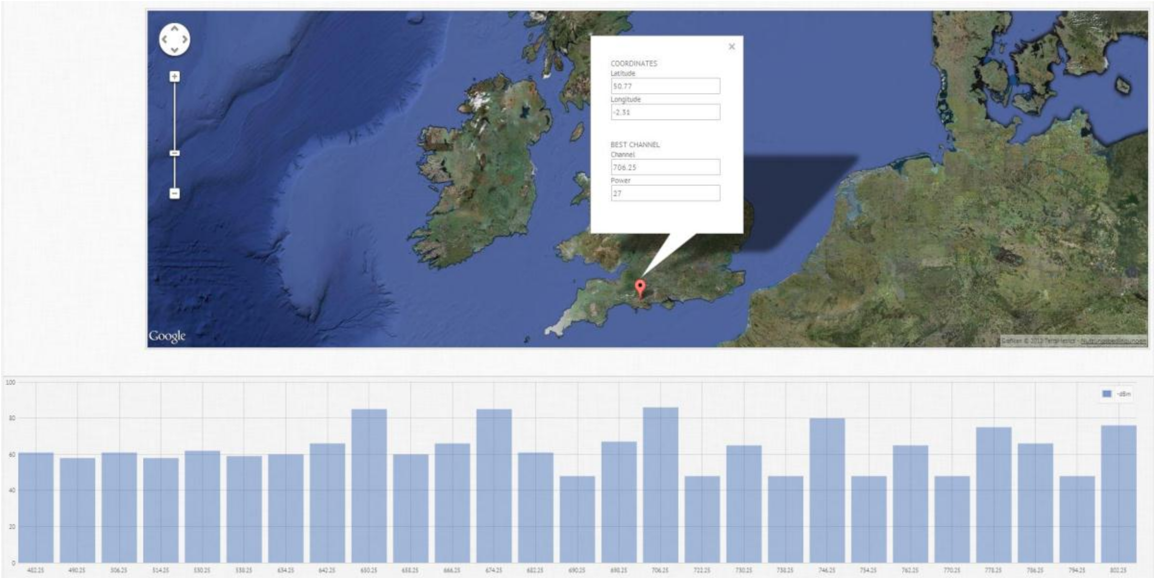


Figure 18 Web-browser interface

3.1.5 Appendix: coding API and examples

The prototype TVWSDB server is implemented in python. The transport protocol is https, and OpenSSL certificates are used for security. The server exports two XML-RPC functions :

1. `occupancy_lists_at_lonlat(longitude, latitude)`
2. `make_map((long_min,long_max),(lat_min,lat_max),image_size=(500,500))`

The first computes spectrum occupancy data in a single pixel; the second computes spectrum occupancy data in a set of pixels corresponding to a rectangular region in longitude-latitude space. (Note that *longitude precedes latitude*, corresponding to the usual x-y coordinate system, but differing from the usual convention in geography.)

Because XML-RPC can only return a restricted range of data types (integers, floats, strings, and lists), both functions return a list:

1. occupancy_lists_at_lonlat(): returns a list of two equal-length lists; the first list contains frequencies of TV channels in MHz, and the second list the corresponding estimated signal strength.
2. make_map(): returns a list of two lists: the [0] element is a list of two integers given the image width and height in pixels; elements [1,2,3,...] are a flattened list of (red, green, blue) pixel colour values. From this an image (e.g. in png format) can easily be reconstructed.

A python code fragment to query the TVWSDB follows:

```
import xmlrpclib
server=xmlrpclib.Server('https://132.146.249.18:8008')
print server.occupancy_lists_at_lonlat(1.2800789,52.0601958)
```

A typical output is:

> python tvws_client.py

```
474.25 -87
482.25 -92
```

```
490.25 -58
498.25 -87
506.25 -58
514.25 -70
522.25 -87
530.25 -59
538.25 -71
634.25 -48
642.25 -62
658.25 -48
666.25 -63
682.25 -49
690.25 -59
698.25 -59
706.25 -59
714.25 -59
730.25 -59
738.25 -60
746.25 -64
754.25 -49
770.25 -49
778.25 -68
786.25 -50
802.25 -67
```

A C++ code fragment follows:

```
xmlrpc_env              env;
xmlrpc_value*           resultP;
xmlrpc_value*           firstElementP;
xmlrpc_value*           secondElementP;
xmlrpc_value*           firstValP;
xmlrpc_value*           secondValP;
xmlrpc_client*          clientP;
const char*                    url="https://193.113.58.233:8008";
const char* const              name = "TVWS database client";
const char* const              ver = "0.1";
const char* const              methodName = "occupancy_lists_at_lonlat";
struct xmlrpc_clientparms      clientParms;
struct xmlrpc_curl_xportparms         curlParms;
curlParms.no_ssl_verifypeer =         true; // we ask the client to use cURL only
curlParms.no_ssl_verifyhost =         true;
curlParms.user_agent        =         name;
clientParms.transport            = "curl";
clientParms.transportparmsP      = &curlParms;
clientParms.transportparm_size = XMLRPC_CXPSIZE( user_agent );
xmlrpc_env_init( &env );
xmlrpc_client_setup_global_const( &env );
xmlrpc_client_create(    &env,    XMLRPC_CLIENT_NO_FLAGS,    name,    ver,    &clientParms,
XMLRPC_CPSIZE(transportparm_size),&clientP );
xmlrpc_client_call2f( &env, clientP, url, methodName, &resultP,"(dd)", (xmlrpc_double) 0.5,
(xmlrpc_double) 51.0 );
xmlrpc_array_read_item( &env, resultP, 0, &firstElementP );
xmlrpc_array_read_item( &env, resultP, 1, &secondElementP );
int nList = xmlrpc_array_size( &env, firstElementP ) ;
for ( int n = 0; n < nList ; n++ ) {
  xmlrpc_array_read_item( &env, firstElementP, n, &firstValP );
  xmlrpc_array_read_item( &env, secondElementP, n, &secondValP );
  xmlrpc_double         freq;
  xmlrpc_int            att;
  xmlrpc_read_double( &env, firstValP, &freq );
  xmlrpc_read_int( &env, secondValP, &att );
  cerr << "  at " << freq << " MHz: " << att << " dB" << endl;
}
xmlrpc_DECREF( resultP ); xmlrpc_env_clean( &env ); xmlrpc_client_cleanup();
```

The following procedure is used for server validation;

1. The QoSMOS CM-RM requests a band in the TVWS
2. CM-SM checks if it has one available in its local repository
3. CM-SM does not find a local portfolio and suspends processing of the request.
4. CM-SM initiates a request to the Geolocation Database gateway
5. GDB gateway receives the request and opens a connection to the TVWS database.
6. GDB gateway request the occupancy map for the location given in the request and selects a band.
7. GDB gateway closes the connection to the TVWS database and constructs a portfolio from the occupancy map.
8. GDB gateway sends the portfolio to the CM-SM
9. CM-SM adds the portfolio received to its local repository and resumes processing of the CM-RM request
10. CM-RM receives the reply and adds the portfolio to its local repository for local use.

The whole process was observed to take approximately 170ms, with a maximum of 1050ms.

3.1.6 Conclusion

A demonstration system for computing a TVWS database has been described, with a fully functional and secure client-server access system. This is intended as a proof-of-concept, and the three interfaces tested (python, C++, and web) have proved the usefulness of the system design.

4 Evaluation and capacity of common spectrum control (CSPC)

4.1 Multi-objective portfolio optimization for spectrum selection and aggregation in cognitive radio

4.1.1 Summary

The advent of cognitive radio systems will allow for the simultaneous and adaptive use of licensed and unlicensed portions of the spectrum. This, in turn, will improve end-user satisfaction and will partially alleviate the spectrum scarcity problem. However, different bands of the spectrum not only have different network (e.g., radio and load) conditions, but also different licensing and/or billing agreements. Therefore, the allocation of spectrum should target both technical and economical objective functions. This problem can be conveniently formulated as a multi-objective or multi-criterion optimization problem. This section aims to adapt the concepts behind portfolio and multi-objective optimization, which are commonly used in the area of economics, to the particular case of cognitive radio systems. The objective is also to observe the advantages and potential interactions between technical and economical approaches. This section proposes a framework for such coexistence and provides some simulation results considering the Pareto optimal trade-off region obtained when using a multi-objective function that considers the average return and average risk values per portion of spectrum to be allocated. The results show that network and economical targets can be simultaneously met inside the optimality region given by the Pareto optimality criteria. Guidelines towards achieving such optimality in both domains are further provided.

4.1.2 Introduction

The increasing demand for higher data rates means that larger amounts of bandwidth are required to accommodate all user transmissions. However, spectrum is a limited resource which has been already allocated in its majority to other legacy services. These issues have led to a spectrum scarcity problem for next generation services. Despite this fact, it has been observed that several portions of the licensed spectrum remain largely under-utilized for long periods of time [PAL06]. This opens the possibility for a new type of smart devices that will be able to opportunistically access different frequency bands without disturbing primary/licensed user transmissions and that promises to solve the bandwidth demand problem. This type of smart device has been termed cognitive radio [LU12]-[MIT93].

Cognitive radio systems promise a networking paradigm with a fully dynamic and opportunistic access to different parts of the spectrum that will considerably improve end-user satisfaction. However, different parts of the spectrum are subject not only to different propagation and load conditions, but also subject to different licensing and billing agreements. Cognitive radio promises an adaptive and opportunistic access to different portions of the spectrum so as to improve end-user satisfaction and solve bandwidth scarcity problems. However, different parts of the spectrum will experience not only different propagation and load conditions, but also different licensing and billing agreements. For example, consider the terminal licensed to operate in LTE (long term evolution) in Figure 19. With the help of cognitive radio this terminal can also access in an opportunistic fashion an unlicensed frequency band such as the ISM band (industrial, scientific, and medical), which is commonly used for WiFi. Whenever the LTE service is degraded due to load or fading, and the ISM band is sensed as available, then the user can be handed over to the latter one without losing connection.

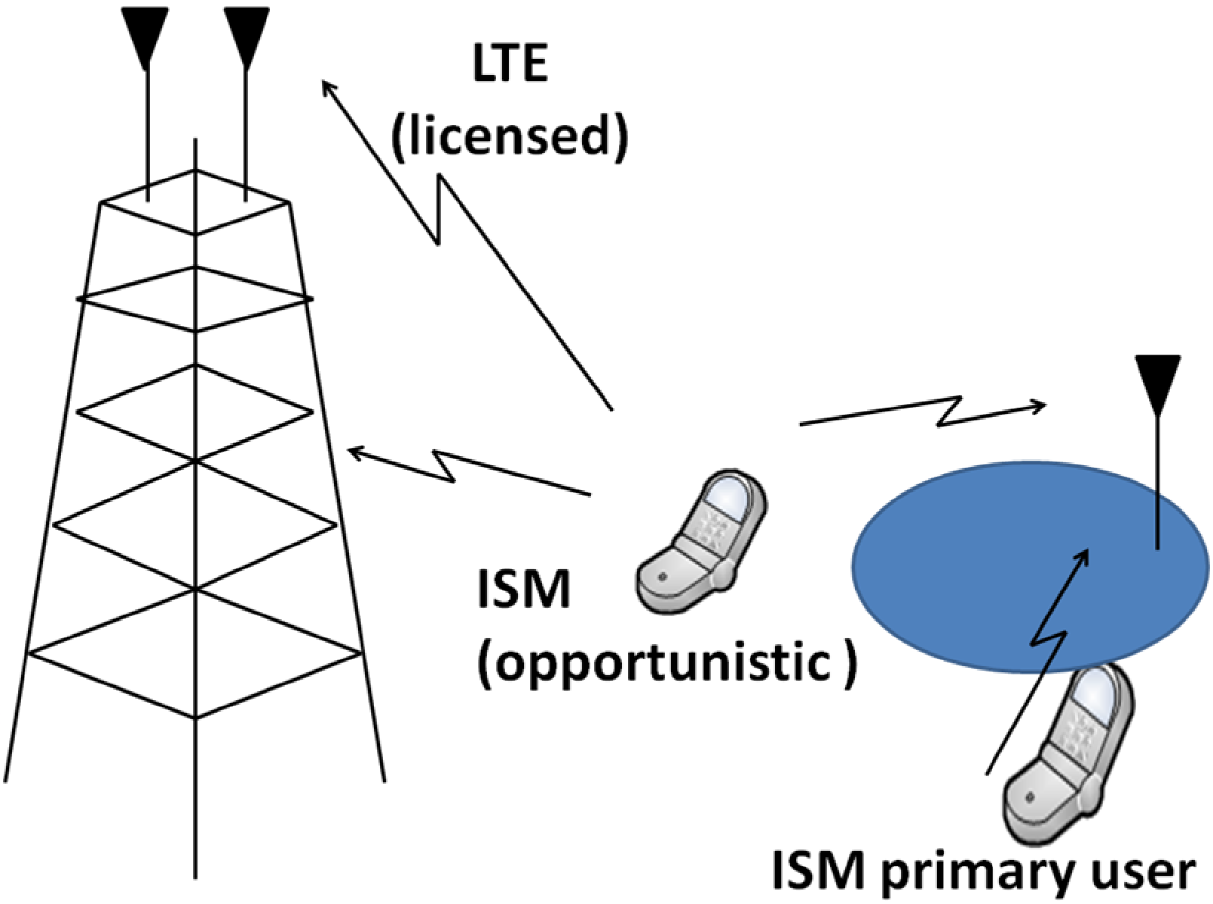


Figure 19 Cognitive radio system

In LTE, radio resources are usually allocated by the central scheduler in a dedicated fashion, thereby ensuring a good level of quality of service. By contrast, in the ISM band the availability of such resources is less reliable due to the contention process with all other primary users in contiguous WiFi networks or other ISM services. Regarding the billing of services, in the licensed LTE band the user can be charged according to the amount of data exchanged over the connection. This is because the operator has invested a considerable amount of money in spectrum licences and needs a return for its investment as well as a margin profit. By contrast, in the ISM band there are no fees for spectrum licenses. Therefore, if the connection is by means of WiFi, the user can be charged in terms of the duration of the connection with a fee relatively lower than in LTE. Conversely, it is also possible that the LTE operator does not owe ISM transmission technology, but instead it is provided by a third party operator that charges for hot-spot services over different geographical areas. In this case, the LTE operator might charge users with an additional fee for opportunistic services in the ISM band to cover the cost of the third-party networking services.

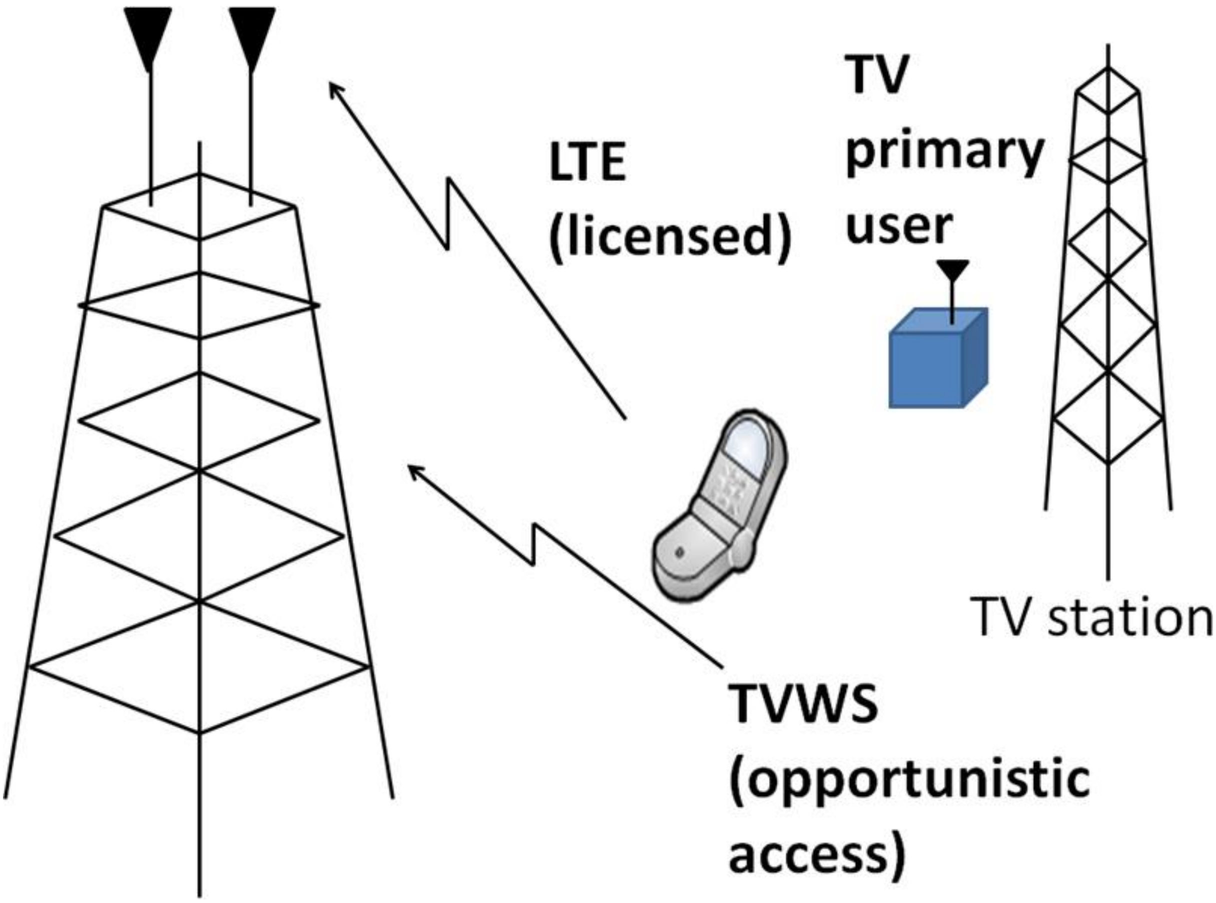


Figure 20 Example of cognitive radio

In another potential scenario in Figure 20, the same user licensed in the LTE network can also opportunistically access a TV white space (TVWS). In this case, the LTE operator could pay a licence to the owner of the TV spectrum for opportunistic access, and thus charge the user with a fee possibly higher than when using the licensed LTE connection to cover for additional expenses. In terms of quality of service, access to TVWSs can use a fixed resource allocation or a contention scheme depending on the requirements/policies to protect primary TV users.

As it can be observed with these examples, the licensing/billing agreements of different portions of spectrum as well as the resource management requirements can be considerably different and complex. In this heterogeneous landscape, operators face a three-fold problem with cognitive radio networks:

- How to allocate the spectrum resources so as to obtain the maximum network and spectral efficiencies?
- How to allocate spectrum resources in order to maximize the return of investment and margin profit while reducing risk or variations of the return?, and
- How to bill the users for the licensed and opportunistic services so that the prices are competitive and attractive in the market and the return and network resource efficiency are kept at an acceptable level?

In contrast to conventional networks, the allocation of the different radio resources must now reflect more accurately the complex interaction and trade-offs between network and economical objective functions. This problem can be formulated as a multi-objective portfolio optimization problem, which has been investigated in detail, for example, in the area of economics [FAB07]-[KON05].

The objective of this section is to reformulate the multi-objective portfolio optimization problem commonly used in economics, and use it in the context of spectrum selection/aggregation in cognitive radio networks. The idea is to propose a framework for radio resource allocation that satisfies both network and economical objective functions. The section also provides a review of the concepts of convex multi-objective optimization applied to the problem of portfolio allocation commonly found in the literature of economics. A modified multi-objective optimization problem is formulated that combines technical and economical factors for radio resource allocation in a scenario with two WiMAX-based networks (one indoor and one outdoor). System-level simulation results show that both economical and radio performance objectives can be simultaneously met, thereby increasing end-user satisfaction and operator revenues.

4.2 Convex Optimization (Overview)

This subsection provides a review of the concepts behind convex optimization, focusing on the multi-objective optimization problem. Techniques for solving this type of optimization are also discussed.

4.2.1 Scalar optimization problem

Let us start with the simplest optimization problem, where the objective function is a scalar value that depends on a vector variable. A scalar optimization problem has the form:

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) \\ \text{subject to} & f_i(\mathbf{x}) \le b_i \quad i = 1,\dots,m \end{array}$$

where

- $\mathbf{x} = [x_1,..,x_N]^T$ $\mathrm{x} = [\mathrm{x}_1, \dots \mathrm{x}_N]^T$is the optimization vector variable,
- $(.)^T$ is the vector transpose operator,
- $f_0 : \mathbf{R}^N \to \mathbf{R}$ is the objective function
- $f_i : \mathbf{R}^N \to \mathbf{R}$, $i = 1,\dots,M$ are the inequality constraints, and
- $b_1,\dots,b_M$ are the bounds of the constraints.

A vector $\mathbf{x}_{opt}$ is called optimal if $f_1(\mathbf{z}) \le b_1,\dots, f_m(\mathbf{z}) \le b_m$ and $f_0(\mathbf{z}) \ge f_0(\mathbf{x}_{opt})$ for $\mathbf{z} \in \mathbf{R}^N$. If $f_i(\alpha\mathbf{x} + \beta\mathbf{y}) \le \alpha f_i(\mathbf{x}) + \beta f_i(\mathbf{y})$, $\alpha + \beta = 1$, $\alpha \ge 0$ $\mathrm{\alpha} \geq 0$; $\alpha, \beta \in \mathbf{R}$; and $\mathbf{x,y} \in \mathbf{R}$ $\mathrm{x, y \epsilon R^N}$, then the optimization is convex [6]. The importance of the convexity property lies on the fact that it ensures that iterative techniques based on successive approximations (discussed later in this section) can converge to the optimum solution.

4.2.2 Vector optimization problem

Let us now extend the scalar optimization problem to the vector case, i.e. a vector optimization problem. A vector optimization problem has the form [6]:

$$\begin{array}{lll} \text{minimize} & \text{w.r.t}\,\mathbf{K} & \mathbf{f}_0(\mathbf{x}) \\ \text{subject to} & f_i(\mathbf{x}) \le b_i & i = 1,\dots,m \end{array}$$

where

- $\mathbf{K} \subseteq \mathbf{R}^N$ is a proper cone, and
- $\mathbf{f}_0 : \mathbf{R}^N \to \mathbf{R}^Q$ is the vector objective function

If the objective and constraint functions are convex, then the optimization problem is convex too. The optimality region O is given by:

$$O = \{\mathbf{f}_0(\mathbf{x}) \mid \exists \mathbf{x} \in D, f_i(\mathbf{x}) \le b_i, i = 1,...,M\} \subseteq \mathbf{R}^Q$$

where D is the domain of the optimization variable. A vector $\mathbf{x}_{opt}$ is called optimal if it is feasible and $O \subseteq \mathbf{f}_0(\mathbf{x}_{opt}) + \mathbf{K}$. However, in most of the cases, a vector optimization problem does not have a unique optimum solution, but rather a large number of solutions which are optimum for one or only a subset of the components of the vector objective function. To deal with this issue, then the concept of Pareto optimal solution is commonly used. A Pareto optimal solution can be loosely defined as that solution that is optimum for at least one of the vector components of the objective function. More formally, a vector **x** is Pareto optimal if $(\mathbf{f}_0(\mathbf{x}) - \mathbf{K}) \cap O = \mathbf{f}_0(\mathbf{x})$. The set of Pareto optimal vectors P complies with the boundary condition $P \subseteq O \subseteq \mathbf{bd}O$, where bdO indicates the boundaries of the optimality region. To make it more tractable, a multi-objective optimization problem can be reformulated, using the method of scalarization, as follows:

$$\begin{array}{ll} \text{minimize} & \boldsymbol{\lambda}^T \mathbf{f}_0(\mathbf{x}) \\ \text{subject to} & f_i(\mathbf{x}) \le b_i \quad i = 1,...,m \end{array}$$

where $\boldsymbol{\lambda} = [\lambda_1,, \lambda_Q]^T$ is the vector of relative weighting factors for each component of the vector objective function. If $\mathbf{x}^*$ is a solution to this problem then it is a Pareto optimal point of the original problem [6]. This means that we can solve this optimization problem for different values of the weighting factors λ_j.

4.2.3 Multi-objective optimization problem

When a vector optimization problem involves the cone $\mathbf{K} \subseteq \mathbf{R}_+^Q$, it is called a multi-criterion or multi-objective optimization problem. The components of the objective functionf$_0$, say $\mathbf{f}_0 = \lfloor F_1,..., F_Q \rfloor$ can be interpreted as Q different scalar objectives, each of which we would like to optimize. We refer to F_j as the *j*-th objective of the problem. A multi-criterion optimization problem is convex if all inequality constraint functions and the objective functions are convex. Since multi-criterion problems are vector optimization problems, all the material for the latter ones applies for the former ones. The set of Pareto optimal values for a multi-criterion problem is called the optimal trade-off surface in the case of q>2, or the optimal trade-off curve (q=2)

The multi-objective optimization problem can be reformulated, using the method of scalarization, as follows:

$$\begin{array}{ll} \text{minimize} & \boldsymbol{\lambda}^T \mathbf{f}_0(\mathbf{x}) \\ \text{subject to} & f_i(\mathbf{x}) \le b_i \quad i = 1,...,m \end{array}$$

where $\boldsymbol{\lambda} = [\lambda_1,, \lambda_Q]^T$ is the vector of relative weighting factors for each component of the vector objective function. If $\mathbf{x}^*$ is a solution to this problem then it is a Pareto optimal point of the original problem.

4.3 Techniques for solving the optimization problem

4.3.1 Lagrange multipliers

The technique of Lagrange multipliers is perhaps the most used for solving optimization problems. Briefly speaking, it consists of formulating the Lagrange function as follows [6]:

$$g(\mathbf{x},\mathbf{v}) = f_0 - \sum_{i=1}^{M} v_i f_i$$

The dual problem is:

$$\text{minimize } g(\mathbf{x},\mathbf{v})$$

where $\mathbf{v} = [v_1,\ldots,v_M]^T$ is the vector of Lagrange multipliers.

4.3.2 Gradient steepest descent

The technique of Lagrange multipliers cannot always be solved in closed-form. Therefore, numerical methods are often employed. However, even in the latter case it can be difficult to obtain expressions for the derivatives of the objective and constraint functions. In this case, iterative techniques are commonly employed. This type of technique consists of reaching the optimum solution by means of successive approximations. The gradient steepest descent method exploits the property of the gradient vector function $\textbf{grad}\ f_0$ gradf$_0$ of pointing into the direction of maximum growth rate of the objective function, so as to calculate a new position that is closer in the direction of the optimum solution. This can be mathematically written as:

$$\mathbf{x}_{opt}(n) = \mathbf{x}_{opt}(n-1) + \xi \frac{\textbf{grad}\ f_0}{|\textbf{grad}\ f_0|}$$

where ξ is the step size used to adjust the converge properties of the iterative scheme.

4.4 Spectrum aggregation and selection as a multi-objective portfolio optimization problem

Consider a cognitive radio system, where *N* different frequency bands of the spectrum can be allocated (see Figure 21). For simplicity, let us assume that each band *i* of the spectrum is infinitely divisible, so that we can allocate a portion x_i, where $0 \le x_i \le 1$. This can be stated as a portfolio optimization problem where we wish to select a portion x_i of the *i*-th asset of a set of *N* possible assets. The objective in a portfolio optimization problem is to maximize the total return (the economic gain of allocating all the portions of the frequency bands) while minimizing or controlling the risk of such allocation (statistical variance of the return).

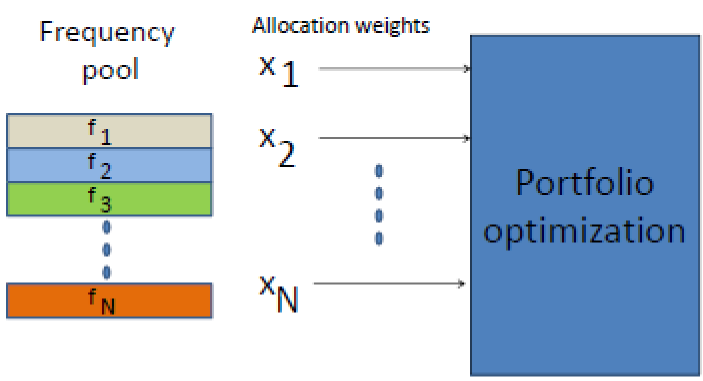


Figure 21 Spectrum portfolio optimization

Consider $\mathbf{x} = [x_1,.., x_N]^T$ as the vector of allocation weights, while the instantaneous return or economic gain for allocating a portion of the *i*-th asset will be denoted as p_i pi. Let us now arrange the instantaneous return values in the vector $\mathbf{p} = [p_1,.., p_N]^T$ and the average or mean vector return as $E[\mathbf{p}] = \bar{\mathbf{p}}$, where $E[.] = \overline{(.)}$ is the statistical average operator. The covariance matrix can be written as $E\left[(\mathbf{p}-\bar{\mathbf{p}})^T(\mathbf{p}-\bar{\mathbf{p}})\right] = \mathbf{C}$. Therefore the total return and risk of a given selection $\mathbf{x}$ can be written as:

$$\text{Average return} = \bar{\mathbf{p}}^T \mathbf{x}$$

$$\text{Average risk} = \mathbf{x}^T \mathbf{C} \mathbf{x}$$

A multi-objective optimization problem, where the aim is to maximize the average return and minimize the average risk, can thus be stated as follows [6]:

$$\begin{aligned} &\text{minimize} \quad \text{w.r.t}\, \mathbf{K}^2 \quad \mathbf{f}_0 = [F_1, F_2] = \left[-\bar{\mathbf{p}}^T \mathbf{x}, \mathbf{x}^T \mathbf{C}\mathbf{x}\right] \\ &\text{subject to} \quad \sum_{i=1}^{N} x_i = 1 \qquad 0 \le x_i \le 1 \end{aligned}$$

which can be reformulated, using the method of scalarization, as follows:

$$\begin{aligned} &\text{minimize} \quad \text{w.r.t}\, \mathbf{K}^2 \quad -\bar{\mathbf{p}}^T \mathbf{x} + \mu \mathbf{x}^T \mathbf{C}\mathbf{x} \\ &\text{subject to} \quad \sum_{i=1}^{N} x_i = 1 \quad {}_i \qquad 0 \le x_i \le 1 \end{aligned}$$

where μ is a scalar value that defines the balance or trade-off between return and risk. The optimization problem can be easily solved using tools of convex optimization. The constrained problem boils down to solving a system of simultaneous equations given by:

$$(\tilde{\mathbf{C}} + \tilde{\mathbf{C}}^T)\tilde{\mathbf{x}} + \mathbf{b} = \mathbf{0} \text{ with the constraint } \sum_{i=1}^{N} x_i = 1$$

where, using MATLAB notation, $\tilde{\mathbf{x}} = \mathbf{x}(1:N-1)$, $\tilde{\mathbf{C}} = \mathbf{B}(1:N-1, 1.N-1)$, while $\mathbf{B}$, and $\mathbf{b}$ are calculated as follows:

$$\mathbf{B}(i,j) = \tilde{\mathbf{B}}(i,j) - \tilde{\mathbf{A}}(N-1,j) \quad \text{and} \quad \mathbf{b} = \mathbf{B}(n-1,:) + \mathbf{d}\text{, where}$$

$$\tilde{\mathbf{A}}(i,j) = \mathbf{A}(i,j) - \mathbf{C}(i,N),\ \mathbf{A} = \mathbf{C}(:1:N-1)$$

$$\tilde{\mathbf{B}} = \tilde{\mathbf{A}}(:,1:N-1)$$

$$\mathbf{B}(i,j) = \tilde{\mathbf{B}}(i,j) - \mathbf{A}(N-1,j)$$

$$d_i = e_i - p_N\text{, and}$$

$$\mathbf{e} = \mathbf{p}(1:N-1)$$

The system of equations can be solved by simply inverting the matrix of the system of equations:

$$\tilde{\mathbf{x}}_{ZF} = -(\tilde{\mathbf{C}} + \tilde{\mathbf{C}}^T)^{-1}\mathbf{b},$$

or, using the least squares solution, which is useful in scenarios with more equations than variables (for example in systems with cooperative sensing):

$$\tilde{\mathbf{x}}_{LS} = -\left(\tilde{\mathbf{H}}^T\mathbf{H}\right)^{-1}\tilde{\mathbf{H}}^T\mathbf{b}$$

where $\tilde{\mathbf{H}} = \tilde{\mathbf{C}} + \tilde{\mathbf{C}}^T$. In systems with imperfect information, we can also use the minimum mean square error (MMSE) solution:

$$\tilde{\mathbf{x}}_{MMSE} = -\left(\tilde{\mathbf{H}}^T\tilde{\mathbf{H}} + \sigma^2\mathbf{I}\right)^{-1}\tilde{\mathbf{H}}^T\mathbf{b}$$

where σ^2 is the variance of the distribution of the available information. Some of the solutions in the above equations can be negative, so an iterative scheme must be used to set their values to zero and recalculate the optimum solution. The iterative scheme is illustrated in the flowchart below:

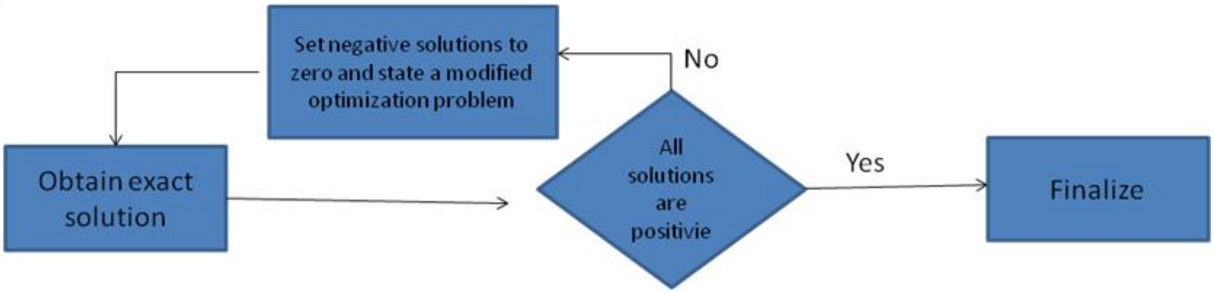


Figure 22 Iterative algorithm for solving the portfolio optimization problem

In a cognitive radio system, however, there might be several restrictions that can affect the form and the solution of the optimization problem. For example, the constraint $\sum_{i=1}^{N} x_i = 1$ means that the same amount of total bandwidth across different bands will be always allocated. This condition is perhaps useful in systems with power constraints. However, if one desires to gain capacity by using as much bandwidth from other bands as possible, then the constraint must be removed, which yields to:

$$\begin{array}{lll} \text{minimize} & \text{w.r.t}\,\mathbf{K}^2 & -\bar{\mathbf{p}}^T\mathbf{x} + \mu\mathbf{x}^T\mathbf{C}\mathbf{x} \\ \text{subject to} & & 0 \le x_i \le 1 \end{array}$$

whose solution is simply given by the solution to the following system of simultaneous equations:

$$(\mathbf{C} + \mathbf{C}^T)\mathbf{x} + \bar{\mathbf{p}} = \mathbf{0}$$

The system of equations can be solved by simply inverting the matrix of the system of equations:

$$\mathbf{x}_{ZF} = -(\mathbf{C} + \mathbf{C}^T)^{-1}\bar{\mathbf{p}}$$

or using the least squares solution, which is useful in scenarios with more equations than variables (for example in systems with cooperative sensing):

$$\mathbf{x}_{LS} = -(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\bar{\mathbf{p}}$$

where $\mathrm{H} = \mathrm{C} + \mathrm{C}^\mathrm{T}$. In systems with imperfect information, we can also use the minimum mean square error (MMSE) solution:

$$\mathbf{x}_{MMSE} = -(\mathbf{H}^T\mathbf{H} + \sigma^2\mathbf{I})^{-1}\mathbf{H}^T\bar{\mathbf{p}}$$

where σ^2 is the variance of the distribution of the available information.

An example of the portfolio multi-objective optimization problem is given by using the values in the table below for 5 spectrum bands with different values of return and risk. The optimum weight allocation for each band is displayed in Figure 23 and the Pareto optimal trade-off region is displayed in Figure 23 for the case of constrained optimization. The unconstrained case is displayed in Figure 25 and Figure 26. Note that in the constrained case, as the variance is increased, spectrum band number 1 dominates the allocation weights, while at low values of risk spectrum band number 5 dominates the allocation. In the unconstrained case though, the allocation is completely different. Spectrum band 5 is always allocated, while spectrum band 4 is never allocated.

Spectrum band	return (%)	risk(%)
1	15	30
2	10	20
3	5	10
4	2	5
5	2	0

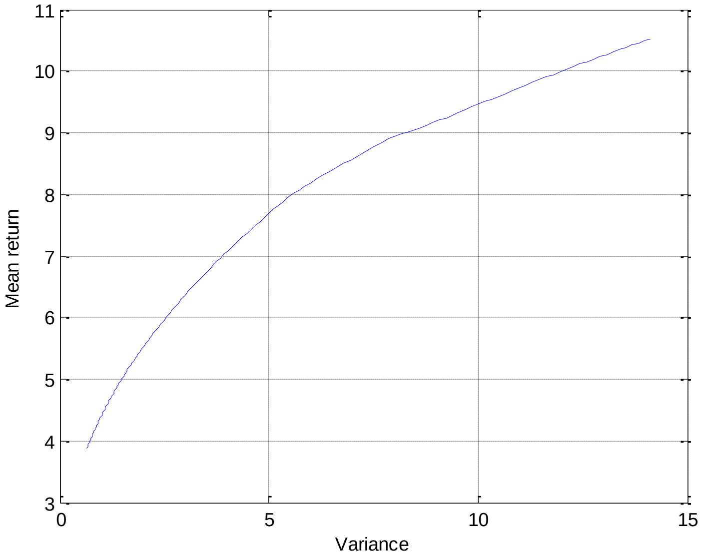


Figure 23 Pareto-optimal trade-off region - constrained case.

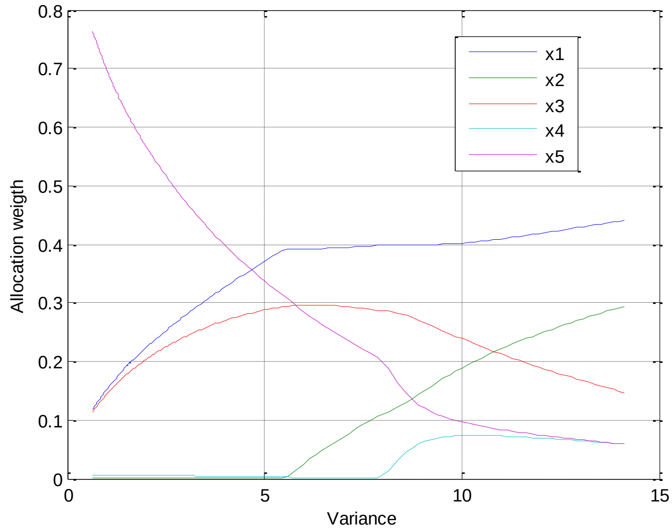


Figure 24 Optimum weights portfolio optimization - constrained case

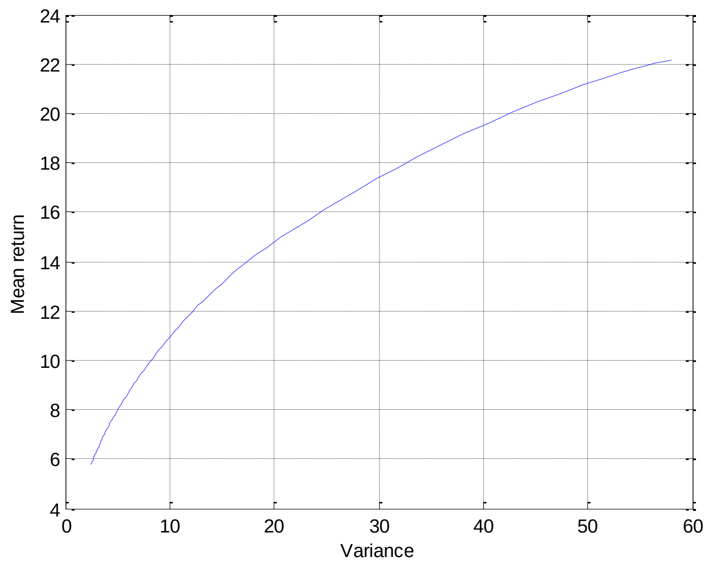


Figure 25 Pareto-optimal trade-off region - unconstrained case

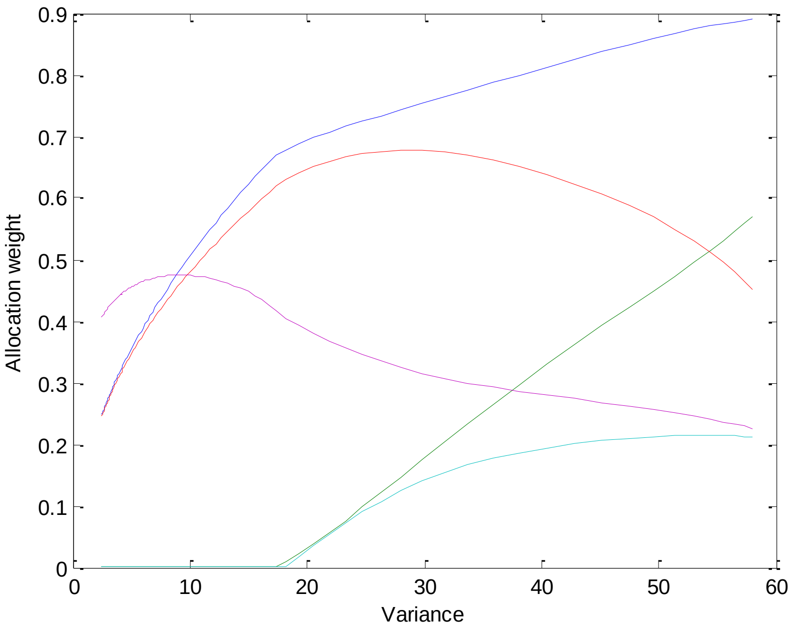


Figure 26 Optimum weights - portfolio optimization - unconstrained case

Another issue in the derivations presented above is that the objective functions are considering only economic or pricing information. Radio and network parameters need to be included as well in order to improve the final decision on frequency selection. For example, some parts of the assigned spectrum may have bad channel conditions, so they should not be assigned. Another aspect closely related to this issue is the pricing or licensing scheme used by the operators. The simplest pricing scheme is to have a fixed or flat rate per portion of spectrum for different users (primary and opportunistic or secondary). However, the licensing scheme can be more dynamic, depending on the type of connection, application, data-rate, and on the current network conditions. In these complex conditions, the optimization problem must include more parameters, both from the radio and the business worlds. This implies a fusion between the spectrum selection and the radio resource managers. The level of fusion between the two entities in the QoSMOS architecture will therefore depend on the pricing scheme used and the portfolio optimization method chosen. Finally, in some networks with ad-hoc and self-configuration features, the decisions might be taken in a decentralized manner, where pricing and network information are unreliable or available only in some parts of the network. Therefore, the portfolio optimization problem should also be modified to cope with these important constraints. To illustrate the different approaches (one with conventional optimization and another one with a modified objective function) the following two subsections address in more detailed each scenario.

4.5 Test scenario with conventional multi-objective portfolio optimization

Consider the deployment scenario depicted in Figure 27 with two WiMAX-based networks, one indoor and one outdoor, operating in two different frequency bands denoted by f_{in} and f_{out}, respectively. The outdoor network consists of a base station (BS) located at the centre of the deployment. The indoor network operates inside a rectangular building of dimensions denoted by *a* and *b*, and its BS is located at the center of the building which is located at the coordinates denoted by x_{in} $\mathrm{x_{in}}$ and x_{out}. The analysis is focused on the downlink considering a bandwidth of 10 MHz using the transmission parameters described in [WIXSLS] for TDD WiMAX systems. The resource management entity of the two networks is centralized at the BS of each network. A cognitive radio is assumed to be operating at the BS of each network. Each BS is able to allocate all the time and frequency resources of its operating frequency band and, in addition, allocate some users in an opportunistic manner in the neighbour frequency band of the adjacent network. The propagation model for outdoor is a non-dispersive channel using a pathloss model defined by the WINNER B1 model [WIND112]. Indoor propagation model is a B5 WINNER model [WIND112] also with flat fading statistics. A building penetration loss of 8 dB is used in all simulations.

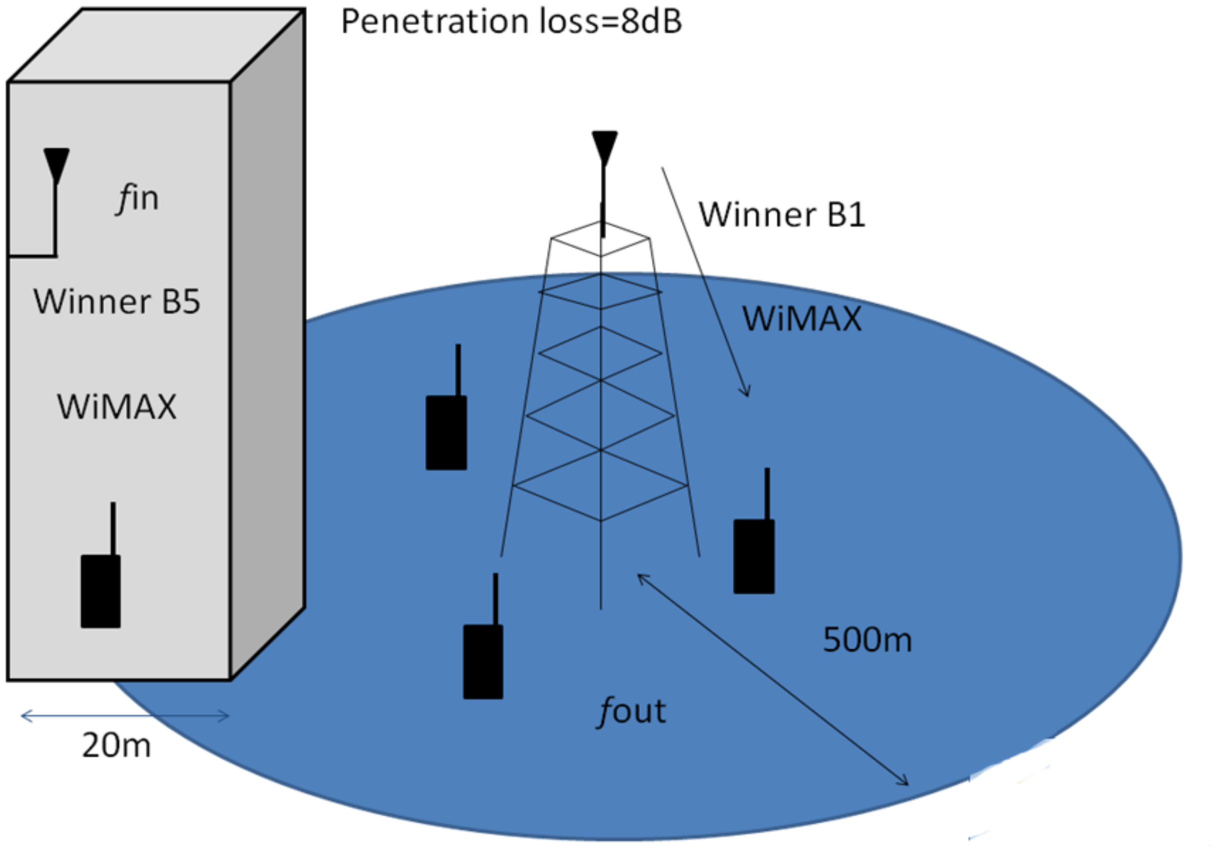


Figure 27 Testing scenario

Each TTI (transmission time interval), the RRM of each BS calculates the amount of bandwidth to be utilized in the primary and secondary frequencies using multi-objective portfolio optimization as described in previous subsections. Once the portion of each spectral band has been calculated, the RRM entity proceeds to allocate the users in the most appropriate radio resources. This means that frequency selection and resource allocation have small interaction, which means also low interaction between the spectrum and the radio manager in the QoSMOS architecture. An example with tighter interaction between the RRM and the spectrum selection entities is presented in the following subsection. Figure 28 displays the architecture of system-level simulator used for the evaluation of the proposed multi-objective optimization framework for cognitive radio. Figure 29 shows the frame for radio resource allocation considering OFDMA (orthogonal frequency division multiple access).

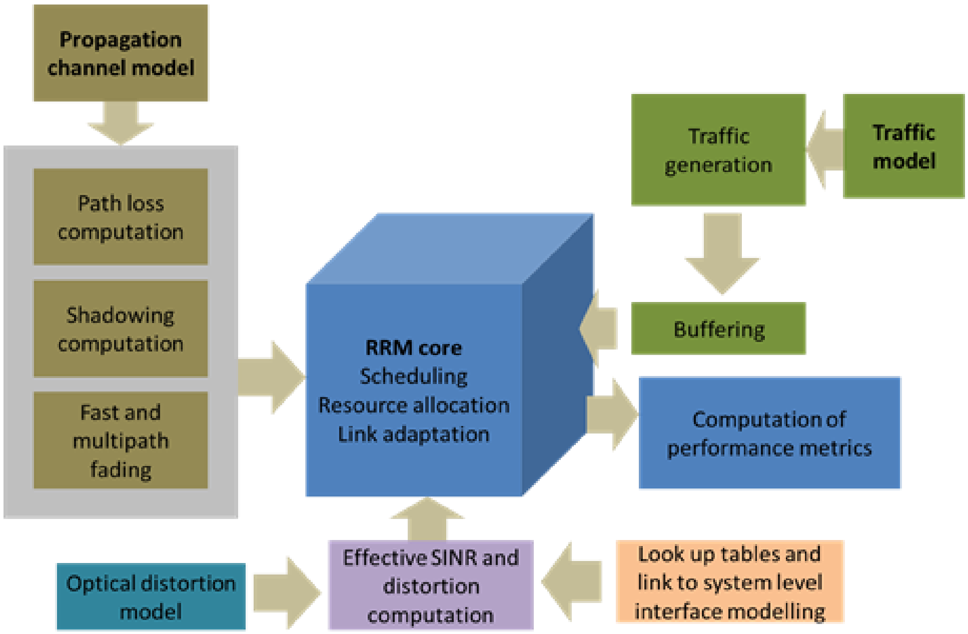


Figure 28 Block diagram of the system-level simulator.

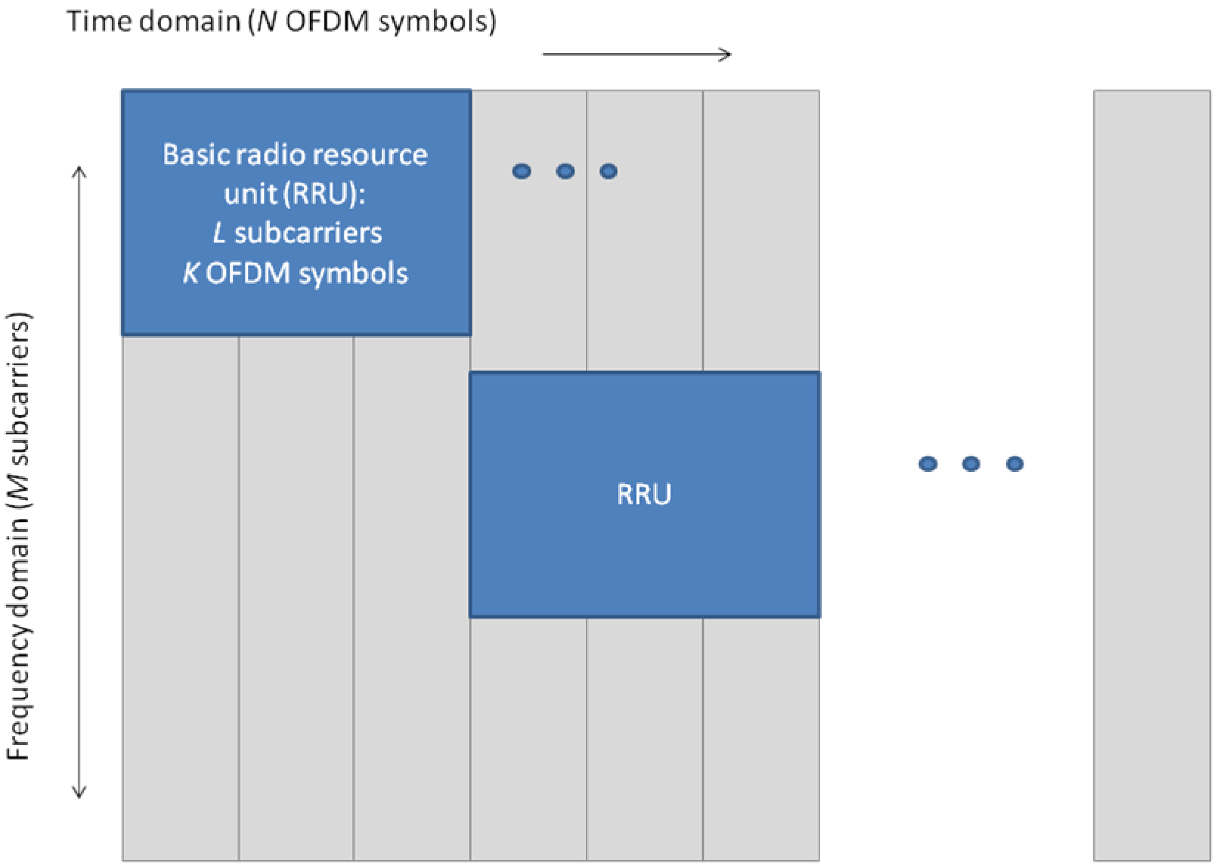


Figure 29 OFDMA-frame for system-level simulation

Table 1 System-level simulation parameters

Parameter	Explanation
Layout	Manhattan grid
Building width	20 metres
Channel model	WINNER B1
Antenna radiation pattern	Omnidirectional
BS antenna gain	15 dB
Mobile antenna gain	0 dB
Wall penetration loss	5 dB
Simulator mode	Combined snapshot
Traffic model	Full queue
N. freq. bands	2
Bandwidth	10 MHz
Subcarriers per symbol	1024
Data subcarriers per symbol	720
Frame length	20 symbols
Radio resource unit	10x720
Frame duration	5 ms
CQI feedback delay	10 ms
LSLI compression	EESM

The portfolio optimization problem provides as a result the fractions of the frequency bands of primary and secondary systems to be allocated to user traffic requests. If the fraction is equal to 1 that means that all radio resources of the band must be used. On the contrary, if the value is 0 then no resources must be used. Since we are allocating two different frequency bands, let us consider two possible combinations of values of return and risk. In the first case we consider that the band with the highest return has also the lowest value of risk. In the second case we will consider the opposite situation with the band having the highest value of return will also have the highest value of risk. We consider that these two cases cover the two extreme cases of the possible economic values for frequency bands. In all the results presented in this subsection, both neighbour networks experience the same economical values for primary and secondary bands.

Table 2 Investment returns

Case	Return primary	Risk primary	Return secondary	Risk secondary
1	High	Low	Low	High
2	High	High	Low	Low

Figure 30 presents the results for the throughput vs. Tx power-to-noise ratio for the case where the frequency with the highest value of return has the lower value of risk. The results displayed use different values of μ, which is the parameter that defines the balance between return and risk in the multi-objective optimization objective function. It can be observed that as the value of μ increases from 0 to 5 the total throughput increases too. However, it starts decreasing from μ=5 onwards. This means that the maximum throughput in the system is reached when we use a value of μ=5 approximately, which indicates that the balance between risk and return can also reach a maximum throughput if appropriately optimized.

Let us now consider the second case, where the frequency band with the higher return will now have the highest risk. The results for throughput are displayed in Figure 31. We can observe the same tendency as in Figure 30, reaching a maximum throughput using a value of approximately μ=4. However, we can observe that, in comparison with the results in Figure 30, the throughput is considerably lower. This suggests that it is convenient to have a frequency band with reliable economic return and low risk rather than with high risk.

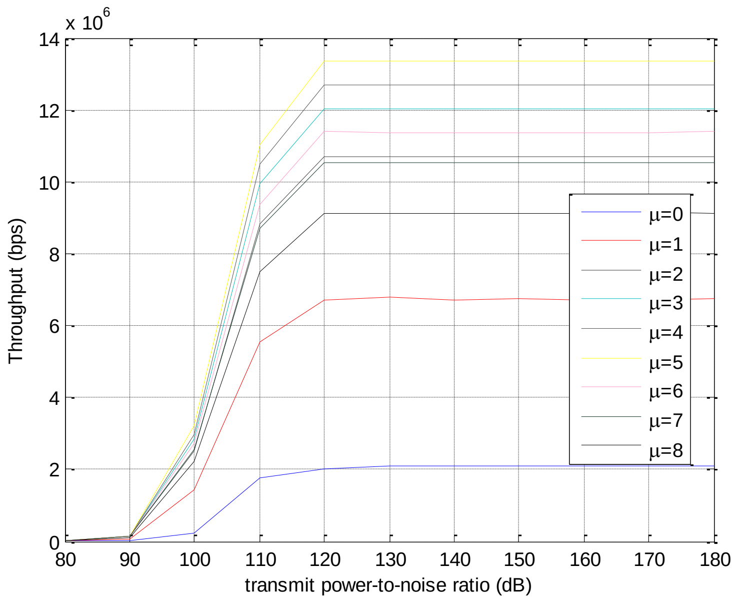


Figure 30 Throughput vs Tx power to noise ratio for different values of μ for case number 1.

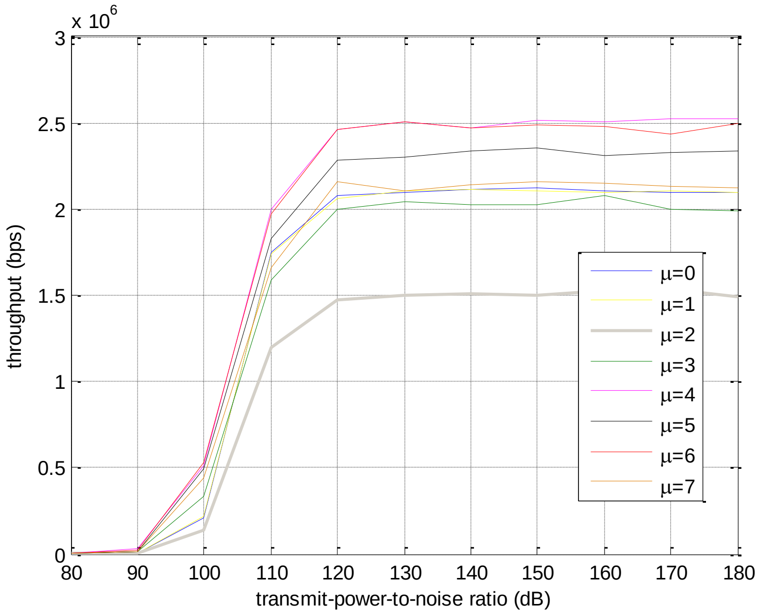


Figure 31 Throughput vs Tx power to noise ratio for different values of μ for case number 2.

In view of the dependency of the total throughput in the network on the value of μ, Figure 32 shows the results for throughput vs. μ for a fixed value of transmit power to noise ratio of 120 dB, which corresponds approximately to a transmission power of 30 dBm using a bandwidth of 10 MHz and a noise floor of -160dBm/Hz. We can observe that case 1 reaches a maximum more or less at a value of μ=5. By contrast, case number 2 exhibits a more flat shape with several maximum values, one at μ=4 and another at μ=6. As previously explained, case 2 never reaches a throughput as high as in case 1, which suggests that it is convenient to have a frequency band with reliable economical characteristics to maximize performance of cognitive radio systems. Optimum allocation in Figure 33 shows how case 2 uses less relative weight allocation in comparison with case 1. Another interesting issue can be observed in Figure 34, which displays the value of the multi-objective function. We can observe that the value of the objective function does not necessarily reflect the performance of the throughput, particularly in the points of maximum performance. This suggests that a further optimization for the value of μ to maximize throughput, in addition to the economical objective function, is required. In practice, it is possible to use an adaptive scheme to find the value of μ that maximizes the global throughput of the network.

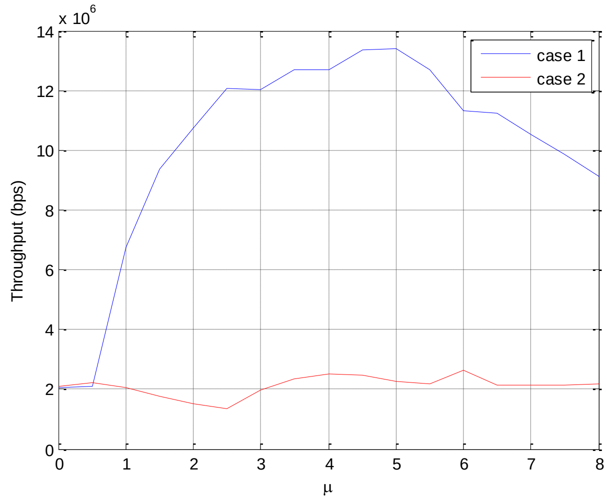


Figure 32 Throughput vs μ.

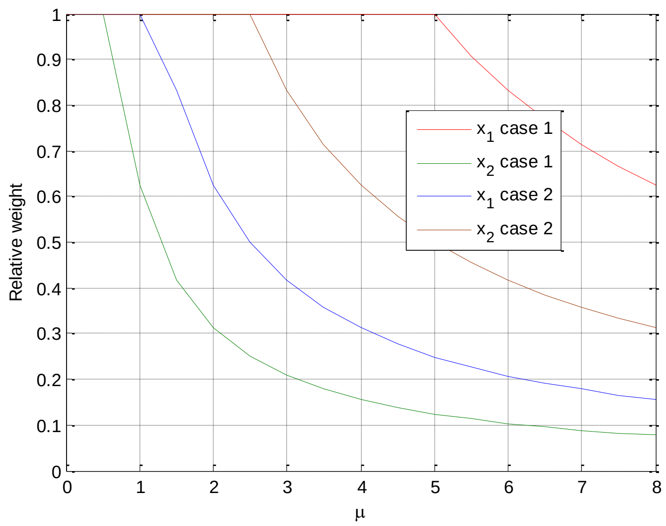


Figure 33 Optimum allocation weight.

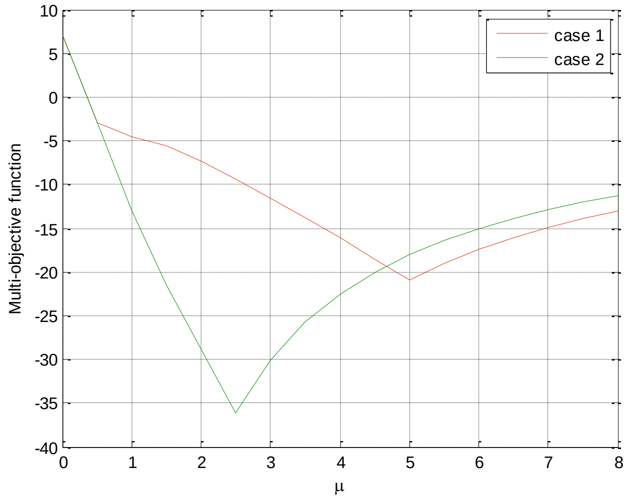


Figure 34 Value of the multi-objective function.

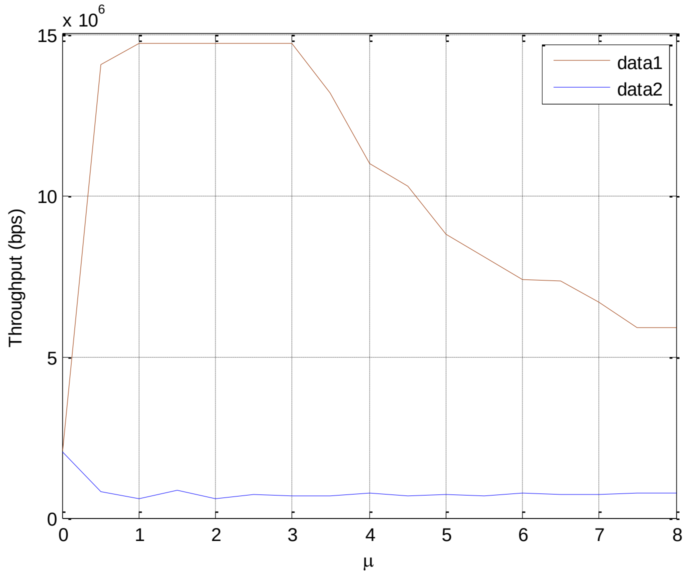


Figure 35 Throughput vs μ.

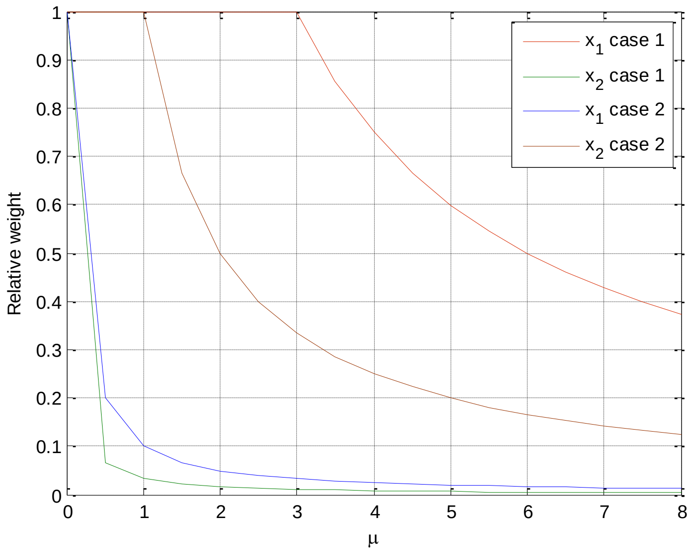


Figure 36 Optimum allocation weight.

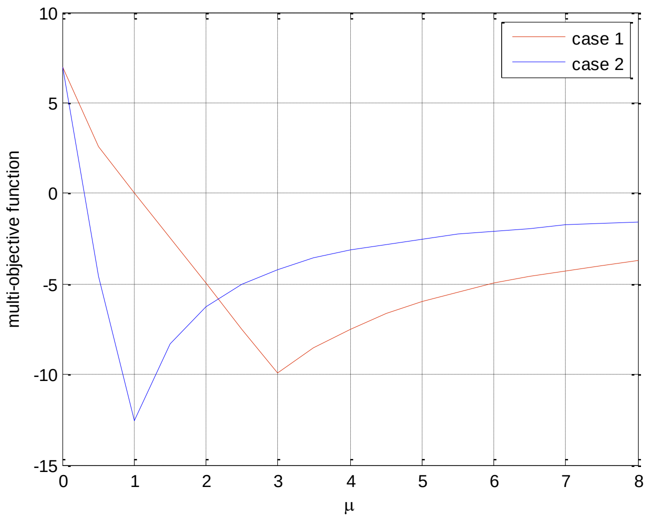


Figure 37 Value of the multi-objective function.

4.6 Modified multi-objective optimization problem

The portfolio optimization problem described in the previous subsections only considers economical factors. Therefore, it will always produce the same result under a fixed pricing scheme even if the network conditions are different. To solve this issue, the objective functions should be modified to take into account some relevant network parameters. In the proposed network scenario, each BS knows the average load conditions and the average interference level of the adjacent network thanks to a cognitive radio. Therefore, the portfolio objective functions can be modified to consider these two network parameters.

In order to incorporate network parameters in the calculation of these fractions of the spectrum, let us visualize the extreme conditions of operation between the two neighbour networks. If the two networks experience an extremely high load while the interference between the two is extremely high then both systems should only allocate users in the primary frequency. If the interference is medium and the load is not as high, then secondary transmission can be enabled. If the interference is almost zero then it doesn't matter the load conditions, allocation should be only based on economical factors. To reflect this scenario, we propose a modified multi-objective portfolio function for optimization. The modified composite return functions for primary and secondary systems, respectively, are as follows:

$$\bar{p}_{pri}^{(\text{mod})} = \bar{p}_{pri}\left(1+\frac{w_{pri}(1-L_{pri})}{1+\eta^{-1}}\right) \qquad \text{and} \qquad \bar{p}_{\sec}^{(\text{mod})} = \bar{p}_{\sec}\left(1+\frac{w_{\sec}(1-L_{\sec})-1}{1+\eta^{-1}}\right)$$

where L_{pri} and L_{sec} denote the normalized load of the primary and secondary systems, respectively, η is a measure of the interference level between primary and secondary systems, and w_{pri} and w_{sec} is a variable set to control the weight of network parameters in the optimization. Similarly, the composite return functions for primary and secondary systems, respectively, are as follows:

$$C_{pri}^{(\text{mod})} = C_{pri}\left(1+\frac{w_{pri}(L_{pri})^2}{1+\eta^{-1}}\right) \qquad \text{and} \qquad C_{\sec}^{(\text{mod})} = C_{\sec}\left(1+\frac{w_{\sec}(L_{\sec})^2-1}{1+\eta^{-1}}\right)$$

The optimization problem can now be reformulated as:

$$\begin{aligned} &\text{minimize} \quad -\bar{p}_{pri}^{(\text{mod})}x_{pri} - \bar{p}_{\sec}^{(\text{mod})}x_{\sec} + \mu\left(-C_{pri}^{(\text{mod})}x_{pri}^2 - C_{\sec}^{(\text{mod})}x_{\sec}^2\right) \\ &\text{s.t.} \qquad\qquad 0 < x_{pri}, x_{\sec} < 1 \end{aligned}$$

Alternatively, only interference can be used to modify the components of the objective function as follows:

$$\bar{p}_{pri}^{(\text{mod})} = \bar{p}_{pri}\left(w_{pri}+\frac{1}{1+\eta}\right) \qquad \bar{p}_{\sec}^{(\text{mod})} = \bar{p}_{\sec}\left(\frac{w_{\sec}}{1+\eta}\right)$$

$$C_{pri}^{(\text{mod})} = C_{pri}\left(w_{pri}+\frac{1}{1+\eta}\right) \qquad \text{and} \qquad C_{\sec}^{(\text{mod})} = C_{\sec}\left(\frac{w_{\sec}}{1+\eta^{-1}}\right)$$

Using this approach we can recalculate the multi-objective portfolio optimization problem under different network conditions. The results are displayed in the following figures. Figure 38 and Figure 39 display the optimum allocation weights and trade-off region, respectively, for a case with high interference and low load in the secondary system, while experiencing a high load in the primary system. It can be observed that more weight is given to the primary transmission, as the secondary transmission is limited by interference. The secondary transmission increases when the primary system has been all allocated ($x_{pri}=1$) and more risk is allowed. Figure 40 and Figure 41 display the optimum allocation weights and trade-off region, respectively, for a case with low interference and low load in the secondary system, while experiencing high load in the primary system. It can be observed that only a slightly larger weight is given to the primary transmission, as the secondary transmission is now not severely limited by interference or by load. Figure 42 and Figure 43 display the optimum allocation weights and trade-off region, respectively, for a case with low interference and high load in the secondary system, while experiencing high load in the primary system. It can be observed that more weight is given to the primary transmission, as the secondary transmission is now limited by load rather than interference. The secondary transmission increases when the primary system has been all allocated ($x_{pri}=1$) and more risk is allowed. Finally, Figure 44 and Figure 45 display the optimum allocation weights and trade-off region, respectively, for a case with high interference and high load in the secondary system, while experiencing high load in the primary system. It can be observed that more weight is mainly just given to the primary transmissions, as the secondary transmission is severely limited by both interference and load. The secondary transmission increases when the primary system has been all allocated ($x_{pri} = 1$) and more risk is allowed, but keeping a very small relative weight.

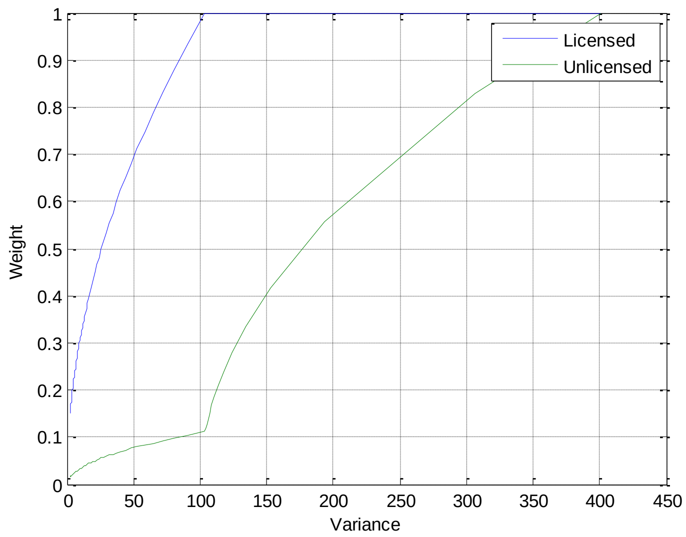


Figure 38 Optimum weights primary (licensed) and secondary (unlicensed) users in a scenario with high interference and low load in secondary system.

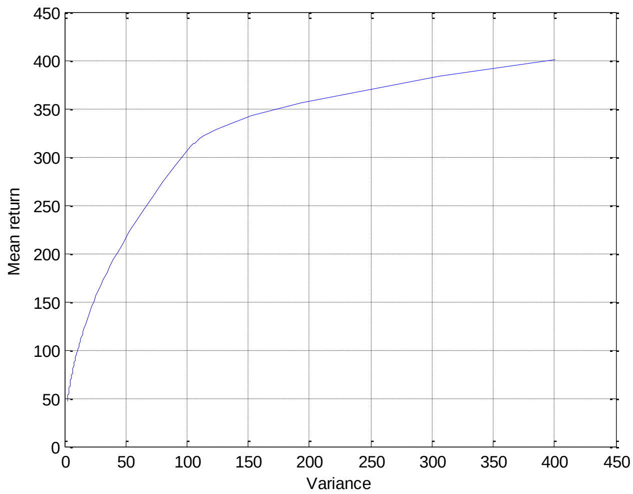


Figure 39 Pareto-optimal trade-off region in a scenario with high interference and low load in secondary system.

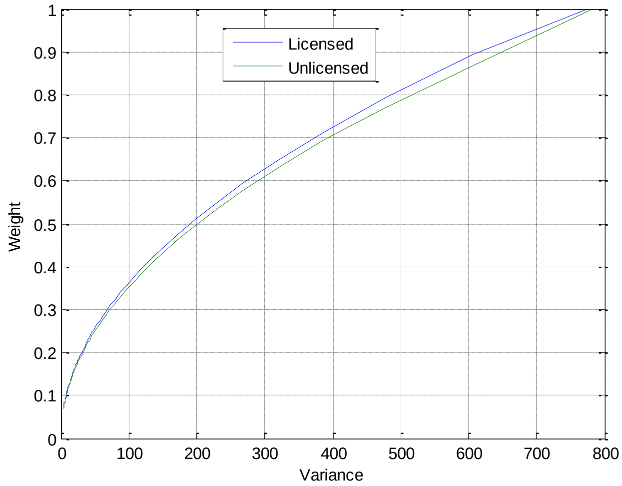


Figure 40 Optimum weights primary (licensed) and secondary (unlicensed) users in a scenario with low interference and low load in secondary system.

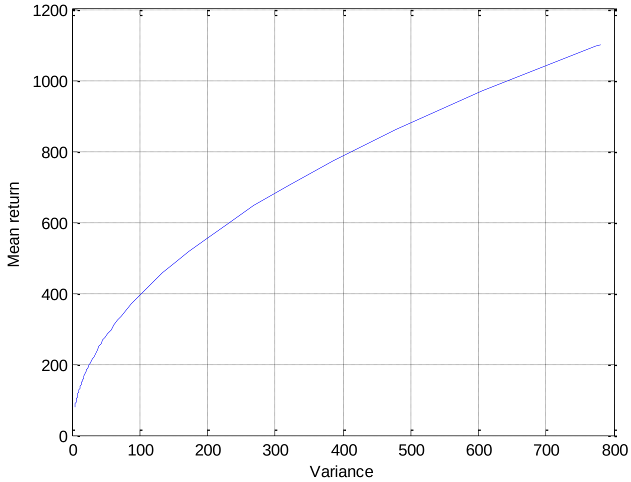


Figure 41 Pareto-optimal trade-off region in a scenario with low interference and low load in the secondary system.

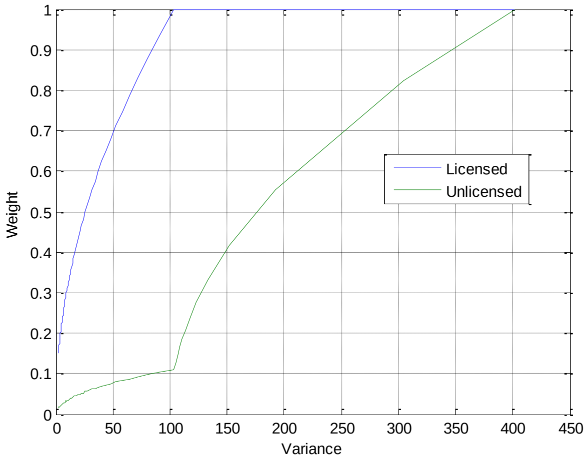


Figure 42 Optimum weights primary (licensed) and secondary (unlicensed) users in a scenario with low interference and high load in secondary system.

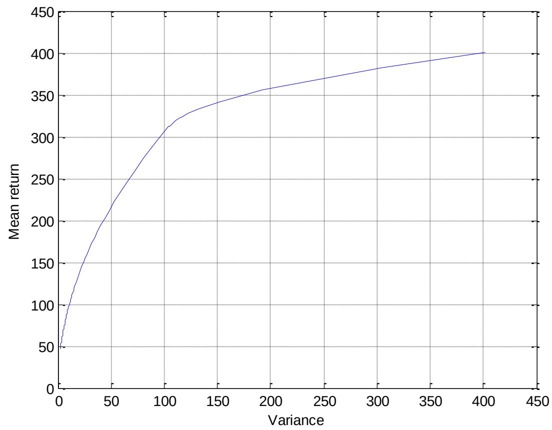


Figure 43 Pareto-optimal trade-off region in a scenario with low interference and high load in the secondary system.

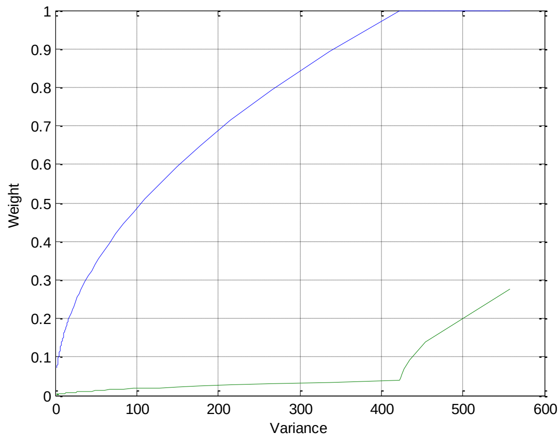


Figure 44 Optimum weights primary (licensed) and secondary (unlicensed) users in a scenario with high interference and high load in secondary system.

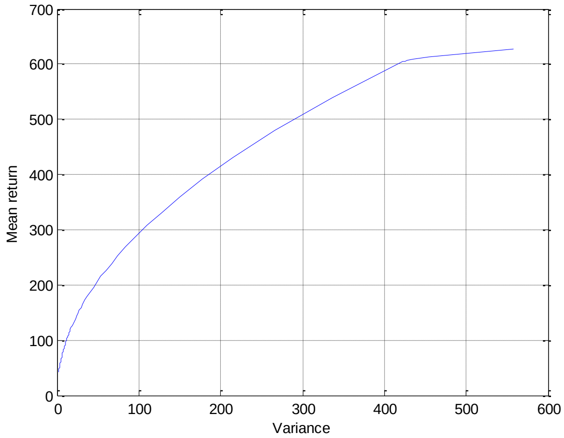


Figure 45 Pareto-optimal trade-off region in a scenario with high interference and high load in the secondary system.

Figure 46 displays of the total throughput achieved by the indoor system by allowing some users to opportunistically operate in the frequency of the adjacent network for various levels of risk. The indoor network has a relatively high load whereas the outdoor network has a relatively low load. Interference is present in the system but it is attenuated by the building penetration loss used in the simulation. The economic return of the secondary system $\overline{p}_{\text{sec}}$ is always lower than the economic return of the primary system $\overline{p}_{pri}$. The opposite is true with the economic risk for primary and secondary systems (C_{pri} and C_{sec}). The figure also includes the results for a conventional system without cognitive radio operation. It can be observed that the performance of the network can increase by allowing more risk (lower μ) or more interference towards the adjacent network. The results in the figure constitute the Pareto optimal solution of the optimization problem. These results mean that we can effectively increase performance and the economical return of the network simultaneously, while controlling or minimizing the risk and the interference to adjacent networks due to secondary user transmissions using the Pareto optimal solution of the system. The advantage of the modified multi-objective optimization approach presented in this subsection is that an appropriate selection of the multi-objective function parameters will produce optimum values of μ that are useful for a wide range of network conditions. This is in contrast to the approach presented in the previous subsection, where μ has to be adjusted to the network conditions. The disadvantage of the modified objective function is that we need to find suitable values for the parameters for the objective function that can fuse network and economic variables. A further improvement over this approach is presented in QoSMOS deliverable D6.7.

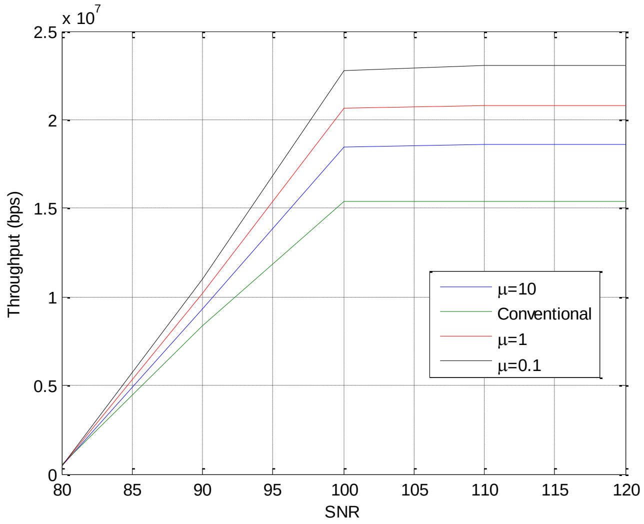


Figure 46 Throughput of primary+secondary users indoor network using modified multi-objective portfolio optimization

4.7 Conclusions for multi-objective portfolio optimization for cognitive radio

This section has demonstrated that multi-objective portfolio optimization schemes commonly used in the area of economics can be successfully used in frequency selection/aggregation problems for cognitive radio networks. In the testing scenario proposed with two WiMAX networks, it was observed that an optimum value of the parameter that measures the balance between return and risk in the multi-objective function can also lead to maximum throughput in the Pareto solution if one of the frequency bands (primary transmission) has higher return an also lower risk that the other frequency bands (secondary or opportunistic transmission). This also means that despite not using realistic values of return and risk, the results in this chapter give an indication of the range of values that provide optimum performance of the cognitive radio network. Two approaches with different levels of interaction between the RRM and the frequency selection entities were also proposed. In the first approach it is necessary to obtain different values for μ that maximize the throughout under different network conditions. In the second approach, only one value can be used that works for a variety of network conditions. The disadvantage of the latter one is that the objective function parameters need to be adjusted, which requires either simulation work or adaptive techniques. The main conclusion in this section is that technical and economical target functions can be conveniently reached in a cognitive radio network by means of the Pareto optimal solution of the multi-objective port-folio optimization problem given an appropriate balance between return and risk components and some economical parameters of the available frequency bands.

5 Evaluation and capacity of local spectrum control (LSPC)

5.1 Implementing spectrum mal-usage detection in spectrum-management framework

This chapter describes the management and system control issues in mobile cognitive radio (CR) systems exploiting "TV white space" (TVWS). The scenario is a cognitive management (CM) system operating in a geographical area with fixed TV broadcast transmitters, supplying fixed users (incumbents) and mobile and opportunistic users in the same UHF frequency band. The outcome of this work is the improvement of the spectral efficiency in the geographical area by exploiting the spatial opportunities of the mobile terminals. The map of the incumbent system is a static radio environment that can be characterized by signal-to-interference-noise (SINR) ratio, which determines the sufficient radio transmission power levels necessary to service the incumbent users. There are specific areas where the system allows for opportunistic usage of the same frequency band without disturbing the incumbents. In order to ensure the required level of the service quality, such a system contains a central control and management unit with cognitive resource and spectrum management (CM-RM and CM-SM) functionalities. CM is responsible for control of the opportunistic system, so guaranteeing operation of the incumbents by collecting, storing and processing information and performing decision processes on the spectrum usage of the opportunistic users. Applying a geographical database of the SINR guaranteeing incumbent's services, the gathered location information of the opportunistic users determines the possibility of the secondary usage of white spaces. This requires periodical position updates at the CM reported by the moving user. As the exact location is unknown between two consecutive reports the opportunistic user may violate SINR requirements of the primary system. We investigate the effect of location sampling on the system conformance measured by mal-usage of the available resources.

Moving cognitive radio in fixed transmitter area

In this approach a static arrangement of transmitters and receivers is assumed, as parts of the incumbent radio system. Important system parameters are geographical positions of equipment, transmitter frequency, power and receiver and transmitter antenna characteristics. The following equation gives the value of path attenuation on the radio channel where P_{out} is the power fed to the transmitter antenna and P_{in} is the effective power on the receiver antenna:

$$a_{path}^{[dB]} = 10\lg\frac{P_{out}}{P_{in}} \qquad (1)$$

For free space propagation the path attenuation can be expressed also as the function of path length D, the wavelength λ and the transmitter and receiver antenna gain, G_T and G_R:

$$a_{path}^{[dB]} = 20\lg\frac{4\pi D}{\lambda} - G_T^{[dB]} - G_R^{[dB]} = 32{,}44 + 20\lg f^{[MHz]} + 20\lg D^{[km]} \qquad (2)$$

Besides the useful signal usually there are interfering radio sources, that may cause disturbances if their frequency range is close to the frequency band of the useful signal. In the investigated topology there are multiple transmitters and we assume that a specific incumbent connects always to the transmitter that provides the best service in the area. In this system the signal of the other transmitters cause interference, therefore degrade the quality of service in the system.

For this topology, if we know the power of the transmitter, where the incumbent is connected, by using (1) the received power can be calculated. The power of the interfering signals can be taken into account as the sum of the received power from the remaining transmitters in the system.

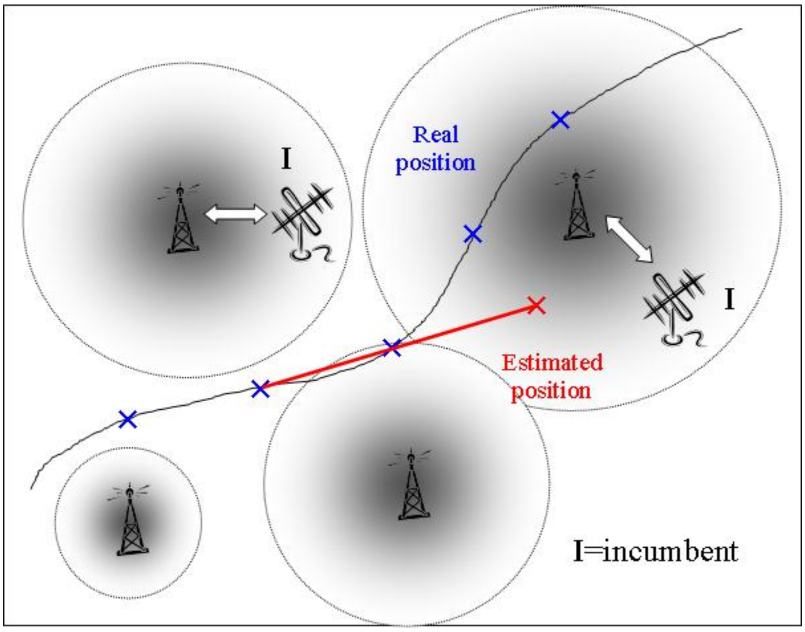


Figure 47 The modelling environment

The ratio of the useful and the interfering signal powers results the Signal to Interference Ratio (SIR). If the power of the system noise has significant value with a level comparable to the interferers, the Signal to Interference Noise Ratio (SINR) can be calculated, but in the current study we do not apply additional noise.

The quality of service for the incumbents is determined by the SIR/SINR at the receiver location. Depending on the applied modulation system, an SIR/SINR margin can be determined to ensure error-free operation.

In the investigated cognitive system opportunistic users may exist in the above topology. Their transmitters appearing as an interferer source, therefore its contribution to the incumbents SIR/SINR must be taken into account. It is the task of the cognitive manager to register the location of the opportunistic user(s) and determine the actual SIR/SINR map by using the geolocation database of the incumbent system topology. In our study the movement of the opportunistic users is also allowed, and this option significantly increases the tasks of the CM. The moving of the cognitive radio continuously changes the signal/interference relations in the system. However the cognitive radio may continuously monitor and report its own conditions and environmental data to the CM (spectrum sensing, position recording with GPS, etc.), the amount of the transmitted information should be limited in order to reduce the load of the service channel. Our goal is to find a solution to reduce the traffic of these service data and find an intelligent solution that can be implemented in the CM to estimate the unknown parameters of the moving opportunistic user between its two reporting period.

Heuristic movement simulation and estimation of the system parameters

To utilize the white spaces by the opportunistic users we have to know exactly, if in the current position the incumbent's signal quality will not be degraded. In order to achieve this goal, one of the main tasks is that the CM manages the physical architecture of the system as a geolocation database and provides the relevant information to the mobile terminals.

The applicability of a selected frequency depends on the SIR value of the current location and frequency. During our simulations we locate multiple transmitters in a hypothetical geographical area.

The incumbents are always serviced in the coverage area by their closest transmitter. The coverage area of the multiple transmitters designates an SIR map of the area. In the area where the SIR is below a defined threshold the incumbents cannot be served, thus for opportunistic users this area and frequency can be utilized. Contrarily, if the moving terminals approach the coverage area they has to cease transmitting, or change frequency, to avoid degradation of incumbent's signal quality.

Since communication with the cognitive radios has significant overhead, it is impossible to provide SIR information to CM with infinite resolution. This means, between two telemetry cycles estimated trajectory should be used to decide about future behaviour.

There are several possibilities for trajectory estimation, such as Markov Chain methods, Neural Networks etc., where training has a key role on the systems overall effectiveness. Without those training sets, we intend to demonstrate a simple heuristic based on the assumption, that the opportunistic user most likely goes straight forward, without direction change. Hence the heuristic simply gets the two last known positions (where former telemetry event was placed), and presumes the terminal will move with the same speed to the same direction.

The realised test doesn't use different frequencies, so the decision condition based on the heuristic is whether to transmit or not. The goal of the test is to evaluate how effective the heuristic is, by counting the mal-usage in the system. In our case *false positive* means the mobile terminal violates primary communication, while *false negative* means it doesn't use the communication channel, however it is possible.

We performed simulations over a hypothetical 35*35 km area with four low-power TV transmitters with commonly used frequencies. The parameters of the simulation (transmitter frequencies, power and antenna characteristics) are summarized in Table 1.

Table 1. Simulation parameters (4 transmitters)

	Rel. location x, y [km]	Frequency [MHz]	Power [dBm]	Antenna gain [dB]
1.	6, 6	498	40	15
2.	27, 12	610	40	15
3.	24, 27	722	40	15
4.	10, 20	818	40	15

During the calculations we suppose line of sight conditions, furthermore no shadowing or reflecting objects are considered. The opportunistic user was considered to move across this area with constant speed while its transmission is controlled by the CM. The geolocation database and the periodically reported position of the cognitive radio provide the required information to the control process. From the viewpoint of incumbents the transmitter of the cognitive radio appears in the area as an additional interferer source. If we suppose that the transmitter operates as an isotropic radiator, the power $P_{interferer}$ at d distance can be calculated from the transmitter power $P_{transmitter}$ with the following equation:

$$P_{\text{interferer}} = P_{transmitter}\left(\frac{\lambda}{4\pi d}\right)^2, \quad (3)$$

where λ is the wavelength. During the simulations we utilized a constant 27 dBm transmit power and 0.49 m wavelength (610 MHz).

The following figure depicts the SIR map in the transmitter's coverage area. The black path marks the route of the cognitive radios; the dots are the positions that are reported to CM. The time and geographical distances between the reported positions are depending on the speed of the object and the sampling period; in this simulation the number of this positions was 30. The white crosses are the estimated positions using the simple heuristics.

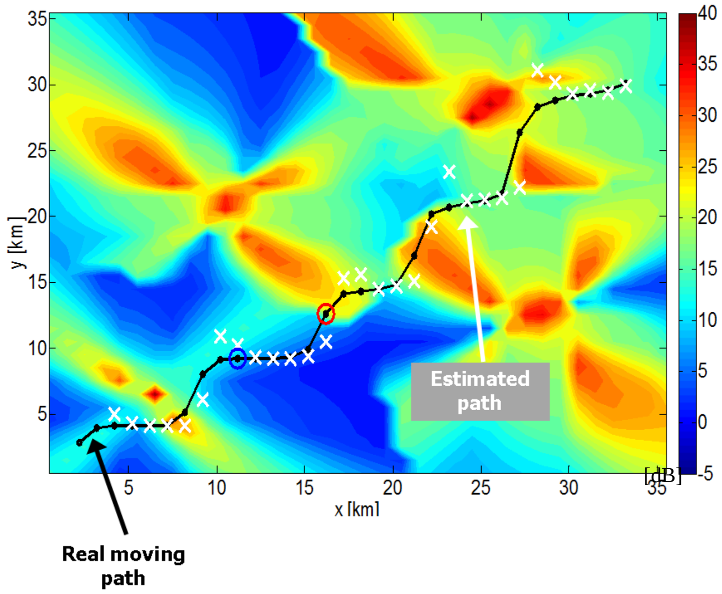


Figure 48 SIR map around the transmitters

We applied a 15 dB SIR threshold to enable or disable the opportunistic user; if the SIR is below this value at the investigated location the operation of the cognitive radio is allowed, otherwise disabled. In case of a false positive decision the point is marked with red circle, while the false negative case is denoted with blue. In the following figure the area around the two types of false cases is detailed:

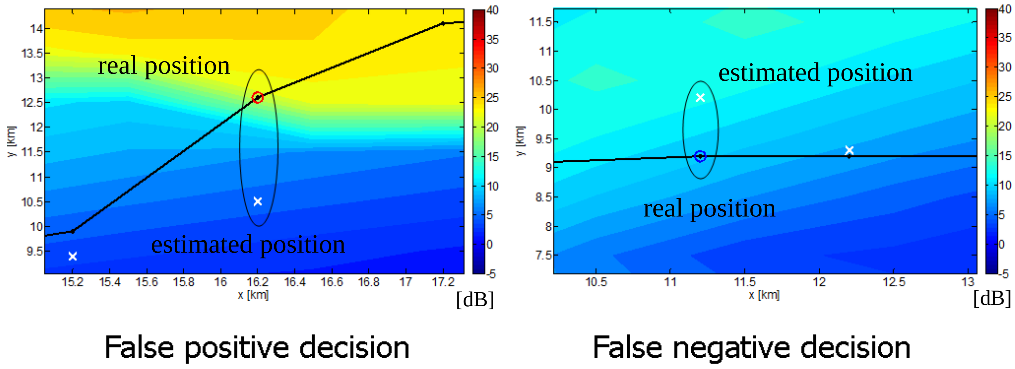


Figure 49 False positive and false negative cases

Further simulations with different moving path may result different number of false detections, however with the above settings this number is usually below 10 for the current simulation area.

5.1.1 Summary of contribution to implementation of QoSMOS spectrum management framework

The method is applicable to ensuring QoS for fixed incumbents in TVWS, and for serving the moving opportunistic users without radio interference. We have highlighted one of the roles of cognitive manager in CR environment, and a heuristic method for estimating the next position was developed.

6 Evaluation and capacity of spectrum analyser and selector

6.1 SAN functions

6.2 Statistical characterisation of metrics for spectrum selection

The CM-SM is responsible for building the spectrum portfolios based on a set of external constraints, such as regulatory and operator policies, and on spectrum sensing results. The CM-SM is responsible as well for the management of the spectrum portfolios, including cognitive spectrum management (decision making) methods to decide how to allocate portions of the available spectrum to the requesting entities or spectrum users (e.g., the base stations of a cellular system). The SSE is the module responsible for this particular task, which is in charge of selecting the most suitable spectrum to meet the communication requirements of the opportunistic users. The process of selecting adequate spectrum bands can be split into two different stages:

- *Spectrum analysis or characterisation*: In this stage, the characteristics of the spectrum holes (spectrum opportunities) in candidate bands are analysed and quantified based on certain set of predefined metrics. The purpose of the spectrum analysis/characterisation function is to quantify the properties of the identified spectrum opportunities so that the spectrum decision function can more easily determine the most convenient choice. Thus, it is essential to define metrics and parameters that can represent the features of the spectrum opportunities and its suitability to meet the needs and requirements of the opportunistic users. In the CM-SM reference model, this function is mainly carried out by the SAN module.
- *Spectrum decision*: Based on the characteristics of the identified spectrum opportunities, a decision on the most suitable spectrum is made in order to meet specific requirements. The function of spectrum decision needs to be supported by a set of appropriate metrics and parameters that quantify numerically the properties of the candidate bands and spectrum opportunities in order to ease the identification of the most convenient band to meet a particular set of requirements. In the CM-SM reference model, the SSE module is the main responsible for this task.

The focus of this section is on spectrum analysis/characterisation. A wide variety of metrics and parameters can be considered for spectrum selection, including RF metrics (i.e., metrics related to the RF properties of the channels where the spectrum opportunities are identified, such as the frequency, bandwidth, interference levels, emission limits, etc.) as well as primary activity metrics (i.e., metrics related to the spectrum occupancy patterns of the primary system in a certain spectrum band, such as the duty-cycle, expected occupancy/vacancy duration, arrival rate of idle/busy periods, etc.) [Lop13]. From the point of view of spectrum selection, primary activity metrics provide a more intuitive characterisation of the amount of vacant spectrum than can be found in a particular channel or band and therefore constitute the type of metrics analysed in this section.

The considered metrics can be divided into time-dimension metrics and frequency-dimension metrics. The metrics considered for the time domain include the channel duty-cycle, the duration of idle/busy periods, the inter-arrival time of spectrum opportunities, and the instantaneous channel availability. The metrics considered for the frequency domain include the band duty-cycle and the number of free channels in a spectrum band. For each of the considered metrics, appropriate statistical properties are derived analytically and possible ways of application in the design of spectrum selection algorithms are suggested and discussed as well.

6.2.1 Time-dimension metrics

Time-dimension metrics describe the properties of the spectrum opportunities in individual channels and are therefore well-suited to select a channel out of a set of candidates. The time-dimension metrics analysed in this study include the channel duty-cycle, the duration of idle/busy periods, the inter-arrival time of spectrum opportunities, and the instantaneous channel availability.

6.2.1.1 Channel duty-cycle

The duty-cycle of a channel can be defined as the probability (theoretical definition) or fraction of time (empirical definition) that the channel is busy (i.e., occupied by an incumbent transmission). This parameter provides an estimation of the overall fraction of time that the channel can be expected to be available for opportunistic transmissions. Thus, channels with lower duty-cycles are more attractive since they will provide more transmission opportunities. The average duty-cycle is a common metric widely used to characterise the occupancy level of a channel. However, in some spectrum bands where the channel load is not constant, the duty-cycle, when observed over shorter periods, may exhibit a time-dependent varying pattern [Lop11]. In such a case, the minimum and maximum duty-cycle values may be a convenient method to determine upper- and lower-bounds for the amount of opportunities that can be expected from a certain band or channel.

The duty-cycle of a channel *j* can be expressed as:

$$\Psi_j = \frac{E(T_1)}{E(T_0) + E(T_1)}$$

where T_0 and T_1 represent the duration of idle and busy periods of the channel, respectively, and $E(T_0)$ and $E(T_1)$, their expected (mean) values, respectively. Based on this metric, an opportunistic system can monitor the duration of idle and busy periods in a channel and compute the duty-cycle of each channel of interest based on the average (mean) of past observed channel durations. The SSE module can then select the channel with the lowest observed duty-cycle in order to maximise the expected long-term free capacity available for opportunistic use and therefore the overall system capacity. This spectrum selection strategy can be appropriate for services without stringent QoS requirements (e.g., best-effort type services), but may be questionable for services with more demanding requirements, for which spectrum selection approaches based on alternative metrics may be more convenient.

6.2.1.2 Duration of idle/busy periods

This metric quantifies the length of the idle (inactivity) or busy (activity) intervals of a channel, which can be useful to determine the time an opportunistic user will have to wait before accessing the channel (duration of busy periods) or how long the channel can be expected to be available for opportunistic transmission (duration of idle periods).

A common assumption widely employed in the existing literature is that period durations (both idle and busy) are exponentially distributed. Based on this assumption, the statistical characteristics of the occupancy/vacancy durations can be expressed as follows.

The PDF of the exponential distribution is:

$$f_{T_i}(T) = \lambda_i e^{-\lambda_i (T-\mu_i)}, \; T \geq \mu_i$$

where T_i indicates the period duration (i=0 for idle periods, i=1 for busy periods), $\mu_i > 0$ is the minimum period duration and $\lambda_i > 0$ is the rate parameter of the exponential distribution.

The corresponding CDF can be obtained by integrating the PDF, which yields:

$$F_{T_i}(T) = \int_{\mu_i}^{T} f_{T_i}(\tau) d\tau = 1 - e^{-\lambda_i (T-\mu_i)}, \; T \geq \mu_i$$

The mean and variance are given by:

$$E(T_i) = \int_{\mu_i}^{\infty} \tau \, f_{T_i}(\tau) d\tau = \mu_i + \frac{1}{\lambda_i}$$

$$Var(T_i) = E\left([T_i - \mu_i]^2\right) = \int_{\mu_i}^{\infty} (\tau - \mu_i)^2 \; f_{T_i}(\tau) d\tau = \frac{1}{\lambda_i^2}$$

These statistics can readily be computed based on past channel observations and employed for spectrum selection according to various criteria. For example, the Maximum Capacity-based Spectrum Decision (MCSD) and Minimum Variance-based Spectrum Decision (MVSD) principles proposed in [Lee11] can be implemented based on the mean and variance of this metric. The MCSD principle is aimed at maximising the total network throughput by selecting the spectrum that provides the highest expected capacity, which corresponds to the channel with the expected highest $E(T_0)$ and/or lowest $E(T_1)$. This criterion can be appropriate for best-effort applications where there are no strict QoS requirements. For delay- and jitter-sensitive applications (e.g., real-time traffic) a more convenient approach is the MVSD principle, which selects the spectrum that provides the minimum capacity variation, which corresponds to $V(T_0)$. This spectrum decision criterion is motivated by the fact that high capacity variations cause delays and jitter. Therefore, a channel where a sustained data rate can be maintained is more convenient for real-time traffic. The distribution of observed channel durations can also be employed to select those channels capable to provide specific QoS guarantees, for example the probability of interruption of an opportunistic transmission caused by the appearance of an incumbent user (i.e., a spectrum handover). The maximum or typical duration for an opportunistic transmission (or a statistically representative value thereof) can be estimated based on the characteristics of the traffic in the opportunistic network. The CDF $F_{T0}(T)$ can then be used to compute the probability that an idle period of such duration (or greater) is observed in the channel, and select a channel capable to provide idle periods of the required transmission duration with a certain probability (i.e., a certain guarantee). Other spectrum selection criteria can also be designed based on this metric.

6.2.1.3 Inter-arrival time of spectrum opportunities

This metric characterises the time elapsed between the beginning of two consecutive periods of the same type (i.e., idle or busy). The inter-arrival time of idle periods (i.e., spectrum opportunities) is an interesting parameter since it determines how often the channel can be expected to be available for opportunistic use. This is particularly important for delay-sensitive services, where each data packet has an associated deadline and the packet needs to be transmitted before the deadline in order to not be discarded. This means that delay-sensitive services need to have access to the channel with certain periodicity in order to avoid delayed (and potentially discarded) data transmissions. The inter-arrival time of spectrum opportunities plays a key role in the performance of this type of services.

The time elapsed between the beginning of two consecutive periods of the same type is given by $T_a = T_0 + T_1$. Assuming exponential distributions for T_0 and T_1, the statistical characteristics of this metric can be derived as follows.

The PDF of the sum of two independent random variables is given by the convolution of their PDFs:

$$f_{T_a}(T) = \int_{\tau} f_{T_0}(T-\tau)\, f_{T_1}(\tau) d\tau = \frac{\lambda_0 \lambda_1}{\lambda_0 - \lambda_1}\left[e^{-\lambda_1(T-\mu_0-\mu_1)} - e^{-\lambda_0(T-\mu_0-\mu_1)}\right],\; T \geq T_0 + T_1$$

The corresponding CDF can be obtained by integrating the PDF, which yields:

$$F_{T_a}(T) = \int_{\mu_0+\mu_1}^{T} f_{T_a}(\tau) d\tau = \frac{\lambda_0 \lambda_1}{\lambda_0 - \lambda_1}\left[\frac{1-e^{-\lambda_1(T-\mu_0-\mu_1)}}{\lambda_1} - \frac{1-e^{-\lambda_0(T-\mu_0-\mu_1)}}{\lambda_0}\right],\; T \geq T_0 + T_1$$

The mean and variance are given by:

$$E(T_a) = \int_{\mu_0+\mu_1}^{\infty} \tau \; f_{T_a}(\tau) d\tau = \mu_0 + \frac{1}{\lambda_0} + \mu_1 + \frac{1}{\lambda_1} = E(T_0) + E(T_1)$$

$$Var(T_a) = E\left([T_a - \mu_a]^2\right) = \int_{\mu_0+\mu_1}^{\infty} (\tau - \mu_0 - \mu_1)^2 \; f_{T_a}(\tau) d\tau = \frac{1}{\lambda_0^2} + \frac{1}{\lambda_1^2} = Var(T_0) + Var(T_1)$$

These statistics of the inter-arrival time of spectrum opportunities can be used in different ways to select a channel based on particular service requirements, in order to match the channel characteristics to the traffic needs. For example, the minimum or typical time interval between the arrival of two consecutive data packets (or a statistically representative value thereof) can be estimated based on the characteristics of the traffic in the opportunistic network. The CDF $F_{Ta}(T)$ can then be used to compute the probability that the inter-arrival time of spectrum opportunities in the channel does not exceed such a value, and select a channel capable to provide statistical guarantees that data packets will not be discarded as a result of delayed transmissions. If the time interval between the arrival of two consecutive data packets is constant for a particular type of service, then the selection criterion may be to select the channel for which $E(T_a)$ more closely resembles such a value and/or $Var(T_a)$ is minimum (in order to minimise the variation in the inter-arrival time of spectrum opportunities). Other effective spectrum selection criteria can also be designed based on this metric.

6.2.1.4 Instantaneous channel availability

The long-term availability of a channel can be expressed in terms of the duty-cycle, as discussed in Section 6.2.1.1. Following a similar principle, an *instantaneous* channel availability parameter can be defined by taking the instantaneous values of the period durations instead of their expectations:

$$\xi = \frac{T_0}{T_0 + T_1}$$

This metric can be used to characterise the availability of a channel over short time intervals and its temporal fluctuations.

The statistics of ξ can be derived under the assumption of exponentially distributed periods. The details of the derivation are omitted due to the complexity and length of the process. The final results are provided below.

The PDF of ξ is given by:

$$f_{\xi}(T) = \frac{1}{T^2} \lambda_0 \lambda_1 e^{\lambda_0 \mu_0} e^{\lambda_1 \mu_1} e^{-\lambda_0 \Xi \zeta} \left(\frac{\zeta}{\lambda_0 \Xi} + \frac{1}{(\lambda_0 \Xi)^2} \right)$$

where

$$\zeta = \max\left(\mu_0, \frac{\mu_1 T}{1-T}\right) = \begin{cases} \mu_0, \; if \quad T \le \dfrac{\mu_0}{\mu_0 + \mu_1} \\ \dfrac{\mu_1 T}{1-T}, \; if \quad T \ge \dfrac{\mu_0}{\mu_0 + \mu_1} \end{cases}$$

and

$$\Xi = 1 + \frac{\lambda_1}{\lambda_0}\left[\frac{1}{T} - 1\right].$$

The corresponding CDF can be obtained by integrating the PDF, and is given by:

$$F_{\xi}(T)=\begin{cases} e^{\lambda_0\mu_0}e^{\lambda_1\mu_1}\dfrac{e^{-\lambda_0\mu_0\Xi}}{\Xi}, & if \quad T\le\dfrac{\mu_0}{\mu_0+\mu_1} \\ 1-\dfrac{1}{\Omega}+e^{\lambda_0\mu_0}e^{\lambda_1\mu_1}\left[\dfrac{e^{-\lambda_1\mu_1\Omega}}{\Omega}-\dfrac{e^{-\lambda_1\mu_1\Phi}}{\Phi}\right], & if \quad T\ge\dfrac{\mu_0}{\mu_0+\mu_1} \end{cases}$$

where

$$\Omega=1+\frac{\lambda_0\mu_0}{\lambda_1\mu_1}$$

and

$$\Phi=1+\frac{\lambda_0}{\lambda_1}\left[\frac{1}{T}-1\right]^{-1}.$$

The mean and variance cannot be derived in a straightforward manner for this metric, but can be approximated by making use of the following relations [Blu01]:

$$E(\xi)=E\left(\frac{T_0}{T_0+T_1}\right)\approx\frac{E(T_0)}{E(T_0)+E(T_1)}\left[1+\frac{Var(T_0)+Var(T_1)}{[E(T_0)+E(T_1)]^2}\right]$$

$$Var(\xi)=Var\left(\frac{T_0}{T_0+T_1}\right)\approx\left(\frac{E(T_0)}{E(T_0)+E(T_1)}\right)^2\left[\frac{Var(T_0)}{[E(T_0)]^2}+\frac{Var(T_0)+Var(T_1)}{[E(T_0)+E(T_1)]^2}\right]$$

These statistics can also be employed to implement the MCSD and MVSD principles proposed in [Lee11] and discussed in Section 6.2.1.2. Based on the instantaneous channel availability metric ξ, the MCSD principle is equivalent to select the channel that maximises $E(\xi)$, while the MVSD principle is equivalent to select the channel that minimises $Var(\xi)$. Other effective spectrum selection criteria can also be designed based on this metric and its statistics.

6.2.2 Frequency-dimension metrics

As opposed to time-dimension metrics, which describe the properties of the spectrum opportunities in individual channels, frequency-dimension metrics describe the characteristics of the spectrum opportunities in a certain band aggregated over the set of channels integrating the band. Therefore, frequency-dimension metrics are well-suited to select a band (instead of a channel) out of a set of potential candidate spectrum bands. Therefore, while time-dimension metrics can be applied in the selection of user-to-channel assignments, frequency-dimension metrics find their field of application in the selection of networks-to-bands assignments.

6.2.2.1 Band duty-cycle

The duty-cycle of a spectrum band can be defined as the probability (theoretical definition) or fraction of time (empirical definition) that the band is busy (i.e., occupied by incumbent transmissions) over all the channels integrating the band. This metric provides, in a single numerical value, an estimation of the availability of the band for opportunistic transmissions. Thus, bands with lower duty-cycles are more attractive since they will provide more transmission opportunities.

The band duty-cycle can be computed by averaging the duty-cycle of all the channels within the band:

$$\overline{\Psi}=\frac{1}{N}\sum_{j=1}^{N}\Psi_j$$

where N is the number of channels in the band and Ψ_j is the duty-cycle of the j-th channel, $j = 1, 2, \ldots, N$, as defined in Section 6.2.1.1. Based on this metric, an opportunistic system can monitor the period durations in various bands and compute the average duty-cycle for each band of interest based on the duty-cycle of each individual channel. The SSE module can then target the spectrum band with the lowest observed duty-cycle (weighted by bandwidth or other relevant parameters) in order to maximise the overall long-term free capacity available to the opportunistic system as a whole.

6.2.2.2 Number of free channels

In the frequency dimension, the spectrum opportunities of a spectrum band at any time instant are the idle channels not being used by the incumbent users at that moment. Assuming that the channels of the spectrum band have the same RF bandwidth, the problem of characterising the spectrum opportunities in the frequency domain reduces to the characterisation of the total number of free channels available at any time instant which are not being used by the incumbent system. The idle/busy state of every channel in a band can be modelled as a Bernoulli (binary) random variable, where the probability of "success" (i.e., probability to be idle) for each channel is known but not necessarily the same. Under this model, the total number of idle channels in the band is the number of "successes" in a set of independent Bernoulli experiments, each of which yields "success" with a different probability according to the duty-cycle of the corresponding channel. In this scenario, the number of free channels is given by the Poisson's binomial distribution (see equation 7 of [Wan93]). This distribution lacks of an analytically tractable closed-form expression, which requires the use of approximations.

If all the channels of the spectrum band had the same duty-cycle, then the number of free channels would follow a binomial distribution where the success probability is given by $1-\Psi$, with Ψ being the duty-cycle for each channel. The binomial distribution, however, can still be used as an approximation to the number of free channels in a band by computing the average band duty-cycle and using it as the success probability parameter for the binomial distribution. As show in Figure 50, this approach provides an accurate model for the real distribution of free channels in the band under various scenarios, outperforming other candidate distributions that were considered and analysed as well.

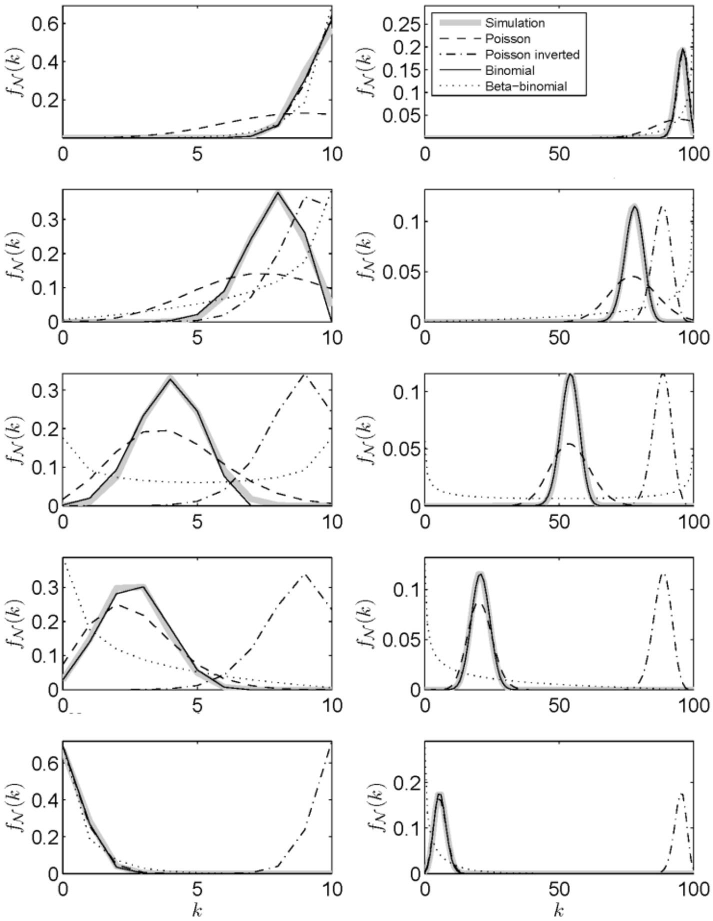


Figure 50 Approximations to the number of free channels in a spectrum band

The PMF of the binomial distribution is given by:

$$f_N(k) = \binom{N}{k}\left(1-\overline{\Psi}\right)^k\left(\overline{\Psi}\right)^{N-k}$$

where N is the total number of channels of the band and $\overline{\Psi}$ is the band duty-cycle.

The CDF of the binomial distribution is given by:

$$F_N(k) = I_{\overline{\Psi}}\left(N-k, 1+k\right)$$

where $I(\cdot,\cdot)$ is the regularised incomplete beta function.

The mean and variance are given by:

$$E(k) = N\left(1-\overline{\Psi}\right)$$

$$Var(k) = N\overline{\Psi}\left(1-\overline{\Psi}\right)$$

Based on this metric, an opportunistic system can monitor the period durations in various bands and compute the average duty-cycle for each band of interest based on the duty-cycle of each individual channel. Based on the band duty-cycle, the statistics shown above can readily be obtained. The SSE module can then select a spectrum band based on the aggregated traffic needs of the opportunistic system. Based on the properties of the aggregated network traffic, the opportunistic system can estimate the number of opportunistic users that are expected to be simultaneously active in the system (i.e., with some data to transmit). The CDF $F_N(k)$ can then be used to compute the probability that a minimum number of free channels is available and then select a band capable to provide the capacity required by the secondary system. Based on the mean $E(k)$ and variance $Var(k)$, it is also possible to select the spectrum band with the highest expected number of free channels (to maximise the overall system capacity) or with the minimum variance in the number of free channels (to ensure low capacity fluctuations and therefore the capability to support a sustained amount of users/traffic). Other effective spectrum selection criteria can also be designed based on this metric and its statistics.

6.2.3 Estimation of the metrics based on empirical observations

The statistics of the time- and frequency-dimension metrics analysed above can be estimated by only computing the minimum μ_i, mean $E(T_i)$, and variance $Var(T_i)$ of the idle and busy period durations.

Regarding the time-dimension metrics, the channel duty-cycle only requires the values of $E(T_i)$ to be known, while the rest of metrics (i.e., duration of idle/busy periods, inter-arrival time of spectrum opportunities and instantaneous channel availability) require the parameters of the assumed exponential distribution (μ_i and λ_i) to be known. However, these parameters can be estimated from the observed period durations in different ways based on the method of moments. For example, the value of μ_i can be obtained as the minimum period duration observed in the channel and the rate parameter can then be estimated either as $\lambda_i = [E(T_i) - \mu_i]^{-1}$ or $\lambda_i = [Var(T_i)]^{-1/2}$. Alternatively, both parameters can be estimated from the mean and variance as $\mu_i = E(T_i) - [Var(T_i)]^{1/2}$ and $\lambda_i = [E(T_i) - \mu_i]^{-1} = [Var(T_i)]^{-1/2}$. These estimates of μ_i and λ_i can then be employed to estimate the statistics of the metrics.

Regarding the frequency-dimension metrics, only the number of channels of the band and the band duty-cycle need to be known. The former can be obtained from regulatory databases while the latter can readily be obtained based on the duty-cycle of individual channels as described in Section 6.2.2.1.

Therefore, based only on the observation of the duration of idle and busy periods, an opportunistic network can estimate the abovementioned metrics and their statistics and apply appropriate spectrum selection methods based on such metrics.

6.2.4 Summary

This section has studied a wide variety of metrics for spectrum selection and their statistical properties. Metrics for both the time-dimension (aimed at facilitating the design of channel selection solutions) and frequency-dimension (aimed at easing the development of band selection solutions) have been analysed. The metrics considered for the time domain include the channel duty-cycle, the duration of idle/busy periods, the inter-arrival time of spectrum opportunities, and the instantaneous channel availability. The metrics considered for the frequency domain include the band duty-cycle and the number of free channels in a spectrum band. For each of the considered metrics, appropriate statistical properties have been derived analytically and possible ways of application in the design of spectrum selection algorithms have been suggested and discussed as well.

6.3 Evaluation of spectrum selection functions based on incumbent user statistics from spectrum sensing

This section will present the evaluation of four spectrum selection (SSE) functions that aim to increase performance and QoS in the opportunistic system while protecting incumbent users;

1. *SSE-Power* selects the channel with lowest received signal strength from the sensing function. This is considered a basic SSE function used to benchmark the other SSE functions.

2. *SSE-Distance* aims to improve performance when the mobile cognitive radio terminals are moving by selecting the channel with longest distance to the incumbent users. The distance is predicted by the SAN function by using sensing results from the cognitive radio devices together with their location.
3. *SSE-OnOff* function aims to improve QoS by calculating the probability that a channel will be available. This is done by the SAN function by using sensing results from the cognitive radio devices to store the ON and OFF periods for the incumbent users.
4. *SSE-Hybrid* combines *SSE-Distance* and *SSE-OnOff* to find the optimal channel.

The QoSMOS scenario considered is cellular extension in white spaces with mobile opportunistic users (OUs) roaming around in the cell. Performance is evaluated by measuring the metrics throughput, delay, packet loss and signal-to-interference ratio. The incumbent users considered are wireless microphones (WMs), and the impact on these is evaluated by measuring the metrics carrier-to-noise ratio (C/I) and the percentage of time the C/I level drops below a critical level.

To evaluate these SSE functions, the cognitive radio system standard IEEE 802.22 [ieee80222] for access in TV White Spaces (TVWS) is implemented in the NS-2 simulator. This standard supports cognitive radio terminals at fixed locations, but mobile terminals are implemented in order to study a mobile cognitive radio system as targeted in QoSMOS. The IEEE 802.22 simulator implementation is presented in Appendix A.

Three hypotheses are defined for the evaluation of the SSE functions:

H.1 *SSE-Distance* will improve performance when OUs are mobile.

H.2 *SSE-OnOff* will improve performance when the density and activity level of WMs is high.

H.3 *SSE-Hybrid* will improve performance both when OUs are mobile and the WM activity and density is high.

In the following, we will first present the network model before we present the proposed SSE functions. Next we present the performance metrics before and finally the scenario considered and performance evaluation of the SSE functions.

6.3.1 Network Model

We consider an IEEE 802.22 system limited to one Base Station (BS) and a set of OUs. An example network setup is illustrated in Figure 51 with *N* OUs and *M* WM Tx-Rx pairs. The small oval illustrates the coverage area for WM_{1a} and the big oval the coverage area for the IEEE 802.22 BS. The simulator supports all channels in the UHF band, but a subset of *M* channels are selected for operation in the studied simulation scenarios denoted $F_1 \ldots F_M$. A personal computer (PC) connected via Ethernet to the BS will establish links to the OUs and run traffic models.

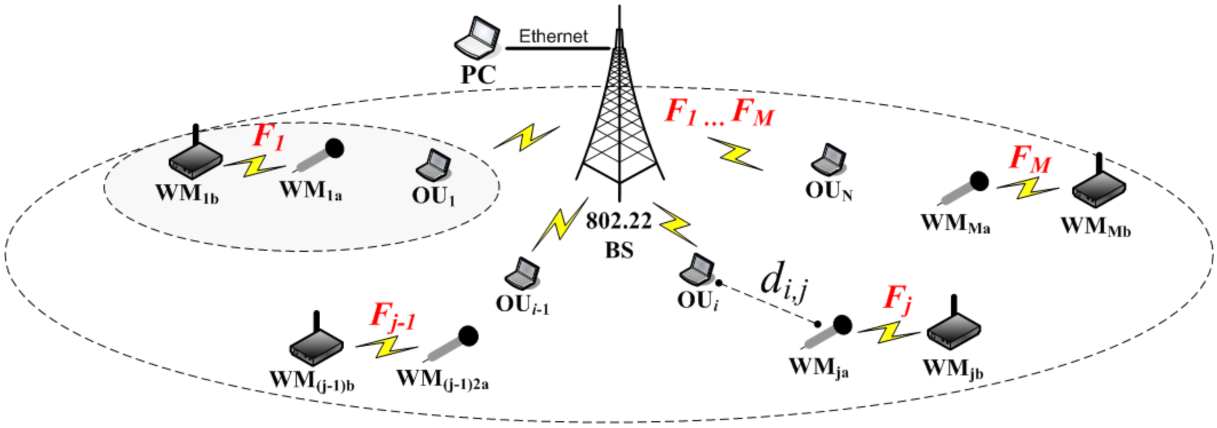


Figure 51 Network Model for evaluation of SSE functions

In this study we assume that TV transmissions are known from the database and focus only on WMs as the incumbent users. WM activity will be detected only by sensing techniques with no support from beacon protocol like IEEE 802.22.1 [ieee802221]. The number of available channels and WMs will be equal in the simulations, where one WM will appear on one single channel. When the WM appear on that channel, the IEEE 802.22 system must vacate the channel and switch to a channel not occupied by any WM. Finally, the IEEE 802.22 self-coexistence protocol is not used since a single cell is considered.

The IEEE 802.22 system communicates with the QoSMOS CM-SM to obtain the selected channel. Sensing results are reported from the OUs to the IEEE 802.22 BS which then sends all sensing results to the SAN (Spectrum Analyser) function in the CM-SM. The SAN function builds the required statistics and sends these statistics to the SSE function. The SSE function then implements an algorithm used to select a channel based on statistics from the SAN and information from the LPFR (Local Portfolio Repository) function. The SSE functions proposed and evaluated in this section will use an optimization algorithm to select the best channel.

6.3.1.1 IEEE 802.22 Traffic Model

The traffic model constant bit rate (CBR) is used for the IEEE 802.22 system. CBR will run over UDP where different traffic rates will be simulated by constantly transmitting UDP packets to each OU of size 1500 Bytes. CBR will use the best effort (BE) QoS traffic profile in IEEE 802.22 which provides fairness between the OUs using the BE profile.

6.3.1.2 Wireless Microphone Traffic Model

A WM pair in the simulator will be two WMs communicating with a typical distance of 100m. When the WM becomes active, its traffic pattern is characterized by a 100% duty-cycle (irrespective whether someone is speaking to the microphone or not) until the WM disappears. Since WMs typically are present in venues such as churches, schools and concert halls, and since these often appear on a channel with specific time intervals (e.g. each evening), we model their appearance pattern according to the ON/OFF model at specific locations. We are not aware of any work that give statistics on WM appearance, hence we assume that all WMs generate new connections according to the negative exponential distribution for the inter-arrival time $1/\lambda_w$ and burst departure time $1/\mu_w$ which is common in wireless communications [Wellens09]. The WM implementation in the simulator is presented in Appendix A.5.

6.3.2 SSE Functions

Four different SSE functions are proposed as described below. These do not replace the spectrum management functionality specified in the IEEE 802.22 standard, but will be complementary and coexist to enhance performance by considering statistics calculated over longer time periods.

6.3.2.1 SSE-Power

The *SSE-Power* function is a basic algorithm that will be used to benchmark the other three SSE algorithms. SSE-Power selects the channel where the spectrum sensor has detected lowest signal from WMs.

The SAN function receives sensing results $r_{i,j}$ from OU_i on channel j and sends this information to the SSE. The *SSE-Power* function then selects the optimal channel ch_{Power} based on the following criteria:

$$ch_{Power} = \min_j \left\{ \max_i \left(r_{i,j} \right) \right\}, \tag{4}$$

$$\text{subject to:} \, r_{i,j} < \delta, j \in [0, M], i \in [0, N] \tag{5}$$

where N is the total number of OUs connected to the BS, M the total number of channels to be sensed and δ the sensing detection threshold (-107dBm over 200kHz in the simulator).

6.3.2.2 SSE-Distance

The second spectrum selection algorithm is one that enhances QoS in presence of mobile cognitive radio terminals. It uses sensing to predict the distance to WMs and then to proactively select the channel with farthest distance from the closest OU to WM on that channel (e.g. each 5th second, which is used in the simulator).

Upon receiving sensing results from the IEEE 802.22 system, the SAN finds the distance $d_{i,j}$ for each OU_i to the closets WM on each channel j. Note that the SAN function is able to predict the location of the WM based on received sensing results from the cognitive radio devices. This information is forwarded to the SSE-Distance function which uses it to select the optimal channel $ch_{Distance}$ based on the following criteria:

$$ch_{Distance} = \max_j \left\{ \min_i \left(d_{i,j} \right) \right\}, \tag{6}$$

$$\text{subject to:} \, r_{i,j} < \delta, j \in [0, M], i \in [0, N] \tag{7}$$

where N is the total number of OUs connected to the BS, M the total number of channels to be sensed and δ the sensing detection threshold.

SSE-Distance and *SSE-Power* will give similar results in many cases. The difference is that the *SSE-Distance* uses historical sensing data from several OUs to predict the location of the WMs, hence it will give more reliable prediction of the WM location.

Furthermore, the *SSE-Distance* function can use information about WM locations to optimize the sensing scheme used. If this is used, it will be referred to with "+"-symbol as *SSE-Distance+*. We focus on the two-stage sensing function in IEEE 802.22, where at the coarse sensing stage (first stage) an energy detector is used for frequent and short sensing duration T_c=1ms. If coarse sensing detects a WM signal it switches to the fine sensing stage (second stage) that uses a more detailed WM detection process for a longer duration T_s=30ms. The coarse sensing period, i.e. the interval between two coarse sensing stages, is given as $T_P = n_c \cdot T_F$, where T_F=10ms is the OFDMA frame length in IEEE 802.22. The channel detection time (CDT) which is set to 2 seconds by FCC gives an upper bound on n_c, hence we have $1 \leq n_c \leq CDT$.

As n_c and hence T_P increases, the sensing overhead will reduce. Note that coarse sensing senses at the end of the OFDMA frame, which will cause a reduction in data transmitted uplink. However, since the number of coarse sensing stages is reduced, the number of false alarms will also decrease reducing the number of fine sensing stages. This will reduce overhead.

We want to set T_P to protect the WM and the IEEE 802.22 nodes from interfering with each other. Hence, when distance $d_{i,j}$ between the OU i and WM j on channel j is low, T_P should be set to a low value such that WMs can be detected faster. This will reduce interference. If the $d_{i,j}$ is high such that the OU is outside detection range, T_P should be set high to increase throughput in the IEEE 802.22 network. To find the WM detection range for the OU, we set up the equation for received transmit power RX_{OU} for the OU:

$$RX_{OU} = TX_{WM} - PL + G_{TX} + G_{RX} \tag{8}$$

where $TX_{WM} = 17$dBm (50mW) is the WM transmit power, PL is the path loss with the Okumura-Hata model, G_{TX} and G_{RX} are antenna gains for the CR and WM respectively ($G_{TX} = G_{RX} = 0$ dBi in the simulator). We then set $RX_{OU} = -107$ dBm for the detection threshold as given by FCC and used in the IEEE 802.22 standard [ieee80222], such that:

$$-107\text{dBm} > 17\text{dBm} - PL\left(d_{i,j}\right)\text{dB} \tag{9}$$

$$PL(d_{i,j}) > 124\text{dB} \tag{10}$$

$$d_{i,j} > 10^{\left(\frac{124 - 69.55 - 26.16 \cdot \log_{10}(f) + 13.82 \cdot \log_{10}(h_b) + C_H}{44.9 - 6.55 \cdot 13.82 \cdot \log_{10}(h_b)}\right)} \tag{11}$$

where f is the frequency in MHz, h_b is the base station height in meters and C_H is a correction factor for user terminal height. Next, we find $d_{i,j}$ for the parameters used in the simulation scenario and find that; for the BS with parameters (f=600, h_b=15, C_H =0), $d_{bs,j} > 0.885$ km, and for the OU with parameters (f=600, h_b=1.5, C_H =0), $d_{i,j} > 0.436$ km.

For *SSE-Distance+*, the value n_c for the coarse sensing period T_P is selected based on the following criteria for the BS with distance $d_{bs,j}$ to the WM j on channel j:

$$n_c = \begin{cases} 10, & if\ d_{bs,j} \leq 0.9km \\ 100, & if\ 0.9km < d_{bs,j} \leq 1km, \\ 200, & if\ d_{bs,j} > 1km \end{cases} \tag{12}$$

and based on the following criteria for the OU:

$$n_c = \begin{cases} 10, & if\ d_{i,j} \leq 0.5km \\ 100, & if\ 0.5km < d_{i,j} \leq 0.6km. \\ 200, & if\ d_{i,j} > 0.6km \end{cases} \tag{13}$$

which gives T_P=0.1, 1 and 2 seconds respectively. A higher margin is considered for the OU compared with the BS since the OUs are mobile. Note that the protection ratio for a WM as set by FCC when using the database approach only is considered to be 1 km.

6.3.2.3 SSE-OnOff

The *SSE-OnOff* function aims to enhance the QoS and performance of the mobile OUs in scenarios where the density of the incumbent WMs is high. The function uses sensing to predict the probability that a channel will not be occupied by a WM. To do this, the SAN function uses sensing results from the cognitive radio devices to calculate the mean values for the ON (channel busy) and OFF (channel idle) periods for the WMs on each channel *j*, denoted $T_{\mathrm{j,ON}}$ and $T_{\mathrm{j,OFF}}$ respectively. The SAN function then sends this information to the *SSE-OnOff* function which selects the optimal channel ch_{OnOff} based on the following criteria:

$$ch_{OnOff} = \max_j \frac{T_{j,\mathrm{OFF}}}{T_{j,OFF} + T_{j,\mathrm{ON}}} \qquad (14)$$

$$\text{subject to: } r_{i,j} < \delta, j \in [0, M] \qquad (15)$$

where *N* is the total number of OUs connected to the BS, *M* the total number of channels to be sensed and δ the sensing detection threshold.

6.3.2.4 SSE-Hybrid

This *SSE-Hybrid* function combines *SSE-Distance* and *SSE-OnOff*. The aim of this function is to enhance QoS when the OUs are mobile and when the density of WMs and/or OUs is high.

We want to use *SSE-Distance* when the distance $d_{i,j}$ between the OU *i* and WM *j* on channel *j* is high such that the probability of a WM appearing within the detection range reduces. On the other hand, we want to use *SSE-OnOff* if the $d_{i,j}$ is low such that the WM might appear within detection range. By using the same detection ranges as calculated for the *SSE-Distance+*, the *SSE-Hybrid* function can be configured to select the optimal channel ch_{hybrid} based on the following criteria:

$$ch_{hybrid} = \begin{cases} ch_{OnOff}, & d_{i,j} \le 0.436km, d_{bs,j} \le 0.885km \\ ch_{Distance}, & d_{i,j} > 0.436km, d_{bs,j} > 0.885km \end{cases} \qquad (16)$$

$$\text{subject to: } r_{i,j} < \delta, j \in [0, M], i \in [0, N] \qquad (17)$$

6.3.3 System Performance Metrics

Different metrics are used to evaluate the performance of the IEEE 802.22 standard depending on the studied scenario. The metrics used to evaluate IEEE 802.22 performance are:

- Throughput: measured at the transport layer for the actual application used. This will not reflect the physical layer throughput which includes management frames for the MAC layer and the general network layer overhead.
- Packet Loss: measured at the network layer as the percentage of packets transmitted but not received.
- Packet Drops: a measure of packets dropped at the end of the BS queue due to congestion on the wireless link. Separate queues are implemented for the BE and UGS QoS profiles in the scheduler at the BS. Also, the BS holds a separate queue for each OU. Queue length is 50

packets. First the BS transmits all packets queued in the UGS queues, next the remaining resources are allocated fairly to packets from the BE queues. All queues drop packets at the tail.

- App Delay: measured at the application layer as the time from the packet is transmitted by the application until it is received. App delay will include delay in all TCP/IP protocol layers.
- SINR: Signal to Interference and Noise Ratio at the OU resulting from Gaussian noise and interference from WMs.

Furthermore, to assess the impact on the WM performance we use the following metrics:

- WM C/I: is the carrier to interference ratio measured at the WM, which is considered as the metric that best describes the performance of the analogue WM. The only interference will be from the IEEE 802.22 system.
- WM outage: is the percentage of time the WM experience C/I values below the required C/I level, with typical required C/I of 25 dB [Ofcom09] and 20 dB [ECC11].

All metrics are measured locally at the actual device, i.e. at the BS for packet drops, at the receiving OU for throughput, packet loss and SINR, at the BS for packet drops and at the WM for C/I. The metrics will mostly be presented as one average value for all OUs and WMs.

6.3.4 Performance Evaluation

6.3.4.1 Simulation Scenario

The QoSMOS scenario considered is cellular extension in white spaces. Furthermore, the scenario can be described as light urban with a BS cell radius of 1.2 km. The mobile OUs move following a random waypoint model with a random speed between 1 and 20 m/s. Their initial location is randomly selected within the BS radius.

The number of available TV white space channels after consulting the geo-location database is assumed to be *M*=4 channels in the simulation scenario. This means that these 4 channels neither are allocated by TV broadcasters nor registered WMs. However, unregistered WMs will appear in these channels in the simulation scenario. There are totally 4 WMs in the scenario, each one appearing separately on one of the 4 channels. Their location is randomly selected within the area of 1.4 km radius from the BS. Their inter-arrival rate and departure rates with average 20 and 5 seconds respectively are selected randomly with ±10 and ±2.5 seconds respectively. Note that during simulation, both the WM inter-arrival and departure rates will vary according to the negative exponential distribution for the initial selected inter-arrival and departure rates.

The basic parameters used in the simulation scenario are given in Table 3. Parameters specific for the IEEE 802.22 system are given in the end column in Table 4.

Parameter	**Value**
IEEE 802.22 BS height	15 m
IEEE 802.22 OU height	1.5 m
IEEE 802.22 BS EIRP	4 W (36dBm)
IEEE 802.22 OU EIRP	0.1 W (20dBm)
IEEE 802.22 Bandwidth	6 MHz
IEEE 802.22 BS cell radius	1.2 km
IEEE 802.22 OU traffic load	200 kbit/s
IEEE 802.22 OU speed	1 – 20 m/s (random), random waypoint model
# IEEE 802.22 OU (*N*)	4 … 16
WM height	1.5m

WM EIRP	0.05 W (17dBm)
WM bandwidth	200 kHz (100 kHz used for peak transmission)
WM inter-arrival time (λ_W)	20 ±10 (uniform distribution)
WM departure time (μ_W)	5 ±2.5 (uniform distribution)
WM pairs distance	100 m
Available channels (*M*)	4
#WM pairs	4
Path loss model	Cost Hata for suburban areas
Frequency bands	600 MHz (channels 605, 611, 617 and 623)
Traffic direction	Downlink

Table 3 Parameters for the simulation scenario

Each simulation is run between 40-50 times, each with a duration of 550 s, and the results are averaged. A warm up time of 50 s is used to ensure that a stable point of network operation is reached. Considering that in NS-2 all nodes will receive packets from all other nodes within detection range, irrespective of actual frequency used, and that the nodes processes the packet fully or partially, each single simulation takes about 1-2 hours depending on the number of OUs on modern computers (RedHat Enterprise Linux v5.8, 64-bit, 8 CPUs, 16GB, 2.3Hz).

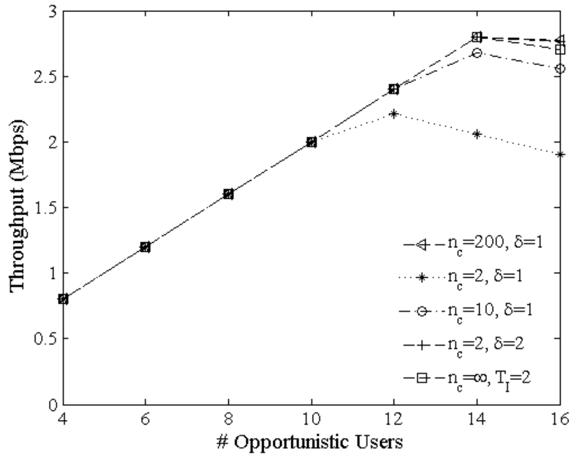
We make a note that more simulations are required to provide better average for the results than have been done for the results that will be presented in the following.

6.3.4.2 Evaluation of Capacity for Different Sensing Configurations without presence of Wireless Microphones

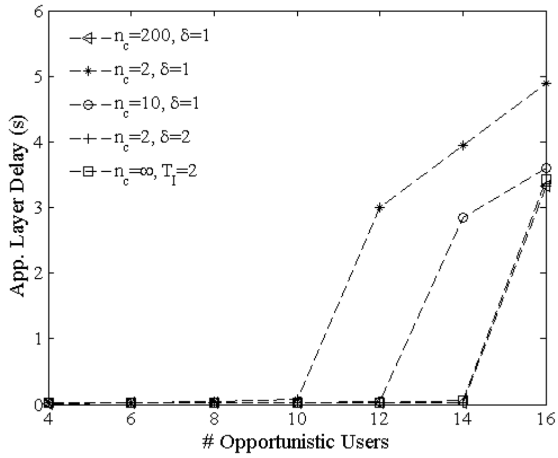
Figure 52(a) and (b) respectively plots the throughput and packet loss when there are no WMs present for the sensing configurations; coarse sensing with $\{n_c=200, \delta=1\}$, coarse sensing with $\{n_c=2, \delta=1\}$, $\{n_c=10, \delta=1\}$, consecutive coarse sensing with $\{n_c=2, \delta=2\}$, and finally fine sensing with $\{n_c=\infty, T_I=2\}$[2]. It can be seen that all sensing configurations achieve similar performance for 4 to 10 OUs. The reason for this is that packets are buffered at the BS, and that the BS buffer not fills up due to the short sensing periods. When zooming in to the plot, it is seen that only some few bits/s differs for the sensing configurations. When the number of OUs increase above 10 it can be seen that throughput degrades for $\{n_c=2, \delta=1\}$ since the number of OUs causes more false alarms. This results in more fine sensing stages that give overhead. This is also observed for $\{n_c=2, \delta=10\}$ for more than 12 OUs. The consecutive sensing strategy $\{n_c=2, \delta=2\}$ becomes similar to $\{n_c=200, \delta=1\}$ since the number of fine sensing stages started by false alarms is reduced, where only a 6 Kbit/s lower throughput is observed for 16 OUs. The fine sensing configuration $\{n_c=\infty, T_I=2\}$ achieves lower throughput for 16 OUs since more time is used for fine sensing (30ms for each 2 seconds).

It can be seen that application layer delay increases as the number of coarse sensing stages and hence false alarms that starts fine sensing stages increases. For instance it can be seen that application layer delay is higher for $\{n_c=2, \delta=1\}$ compared to the other sensing configurations. It can also be seen that application layer delay increases for all sensing configurations when the number of bits transmitted are higher than the channel capacity. This is confirmed by the queue drops in the downlink direction at the MAC layer for the IEEE 802.22 base station as illustrated in Figure 52(c).

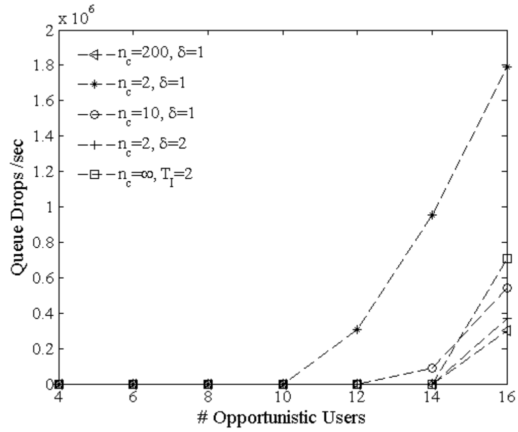
[2] Fine sensing does not use coarse sensing, hence $n_c=\infty$. The interval between the fine sensing periods is $T_I=2$ seconds and the sensing time is $T_s=30$ ms. Fine sensing is described in more detail in Appendix A.



(a) Average IEEE 802.22 Throughput



(b) Average Delay at Application Layer



(c) Queue drops at MAC layer in IEEE 802.22 base station

Figure 52 Evaluation of capacity for different sensing configurations without presence of WMs

6.3.4.3 Evaluation of the SSE Functions for Increasing Number of OUs

The simulation results for increasing number of IEEE 802.22 OUs for the SSE functions; *SSE-Distance*, *SSE-OnOff*, *SSE-Hybrid* and *SSE-Power* is presented in Figure 53, in addition to the case without presence of WMs referred to as "No WMs" (i.e. the results for $\{n_c = 200, \delta = 1\}$ presented above). We note that the variation in the simulations is high and that the number of simulations should be higher to get more accurate results.

From the average system throughput for all OUs given in Figure 53(a), it can be seen that *SSE-OnOff* achieves highest throughput for most number of OUs. This is because when the WM activity level is quite high as in the considered scenario, *SSE-OnOff* will most often select the channel that stays idle for the longest period. Hence, the number of channel switches is reduced and harmful interference to the OUs is reduced resulting in a more stable network. Harmful interference to the WM is also reduced (see average WM C/I in Figure 53(e)).

It can also be seen that *SSE-Distance* achieves slightly higher and similar throughput for 12 and 14 OUs. This comes at the cost of higher interference to the WMs (see average WM C/I in Figure 53(e)). The higher throughput and increase in interference to the WM here can be explained by the fact that the OUs often are located outside of WM detection range while still not experiencing harmful interference from the WM.

SSE-Hybrid does not achieve maximum throughput as desired. Hence the *SSE-Hybrid* function might not always select the optimal SSE function in the considered scenario. This could potentially be different if other values or another optimization were used for the SSE function selection in equation (16).

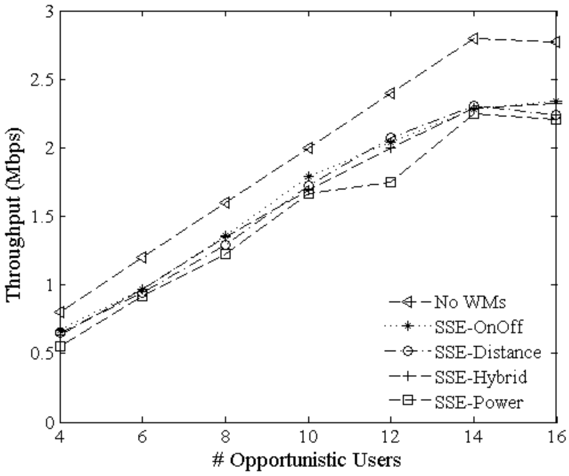
It is observation that *SSE-Distance* achieves better performance than *SSE-Power*. The two will mostly select the same channel in the case that WMs are present on all channels since the distance is derived from the sensor measurements. However, in the case that no WM is present on one or more channels during sensing, *SSE-Power* will in many cases select a channel without knowledge about where the WM might appear. Therefore, the WM might appear close to the OU such that both the OU and WM experience harmful interference. It was observed especially for *SSE-Power* that this resulted in the OU loosing synchronization with the BS which reduced throughput dramatically (explains the dip in throughput for 12 OUs). This was also observed for the other SSE functions, but less frequently.

The average packet loss given in Figure 53(b) show that *SSE-Power* generally gives higher packet loss and that *SSE-OnOff* generally gives lowest packet loss except for OUs 12 and 14, which confirms with the observation in Figure 53(a). One note on the packet loss is that it is measured at the network layer, hence packets sent to the MAC layer and dropped at the end of the BS queue during long quit periods and idle time when no channels are available are also counted.

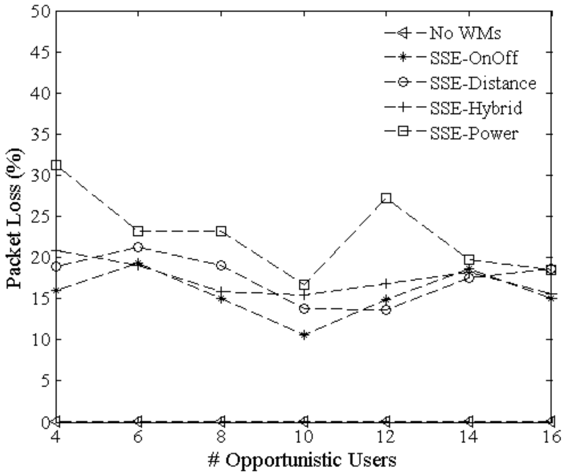
Average application layer delay given in Figure 53(c) is around 18 milliseconds when the number of OUs and traffic load is low, and increases as load increases. It can be seen that packet loss increases dramatically when the IEEE 802.22 OFDMA frame is full.

The average SINR values measured at the OUs given in Figure 53(d) show that there is some variance which would be smoother when the number of simulations increases.

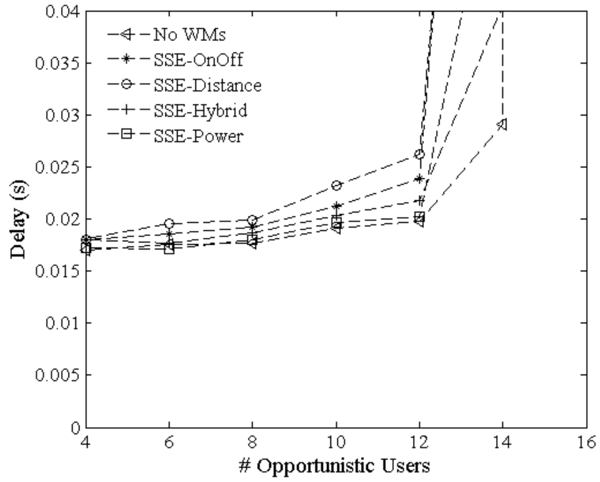
From the average experienced C/I for the WMs given in Figure 53(e), it can be seen that the *SSE-Power* interferes most with the WMs which is because it starts using the channel without historical knowledge of where the WMs might appear. It can also be seen that the percentage of time the WM experience C/I below 25 dB given in Figure 53(f) is highest for *SSE-Power*. Impact on the WM is found to be quite similar for the other SSE functions. One observation is that interference to the WM seems to increase slightly for *SSE-Distance* when the number of OUs increases, which is because the average distance to the nearest WM decreases as the number of OUs increases.



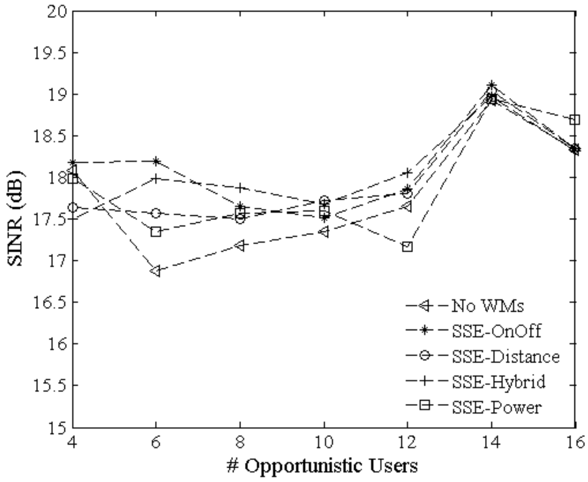
(a) Average IEEE 802.22 Throughput



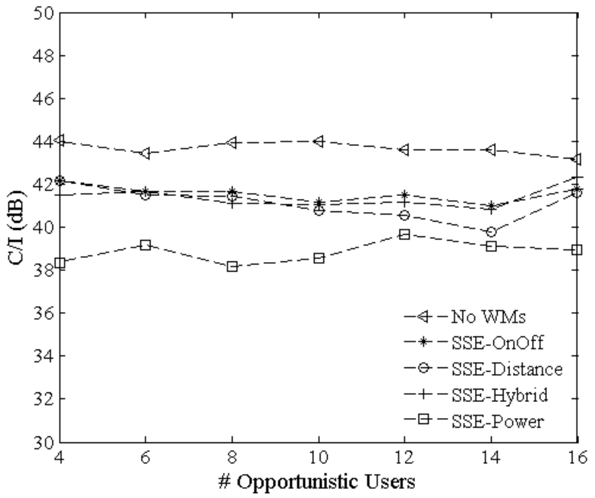
(b) Average Packet Loss for OU



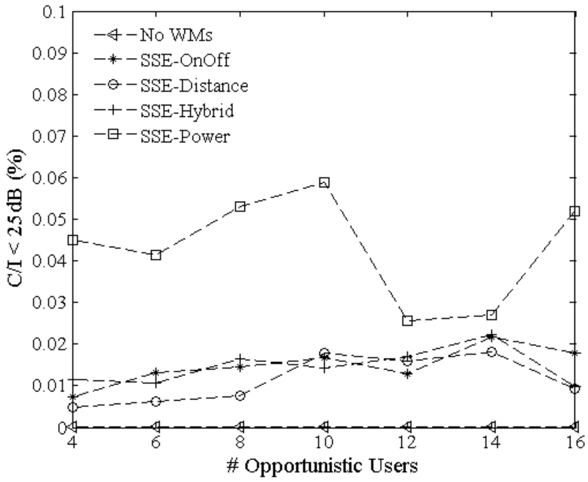
(c) Average Application Layer Delay for OU



(d) Average SINR for OU



(e) Average WM C/I



(f) Average percentage of time WM C/I < 25 dB

Figure 53 Performance evaluation of the SSE function for increasing number of OUs

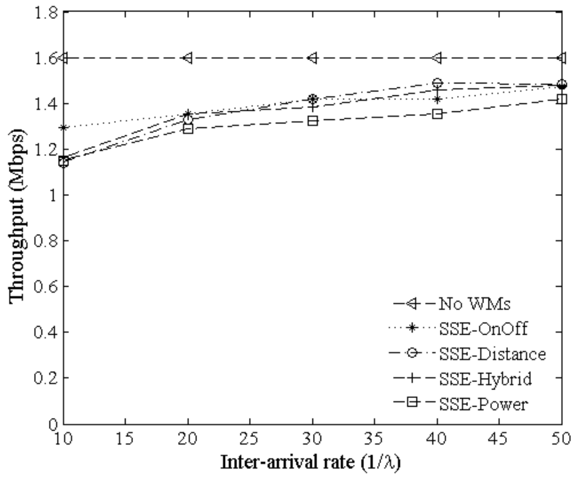
6.3.4.4 Evaluation of the SSE Functions for Various WM Activity Levels

The simulation results for increasing WM inter-arrival rate (i.e. for lower WM activity) for the SSE functions; *SSE-Distance*, *SSE-OnOff*, *SSE-Hybrid* and *SSE-Power* are presented in Figure 54, in addition to the case without presence of WMs referred to as “No WMs”.

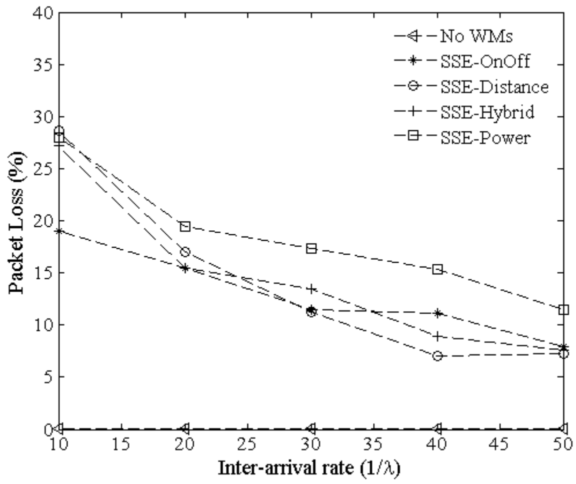
From the average throughput for all OUs given in Figure 54(a), it can be seen that *SSE-OnOff* gives best performance for high WM activity with inter-arrival rate 10 and 20, but that throughput decreases as WM activity reduces. This indicates hypothesis H.2 stated above. It can also be seen that *SSE-Distance* achieves higher throughput as the WM activity level decreases which indicates hypothesis H.1 above. *SSE-Hybrid* does not manage to achieve highest performance all the time as stated in hypothesis H.3. As noted above, this could potentially be different if the values used for SSE function selection in equation (16) were optimized or selected differently, or if the selection criteria is based on other variables. All SSE functions seem to achieve similar throughput as the WM activity level becomes low.

Average packet loss given in Figure 54(b) reduces as the WM activity reduces.

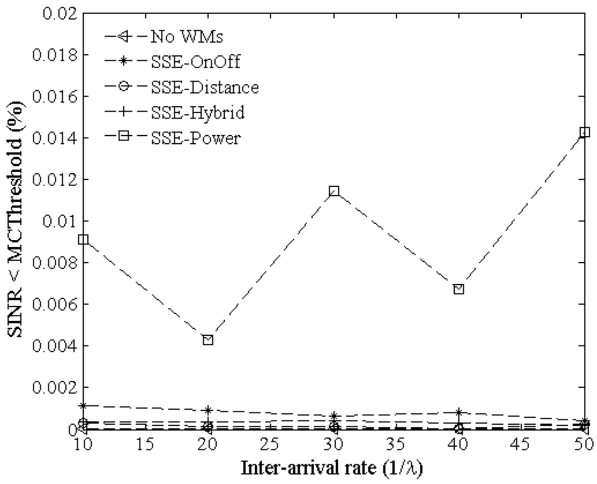
The average percentage of received packets at the OU that obtains SINR value less than the modulation and coding rate threshold given in Figure 54(c) is highest for *SSE-Power* since it more often selects channel without knowledge of how often or where the WMs are located. *SSE-Power* also causes more severe interference to the WMs as illustrated in Figure 54(d) by the average percentage of time the WMs experience C/I below 25 dB.



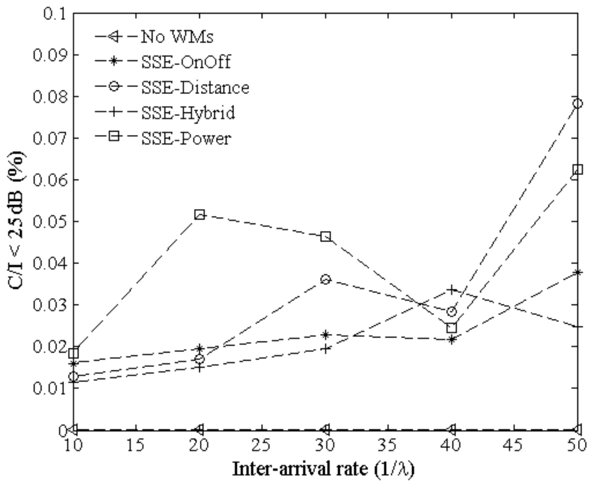
(a) Average IEEE 802.22 Throughput



(b) Average Packet Loss for OU



(c) Average percentage of received packets where SINR < Modulation and coding threshold.



(d) Average Percentage of time WM C/I < 25 dB

Figure 54 Performance evaluation of the SSE functions for different WM activity levels

6.3.5 Concluding Remarks

Three SSE functions were evaluated; *SSE-Distance*, *SSE-OnOff* and *SSE-Power*, and performance evaluation was conducted by simulations to verify three hypotheses:

H.1 *SSE-Distance* will improve performance when OUs are mobile.

H.2 *SSE-OnOff* will improve performance when the density and activity level of WMs is high.

H.3 *SSE-Hybrid* will improve performance both when OUs are mobile and the WM activity and density is high.

The results gave indications that H.1 and H.2 are correct, but did not manage to indicate H.3. It is left for further work to optimize the spectrum selection parameters for *SSE-Hybrid* and to evaluate it in more detail.

It was generally found that the SSE functions using historical and statistical knowledge about the WMs location and WM activity level enhanced performance for the IEEE 802.22 network and reduced impact on WM performance. Harmful interference was reduced for both the IEEE 802.22 network and the WM, which resulted in a more stable network with higher system throughput.

An important note on the work and results presented here is that the number of simulations should be higher to give conclusions. Therefore the work presented here does not give exact averages for the results, but they give a good indication on the performance.

Appendix A. IEEE 802.22 NS-2 Simulator Implementation

We have adapted an extensive implementation of IEEE 802.16e in NS-2 [Golmie09] developed by the WiMAX Forum [WimaxFor] to implement the IEEE 802.22 standard and the proposed SSE functions. The features that are different from IEEE 802.16 are implemented and conform to the best extent to the functional requirements as specified in the IEEE 802.22 standard. The main parameters in our IEEE 802.22 NS-2 simulator are given in Table 4 with a comparison to the main profiles in the IEEE 802.16 and IEEE 802.22 standards.

Parameter	IEEE 802.16	IEEE 802.22	NS-2
Bandwidth	10 MHz	6, 7, 8 MHz	6 MHz
FFT size	1024	2048	2048
Frequency/channels	2.5-269 GHz	54-698 MHz	54-698 MHz
Frame size	5 ms	10 ms	10 ms
Duplexing method	TDD	TDD	TDD
Tx/Rx Trans. gap (TTG)	105.7µs	83.33/190/270 µs	83.33 µs
Rx/Tx Trans. gap (RTG)	60 µs	210 µs	210 µs
Modulation types	QPSK, {16,64}-QAM	QPSK, {16,64}-QAM	QPSK, {16,64}-QAM
Coding rates	1/2, 2/3, 3/4, 5/6	1/2, 2/3, 3/4, 5/6	1/2, 2/3, 3/4
Error corr. coding	CC, CTC, LDPC	CTC/BTC	No (emulated)
Max power	BS:43, OU: 23dBm	BS: 36, OU: 36 (20 if mobile)	BS: 36, OU: 20dBm
Assumed noise figure	BS: 4, OU: 7 dB	BS: 3, OU: 6dB	BS/OU: 4dB
QoS classes	UGS, rtPS, ErtPS, nrtPS, BE	UGS, rtPS, ErtPS, nrtPS, BE	UGS, BE
Cyclic prefix mode	1/4, 1/8, 1/16, 1/32	1/4	1/4
OFDM mapping	Rectangular	DL: vert., UL: horiz.	Vertical
Error protection	HARQ	ARQ	ARQ

Subcarrier spacing	10.94 kHz	3.348/3.906/4.464 kHz	3.348 kHz
Useful symbol length	91.4 µs	298.7/256/224 µs	298.7 µs
Guard time	11.4 µs	37.34/32/28 µs	37.34 µs
Symbol duration	102.9 µs	373.3/320/280 µs	373.3 µs
Sampling frequency	11.2MHz	6.857/8/9.145MHz	6.857MHz
Sampling period	0.18 ms	0.299/0.256/0.224 ms	0.299 ms
Symbols per frame	48	26/30/34	26
Used subcarriers	840	1680	1680
Guard and null subcarr.	184	368	368
Pilot subcarriers	DL: 120, UL: 280	DL/UL: 240	DL/UL: 240
Data subcarriers	DL: 720, UL: 560	DL/UL: 1440	DL/UL: 1440
Subcarriers/subchannel	DL: 24, UL: 28	DL/UL: 24	DL/UL: 24
Subchannels	DL: 30, UL: 35	DL/UL: 60	DL/UL: 60
Pilot location	Distributed	Distributed	Distributed
Power control	Distributed	Distributed	No
Subcarrier allocation	FUSC, PUSC, Cont.	PUSC	PUSC
Sensing strategy	N/A	Optional	Two-stage/fine sensing
Coarse sensing duration	N/A	Optional	1 ms
Coarse sensing interval	N/A	Optional	2 sec
Fine sensing duration	N/A	Optional	30 ms
Fine sensing interval	N/A	Optional	∞ (event based)
Cooperative sensing	N/A	Optional	"OR" rule
WM detection threshold	N/A	Optional	-107 dBm

Table 4 IEEE 802.22 implementation in NS-2 compared with the IEEE 802.16 [ieee80216][Wimax06] and IEEE 802.22 [ieee80222] standards.

A.1. OFDMA and Channelization Structure

Both the IEEE 802.22 downlink and uplink subframes have totally 60 subchannels, where each subchannel consists of 28 subcarriers out of which 24 data and 4 pilot subcarriers. The IEEE 802.22 network can operate on any vacant TV channel not used by the incumbent user. Channel bonding of scattered available channels is not considered in this study, therefore only one available UHF channel will be used by the IEEE 802.22 system at any time. A WM occupies only one channel.

In our IEEE 802.22 NS-2 implementation only the 6 MHz profile is used. For the subcarrier allocation strategy, partially used subcarrier (PUSC) allocation [Wimax06] is used. Guard bands are considered at both ends of the channel bandwidths with a total of 368 guard and null subcarriers.

The transmit/receive transition gap (TTG) is set to 210 µs, which supports a 30 km distance between the BS and OU. A dynamic TTG is needed for greater distances, however this is not implemented since the simulation scenarios considered involve only small distances. The receive/transmit transition gap (RTG) is set to 88.33 µs. There are totally 26 symbols, each of 373.33 µs duration. The DL:UL ratio is set to 2:1.

A.2. OFDMA Traffic Scheduling

The MAC layer of IEEE 802.22 uses linear scheduling to allocate OFDMA slots to traffic from the upper layers in both the downlink and uplink subframes, as opposed to a rectangular scheduling in IEEE 802.16. An OFDMA slot can be characterized as a {subchannel, symbol}-tuple in the frequency and time domain. Vertical striping is used for both DL and UL subframes in the simulator, which means that OFDMA slots are allocated in frequency first and then in time, i.e. the first symbol is filled with data before the next symbol.

A.3. Propagation and Channel Modeling

The propagation model used in the simulator is the Okumura-Hata [Oku68] path loss model, configured for suburban scenarios. Further, the Okumura-Hata model is combined with a Clarke-Gans [Clarke68] implementation of Rayleigh fading, which has been extended to support 2048 subcarriers for IEEE 802.22. The Vehicular A ITU power delay spread model [ITUR97] is used which is suited for the considered IEEE 802.22 scenario. The BS is configured with 36 dBm transmit power and the mobile OU is configured with 20 dBm transmit power. Antenna gains for both the BS and OU are set to 0 dB. Dynamic transmit power adaptation is not implemented.

Interference modeling is done at the subcarrier level by capturing packets from all IEEE 802.22 nodes and WMs. When the received signal to interference plus noise ratio (SINR) on each subcarrier is calculated for each packet, a decision is made to further process or drop the packet. This is done by first finding the exponential effective SIR mapping (EESM) [EESM03] to get the effective SINR and then extracting the block error rate (BLER) from the SINR, modulation and coding rate and block size. Based on the BLER value a decision is made whether to drop the packet or not[3].

A.4. Error Protection

Channel errors are considered in the simulations. For error protection the IEEE 802.22 standard uses ARQ. For error correction coding the IEEE 802.22 uses convolutional turbo codes (CTC) and block turbo code (BTC), although only ARQ is implemented in NS-2. However, if BLER found as described in Section A.3 is above a threshold set to 4% (recommended by the WiMAX Forum IEEE 802.16e implementation [Wimax06]) the simulator emulates that the erroneous bits are corrected.

A.5. Wireless Microphone Implementation

A typical analogue WM in the TV bands uses a narrow bandwidth of 200 kHz, which amount to 68 active subcarriers in the NS-2 OFDMA simulator. Since an analogue WM most of the time focus the transmit power on a narrow part of the 200 kHz bandwidth, we assume that the WM on average uses half the bandwidth with 34 subcarriers when implementing the WM in the NS-2 simulator. In this case, if the WM transmits with 50 mW, the transmit power per subcarrier for the WM is set to 0.05W/34=0.0015W in the NS-2 simulator. It should be noted that the NS-2 simulator is packet transmission-based, hence we assume that WM virtually transmits data packets consecutively during the whole IEEE 802.22 OFDMA frame in order to simulate the 100% duty-cycle traffic pattern. The modulation and coding rate is QPSK 1/2.

A.6. Wireless Microphone Detection Process

The main functions involved in incumbent user detection in the IEEE 802.22 standard are the spectrum manager at the BS and the spectrum automaton at the OU which controls the spectrum

[3] Please refer to [Golmie09] for a detailed description of the OFDMA physical layer implementation and interference modeling. Note that the propagation model, operating frequency range and system profiles are re-implemented to fit the UHF bands, IEEE 802.22 and suburban scenarios.

sensing and the geo-location function. These are described in more detail in [D6.1]. All functions are implemented in NS-2, except for dynamic transmit power control. Since the latter is not implemented, the IEEE 802.22 will switch to a new channel if a WM is detected on that channel. If no channels are available IEEE 802.22 will cease transmission.

Spectrum Manager

The spectrum manager is implemented in the BS and is responsible for deciding which channel to use. It specifies the set of channel lists, i.e. the backup, candidate and protected channel lists. In the IEEE 802.22 standard a channel will get status as backup channel when sensed as unused every six seconds over a period of 30 seconds. In our simulator, these backup-channels are subject for selection in the SSE functions implemented.

In the OU, the spectrum automaton is a lightweight version of the spectrum manager in the BS. The spectrum automaton is controlled by the BS and is mostly responsible for reporting information to the spectrum manager. In rare cases, the spectrum automaton itself is responsible for sensing at initial OU power on, when it loses contact with the BS and during an idle time when there are no tasks pending.

Spectrum Sensing Implementation

The spectrum sensing function can first be classified into in-band sensing, that senses the operating channel, and out-of-band sensing, that senses activity on other channels that potentially can be used by the IEEE 802.22 system. For in-band sensing the two-stage spectrum sensing approach, as specified by the standard, is implemented as the default sensing strategy in the simulator. At the coarse sensing stage (first stage) a simple energy detection is used for frequent and short sensing duration T_c. If coarse sensing detects a WM signal it switches to the fine sensing stage (second stage) that uses a more detailed WM detection process for a longer duration T_s. A simple energy detector will also be used for fine sensing in the simulator. If a WM signal is detected by fine sensing then the operating channel is switched to one of the backup channels.

Probability of detection P_D and probability of false alarm P_{FA} are set in the simulation separately for coarse and fine sensing. In the simulator, P_{FA} will be set to a given value while P_D will be defined by a function of P_{FA} as detailed in the design of the energy detector below. The effect of P_D and P_{FA} on the IEEE 802.22 will mostly be considered for coarse sensing which is more unreliable than fine sensing. Fine sensing that uses more advanced sensing techniques is considered to be more accurate and P_D and P_{FA} are therefore set to one and zero, respectively. Further, cooperative sensing with the OR rule is implemented for all sensing stages, which is mandatory in the US as specified in the IEEE 802.22 standard [Sec. 8.6.1.3]. The total P_D and P_{FA} can be calculated analytically as $P_D = 1 - (1 - P_{D,i})^N$ and $P_{FA} = 1 - (1 - P_{FA,i})^N$, where N is the number of cooperative sensing nodes and $P_{D,i}$ and $P_{FA,i}$ is the probability of detection and false alarm for each single node, respectively.

Coarse sensing is carried out during allocated time periods at the end of the uplink OFDMA subframe with sensing duration T_c. Coarse sensing occurs at every n_c OFDMA frame. Quiet periods for fine sensing are scheduled at the MAC layer. The fine sensing duration is $T_s = 30$ ms, spanning three OFDMA frames. A pre-defined fine sensing period T_I (i.e. time between fine sensing stages) might be set, including infinite, if the event based strategy is used.

Three sensing strategies are implemented for in-band sensing in the IEEE 802.22 simulator:

- Two-stage spectrum sensing : with coarse sensing $T_c = 1$ ms every $n_c = 200$ OFDMA frame, in which coarse sensing detection triggers a fine sensing stage with duration $T_s = 30$ ms;
- Two-stage consecutive spectrum sensing [Jeon08]: same as coarse sensing, but $\delta = 2$ consecutive coarse sensing detections are needed before fine sensing is triggered, for which $T_s = 30$ ms;

- Single-stage spectrum sensing: only fine sensing with $T_s = 30$ ms, time based with a sensing period $T_I = 2$ seconds (i.e. the interval between fine sensing);

Finally, in IEEE 802.22, and our NS-2 implementation, out-of-band sensing is performed during quiet periods allocated for the fine sensing periods, and all relevant channels are sensed during one fine sensing period.

Implementation and Design of Energy detector

The energy detector is simple in that it determines if a signal is present or not on a channel by comparing the energy of a sensed signal with a given threshold. It does not consider specific signal features.

The sampled signal at the energy detector is

$$y(n) = x(n) + w(n) \tag{18}$$

where $x(n)$ is the received signal samples and $w(n)$ is the noise samples, which respectively have power P_S and P_N. To detect whether the signal is present or not, we set up the two hypotheses:

$$H_0: y(n) = w(n) \tag{19}$$

$$H_1: y(n) = x(n) + w(n) \tag{20}$$

The signal energy is measured during a finite time interval and compared to a threshold γ. The test statistic of the energy detector is:

$$T = \sum_{n=1}^{M} |y(n)|^2 \tag{21}$$

Since we have a large number of samples M in the vector y, the test statistic can be approximated as a Gaussian random variable by the central limit theorem:

$$T \sim N\left(P_s + P_N, \frac{(P_s + P_N)^2}{M}\right) \tag{22}$$

where the mean is $E[T] = P_s + P_N$ and the variance $Var[T] = \frac{(P_s+P_N)^2}{M}$. Note that for small number of samples M, the chi square distribution would have been used for T.

The detector threshold can then be found based on the required false alarm probability:

$$\gamma = \mathrm{P_N}\left[1 + \frac{\mathrm{Q}^{-1}(\mathrm{P_{FA}})}{\sqrt{\mathrm{M}}}\right] \tag{23}$$

The probability of detection in an AWGN channel is then given by:

$$P_D = P\{T > \gamma \mid H_1\} \tag{24}$$

$$P_D = 1 - Q\left(\frac{\sqrt{M}}{P_S + P_N}(P_S + P_N - \gamma)\right) \quad (25)$$

In the simulator we will calculate the probability of missed detection when a WM signal is sensed, which is given by:

$$P_{MD} = 1 - P_D = Q\left(\frac{\sqrt{M}}{P_S + P_N}(P_S + P_N - \gamma)\right) \quad (26)$$

The energy detector works well when the noise power is known. However, in a realistic scenario, it is difficult to estimate the noise power exactly. Hence, we assume that the noise power is not known exactly and that it has an uncertainty of $\pm\Delta$ dB. To consider this change in delta we add it to the noise power as follows:

$$P_N = 10\log_{10}(kTB) + \text{NF} \pm \Delta \quad (27)$$

where k is the Boltzmann constant, T=300 Kelvin is the temperature, and B=6MHz is the bandwidth.

The probability of false alarm for a given detector threshold γ can be found by:

$$P_{FA} = P\{T > \gamma \mid H_0\} \quad (28)$$

$$P_{FA} = 1 - Q\left(\frac{\sqrt{M}}{P_N}(P_N - \gamma)\right) \quad (29)$$

Now, the detection threshold γ can be calculated for a considered $P_{FA} = 0.01$ by using equation (23) with P_N as in Eq. 0 with noise uncertainty $\Delta = 0$ dB. Furthermore, we use the calculated detection threshold γ to find the worst case P_D and P_{MD}. Sensing time is set to 1 ms, hence we will have M=6MHz*1ms=6000 samples. The sensing threshold for a wireless microphone is -107dBm (averaged over 200kHz) as specified in the IEEE 802.22 standard [ieee80222]. A noise figure NF = 4 dB is assumed, which can be considered realistic for a sensor. The resulting probability of detection for $\Delta = 0$, 0.5 and 1 dB are given for different received signal strength (RSSI) values in Figure 55. It can be seen that the energy detector without noise uncertainty (i.e. $\Delta = 0$ dB) performs well and will be able to detect the WM for the sensing threshold -107 dBm without problems. It is however interesting to see that the minimum signal that can be detected, referred to as the power wall, for $\Delta = 0.5$ and 1 dB are -108.9 dBm and -105.7 dBm respectively. This indicates that the energy detector with noise uncertainty Δ=1 will not be able to detect the WM at -107 dBm.

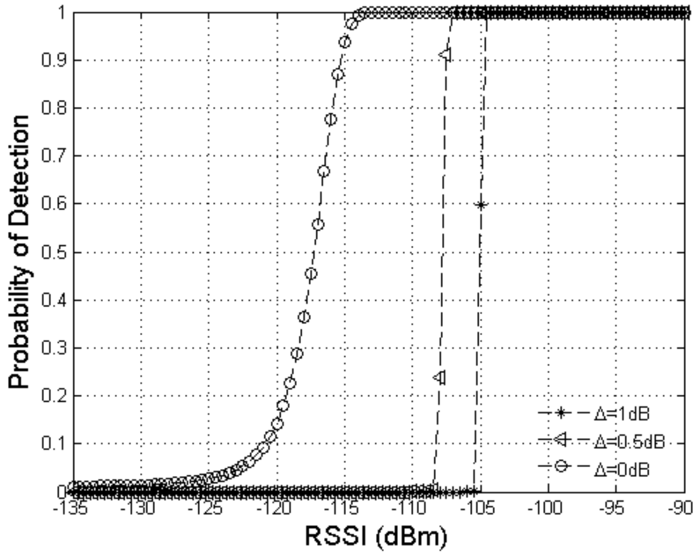


Figure 55 Probability of detection for noise uncertainty Δ=0, 0.5 and 1dB as a function of RSSI for probability of false alarm P_{FA}=0.01. Samples M=6000, sensing time 0.5ms, sensing BW=6MHz, noise figure 4dB, detection threshold -107dBm.

In a realistic scenario, it will be reasonable to consider a noise uncertainty of 1 dB. Hence, we design our sensor to divide the sensing stage into two sub-sensing stages where each sub-sensing stage will sense 3 MHz for 0.5 ms. This will reduce the noise floor with 3 dB, however the number of sensing samples reduces from 6000 to M=3MHz*0.5ms=1500. Also note that the total false alarm for the sensor should be $P_{FA} = 0.01$, hence we need to find the probability of false alarm for each single sub-sensing stage P_{FA}^{sub}. For N=2 sub-sensing stages this can be found by $P_{FA} = 1 - \left(1 - P_{FA}^{sub}\right)^N$, resulting in $P_{FA}^{sub} = 0.005$ (this can also be found by using equation (29)). A potential issue with this sensor design is that a WM can appear on the border between the two 3MHz bands for the two sub-sensing stages. However, we consider this a minor issue since the sensor could be designed with a small overlapping area in the sensed bandwidths for the two sub-sensing stages.

The resulting probability of detection for $\Delta = 0.5$ and 1 dB as a function of RSSI is plotted in Figure 56. It is found that the power wall for $\Delta = 0.5$ and 1dB are -112.2 dBm and -108.8 dBm respectively. It is also found that $P_D = 0.993$ for RSSI $= -107$ dBm with $\Delta = 1$ dB.

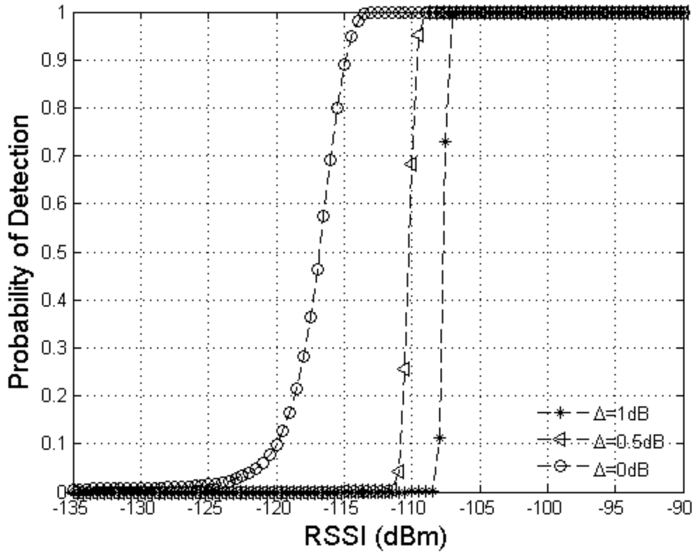


Figure 56 Probability of detection for noise uncertainty Δ=0, 0.5 and 1dB as a function of RSSI for probability of false alarm P_{FA}=0.005. Samples M=1500, sensing time 0.5ms, sensing BW=3MHz, noise figure 4dB, detection threshold -107dBm.

For the simulator implementation then, the false alarm probability is set to $P_{FA} = 0.01$ or more specifically with a probability $P_{FA}^{sub} = 0.005$ for each of the two sub-sensing stages each over 3 MHz. Missed detection is randomly selected with a probability P_{MD} based on lookup in a table for the equation given in equation (26).

6.4 Selection of DTV bands for LTE uplink extension

This section completes the study introduced and discussed in Appendix A of [D6.5]. In particular, a performance evaluation study of the "cellular extension in TeleVision White Spaces (TVWS)" scenario defined in [D1.2] is provided. The objective is to analyse and determine, by means of system-level simulations, the conditions under which the considered scenario would be feasible and its technical implications. This study provides quantitative and qualitative reference results as well as some guidelines to help decision entities, such as the Spectrum Selector (SSE), to make decisions on the adequacy of selecting DTV bands for the extension of a cellular system and evaluate the expectable consequences in terms of protection of the primary DTV system, performance of the LTE cellular system and efficiency of spectrum utilisation. These three aspects are quantified and analysed based on the set of conformance metrics defined in A.1.3 of [D6.5].

6.4.1 Simulation platform

The approach employed to analyse the considered scenarios is based on system-level simulations. Interference interactions could be estimated by means of radio propagation equations (e.g., as it was done in [D6.1], [D6.2] and [D6.3] for the downlink of an LTE system). However, for the analysis of the uplink, a simulation-based approach was considered to be more convenient. In terms of interference interactions, computing the effective aggregated interference generated by a reduced number of static base stations could be relatively simple based on radio propagation equations while it would result infeasible for a high number of mobile terminals. Moreover, system-level simulations allow for a simple analysis of other aspects such as the performance of the secondary LTE system and the efficiency of spectrum utilisation.

The simulated scenario is composed of a complete LTE cellular network where a primary DTV transmitter is present at a certain distance from the geometrical centre of the cellular network. The simulation platform is divided into three main modules: a) main module which integrates general aspects such as the cell deployment, mobility and traffic models, pathloss models, shadow (slow) fading models, multipath (fast) fading models, antenna radiation patterns, etc., b) downlink module which integrates aspects related to the downlink of the LTE system (scheduler, link adaptation, handover and other radio resource management methods for the downlink), and c) uplink module integrating similar aspects related to the uplink.

Table 5 summarises the main configuration parameters relevant to this study. The platform considers an LTE cellular system composed of 19 trisectorial cells arranged in 2 tiers. Three different load levels (low, medium and high) as a function of the amount of available resources are considered as shown in

Table 6. The characteristics of the primary and secondary receivers have been selected in accordance with the set of parameters considered in [D6.1], [D6.2] and [D6.3] in order to provide similar evaluation conditions. The minimum CNR required for a proper operation of DTV receivers is set at 21 dB as indicated in Table 1 of [ECC159] for fixed receivers. Several organisations [ITU1368][FCC04168][ARIBB21] have provided requirements for the DUR and its details. In spite of different standards, the requirement is quite similar for the different broadcast formats, which is generally the same as the required minimum CNR, on the assumption that the interferer has a noise-like characteristic. Following the same assumption as in [D6.1][D6.2][D6.3], the DUR is set equal to the CNR. For pathloss calculations, several empirical radio propagation models can be applied depending on the propagation environment. Examples of such models are the Okumura-Hata model [Oku68][Hat80] and the COST 231 extensions [Cic98], the ITU-R P.1546 model [ITU1546] and the flat-terrain model. The pathloss models employed in this study have been selected based on their range of applicability in terms of the environment, frequency range and distance. As shown in Table 5, the Hata COST 231 models are employed in most of the rays between transmitters and receivers. These models were envisaged for cellular mobile communication systems but are still valid in terms of operating frequency and distance for the DTV signals. However, for the interference computation from the DTV transmitter to the secondary receivers (i.e., eNodeBs), the distances observed in practice were higher than 20 km, beyond which the Hata COST 231 model is not applicable. In such a case, a simple but optimistic alternative would be to make use of the free space model. A more accurate but complex alternative would be the ITU-R P.1546 model [ITU1546]. In this study, the Egli model [Egl57] has been employed, which constitutes an intermediate alternative between both extreme points in terms of complexity and accuracy. Typically used for outdoor line-of-sight point-to-point links, this simple terrain model provides the pathloss as a function of the transmitter and receiver heights, distance and frequency. The model does not account for travel through some vegetative obstruction or other factors accounted for by the ITU-R P.1546 model, thus providing a presumably less accurate estimation of the pathloss. However, the model was derived from real-world data on UHF and VHF television transmissions in several large cities and is applicable to scenarios where the transmission has to go over an irregular terrain, thus providing a much more accurate estimation than the free space model. Besides pathloss models, shadow (slow) fading and multi-path (fast) fading models are considered as well. Shadow fading is modelled as a log-normal process with 0-dB mean and 5.5-dB standard deviation [ITU1546] as in [D6.1][D6.2][D6.3]. Fast fading is assumed to be constant during one Transmission Time Interval (TTI), but independent between consecutive TTIs. A pre-recorded fast fading trace of 3 seconds (3000 TTIs) is employed and repeated periodically for longer simulations.

An important configuration parameter not shown in Table 5 is the simulation length. Shorter simulation times result in a reduction of the overall time required to complete the study, at the expense of a loss of accuracy in the obtained results. Therefore, in order to guarantee statistically representative results, a minimum simulation length must be respected. In order to determine a reasonable simulation time, several simulations were performed and the relative error observed in some performance metrics were analysed as a function of the number of simulated TTIs. Table 7 shows the relative error observed for the throughput, BLER and BER metrics, as a function of the simulation length in terms of the number of simulated TTIs. As it can be observed, increasing the simulation time provides more accurate results. For a simulation length of 500 TTIs the observed relative error is between 2-5%, which decreases to about 1-2% for a simulation time of 1000 TTIs. Increasing the simulation length beyond 1000 TTIs does not provide appreciable improvements on the statistical accuracy of the obtained results. Therefore, a simulation length of 1000 TTIs was selected for this study.

Table 5. Configuration of the simulation platform.

PARAMETER	**VALUE**	**UNIT**
Common parameters		
Carrier frequency	600	MHz
LTE bandwidth	1.4 / 5 / 20	MHz
Cellular layout		
No. of eNodeB	19	–
No. of sectors per eNodeB	3	–
No. of eNodeB rings/tiers	2	–
Inter eNodeB distance	500	M
UE speed	5	km/h
No. UE per sector	See Table 6	–
LTE UE parameters		
Antenna height	1.5	M
Antenna gain	0	dB
Max. transmission power	24	dBm
Modulation	QPSK and 16-QAM	–
LTE eNodeB parameters		
Antenna height	30	M
Antenna gain	14	dB
Noise figure	5	dB
BLER target	0.1	–
DTV transmitter parameters		
Antenna height	100	m
Antenna gain	10	dB
Transmission power	1000	W
DTV receiver parameters		
Antenna height	10	m
Antenna gain	7	dB
Noise figure	7	dB
Required CNR	21	dB
Required DUR	21	dB
Pathloss models		
DTV TX to DTV RX (signal)	Hata COST 231 macrocell	–
DTV TX to eNodeB (interf.)	Egli	–
UE to eNodeB (signal)	Hata COST 231 macrocell	–
UE to DTV RX (interf.)	Hata COST 231 microcell	–
Shadow (slow) fading		
Fading type	2D space-correlated maps [Cai03][Cla05]	–
Map resolution	5 meters/pixel	–
Mean	0	dB
Standard deviation	5.5	dB

Inter-site fading correlation	0.5	–
Multipath (fast) fading		
Power Delay Profile (PDP)	Pedestrian	–
Trace length	3	s

Table 6. Simulated load levels.

		Number of UE per sector		
LTE bandwidth	**No. of RBs**	**Low load**	**Medium load**	**High load**
1.4 MHz	6	1	3	6
5 MHz	25	5	12	25
20 MHz	100	20	50	100

Table 7. Relative error in statistical results as a function of the simulation length.

	Relative error (%)		
Simulated TTIs	**Throughput**	**BLER**	**BER**
200	9.22	11.06	4.50
300	5.46	9.02	2.62
500	4.59	4.59	2.33
1000	1.34	2.34	2.13
2000	1.27	2.32	1.03
5000	1.27	0.004	0.41

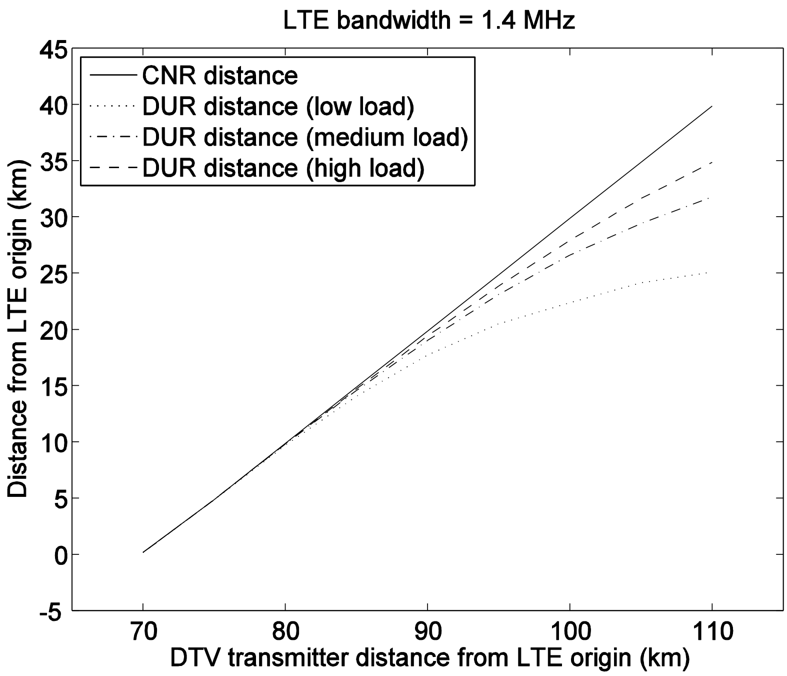


Figure 57. Protection of the primary DTV system (LTE bandwidth of 1.4 MHz).

6.4.2 Results

6.4.2.1 Protection of the primary DTV system

The considered scenario is first analysed in terms of the protection of the primary DTV system. Figure 57, Figure 57, Figure 58 and Figure 59 show the distance at which the minimum required CNR and

DUR are met as a function of the distance between the primary DTV system and the secondary LTE system. The results are shown for various LTE bandwidths and the load levels shown in Table 6. To begin with, the observed results and trends are commented and justified. Afterwards, these results are analysed in terms of: a) the minimum distance between the DTV and LTE system required for the spectrum to be selected by the CM-SM for opportunistic exploitation, b) the amount of primary spectrum that could be selected by the CM-SM for reutilisation by secondary LTE users, and c) the load of secondary LTE users that could be supported. Notice that these three points have relevant consequences in the process of spectrum management and are important aspects to be taken into account by the CM-SM when deciding on the reutilisation of primary spectrum from the DTV band for secondary exploitation.

Results and observed trends

In the first place, it can be appreciated in Figure 57, Figure 58 and Figure 59 that when the DTV and LTE systems are 70 km apart, the distance from the LTE system at which the minimum CNR is satisfied is approximately zero, meaning that the coverage area of the DTV transmitter is this simulation is about 70 km. As the distance between the DTV and LTE systems increases linearly, the minimum CNR distance, measured from the LTE system, increases linearly as well. This is a logical result since the CNR experienced by the primary DTV receivers is independent of the secondary LTE system. Thus, if the primary DTV transmitter is shifted by a certain distance, the minimum CNR distance (i.e., the border of the coverage area) is shifted by exactly the same distance. Therefore, the curve for the minimum CNR distance in Figure 57, Figure 58 and Figure 59 has a unitary slope regardless of the LTE system bandwidth and supported load. On the other hand, as appreciated, the minimum DUR distance does not increase linearly with the separation between the DTV and LTE systems, but logarithmically. This could be explained in terms of radio propagation by the fact that the average interference power received from the LTE system attenuates logarithmically with the distance (pathloss).

Minimum distance between the DTV and LTE systems

Based on the observed behaviours of the CNR and DUR distances, the minimum separation distance between the DTV and LTE system for the spectrum to be selected by the CM-SM for opportunistic exploitation can be determined as follows. As shown in Figure A-2 of [D6.5], the primary system is protected in terms of interference as long as the DUR boundary is out of the coverage area determined by the minimum required CNR. In Figure 57, Figure 58 and Figure 59, this occurs when the DUR distance, as measured from the LTE system, is lower than the CNR distance. When both values are similar, the boundary of the primary DTV system operates on the limit of the minimum required DUR value. However, the lower the DUR distance is with respect to the CNR distance in these figures, the more protected is the primary DTV system against LTE interferences and interference fluctuations caused by radio propagation effects. Based on this observation, the minimum distance required between the DTV and LTE systems can be determined as the point where the CNR and DUR curves begin diverging in Figure 57, Figure 58 and Figure 59. For instance, based on Figure 57 for an LTE bandwidth of 1.4 MHz, the minimum distance between the DTV and LTE systems for low LTE loads is around 80 km. In other words, the LTE system needs to operate 10 km apart from the boundary of the coverage area of the primary DTV system. Similarly, for medium/high loads the minimum distance is around 15 km from the boundary of the DTV coverage area. These values would need to be slightly increased depending on the desired protection margin against interference and radio propagation fluctuations. Such a protection margin could be imposed in terms of the minimum difference required between the CNR and DUR distances, with a higher minimum difference representing a higher protection margin.

Based on the results of Figure 57, Figure 58 and Figure 59, Table 8 shows the minimum distance required between the LTE system and the border of the DTV coverage area in order to guarantee an adequate protection of the primary system in terms of the minimum required DUR. The results are shown as a function of the LTE bandwidth and the supported load. These results have been obtained as the exact distance at which the minimum required DUR value is met (Table 5) and do not account for potential fluctuations on the received interference power from the LTE system. Assuming a protection margin of 10 dB for the desired DUR (i.e., the maximum fluctuation that could be tolerated in the received interference power while still satisfying the minimum required DUR), the distances shown in Table 8 should be increased between 10 km (for 1.4-MHz LTE bandwidth and low loads) and 30 km (for 20-MHz LTE bandwidth and high loads) approximately. Note that these protection results refer to the interactions between the uplink component of the LTE system and the DTV system and do not take into account the downlink component of the LTE system, which was already analysed in Section 7.5 of [D6.1], Section 4.2 of [D6.2] and Section 4.3 of [D6.3].

It is interesting to observe in Table 8 that the minimum separation distance required between the DTV and LTE systems increases with the supported load. This can be explained by the fact that a higher number of UEs results in a higher aggregated interference level at the primary DTV receivers. Therefore, as the load increases, the LTE system needs to operate further away from the DTV system in order to keep the aggregated interference below the maximum tolerable level and meet the required minimum DUR.

Regarding the impact of the selected bandwidth on the minimum required distance, it is interesting to note that a wider block of spectrum implies the availability of a higher amount of radio resources and therefore a higher maximum number of supported users. Therefore, the larger the selected bandwidth is, the higher the maximum potential number of supported UEs and aggregated interference is, and the larger the operation distance should be.

Note that the minimum required distance between the DTV and LTE systems is determined by the maximum tolerable aggregated interference (minimum DUR), which depends on the actual number of supported users rather than the operation bandwidth. In fact, the comparison of Figure 57 and Figure 58 clearly indicates that an LTE bandwidth of 5 MHz at low loads results in a similar DUR curve as an LTE bandwidth of 1.4 MHz at high loads, since in both cases the number of UEs is similar (see

Table 6). Similarly, the comparison of Figure 58 and Figure 59 shows that an LTE bandwidth of 20 MHz at low loads results in a similar DUR curve as an LTE bandwidth of 5 MHz at high loads, since in both cases the number of UEs is similar as well (see

Table 6). As a result, the operation distance from the DTV system should be increased when employing larger operation bandwidths.

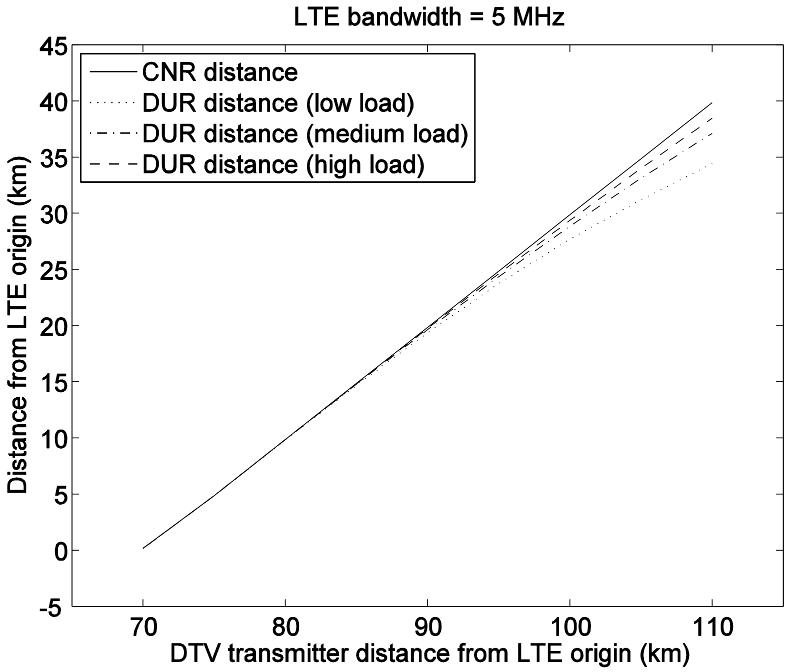


Figure 58. Protection of the primary DTV system (LTE bandwidth of 5 MHz).

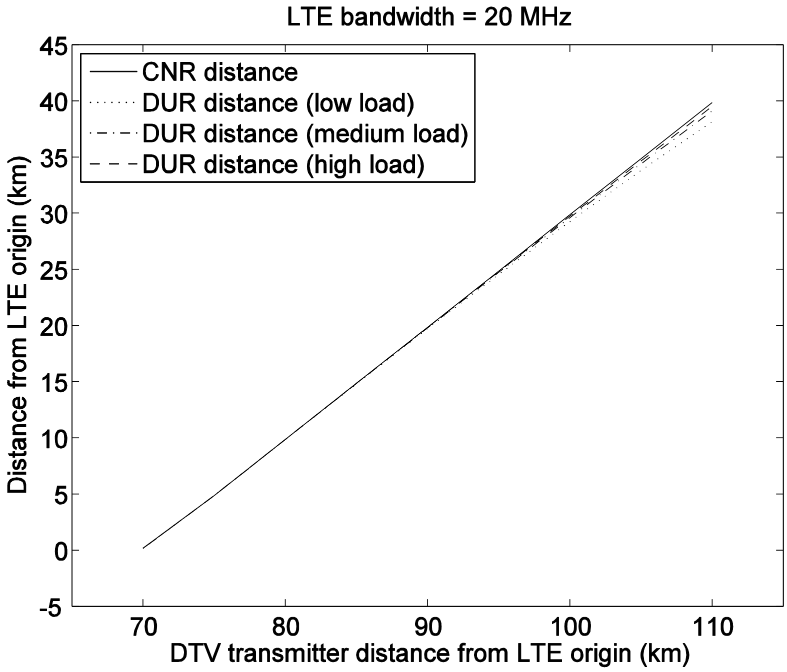


Figure 59. Protection of the primary DTV system (LTE bandwidth of 20 MHz).

Table 8. Protection distances from the border of the DTV coverage area.

		Load		
		Low	**Medium**	**High**
Bandwidth	**1.4 MHz**	9.82 km	14.82 km	19.83 km
	5 MHz	19.83 km	22.41 km	24.82 km
	20 MHz	24.81 km	29.81 km	34.79 km

Selected bandwidth and supported load

The amount of spectrum from the primary DTV band selected by the CM-SM is lower-bounded by the minimum traffic load (i.e., minimum number of UEs) that must be supported by the secondary LTE system and upper-bounded by the maximum aggregated interference level tolerable by the primary DTV system. As previously discussed, the reutilisation of a wider block of spectrum enables a larger traffic load to be supported but requires a larger separation between the DTV and LTE systems, and vice versa, in order to protect the primary system in terms of interference. Notice that power control and other Radio Resource Management (RRM) techniques are responsibility of the Cognitive Manager for Resource Management (CM-RM) and are therefore out of the scope of the CM-SM. From the point of view of the CM-SM, it is assumed that each UE has an allowed maximum transmission power limit, and traffic loads and resulting interference levels are taken into account in terms of the maximum number UEs simultaneously transmitting. Therefore, CM-SM decisions on the amount of bandwidth selected for secondary reutilisation can be taken based on the supported traffic load (in terms of the number of UEs) and the distance from the primary DTV system. Larger distances to the boundary of the coverage area for the DTV system enable the reutilisation of wider blocks of spectrum, when available, and increase the LTE system capacity.

DUR strip

The actual DUR boundary, i.e., the distance at which the minimum DUR is observed, is not constant but shows certain variation around an average value, within certain maximum and minimum values, thus leading to the existence of a *DUR strip*, as illustrated in Figure 60 for an LTE bandwidth of 1.4 MHz (similar trends were observed for other bandwidths). The existence of such DUR strip is mainly the result of the mobility of UEs (i.e., the instantaneous location of every UE) but it also depends on several RRM aspects such as the admission criterion and scheduling sequence (i.e., which UE transmits at every time instant), the power control method (i.e., the interference power from every active UE), etc. Since all these aspects determine the exact location of the DUR boundary and they vary over time, the DUR boundary varies over time as well. Therefore, the exact location of the DUR boundary depends on the considered time instant and the state of the aforementioned aspects. The behaviour of the DUR boundary can be characterised in a simple way by means of the minimum and maximum values within which it is confined, as illustrated in Figure 60. It is worth noting that the most likely value is the average DUR in the middle of the DUR strip (which is the parameter shown in Figure 57, Figure 58 and Figure 59), and the probability of a particular DUR value decreases as it is closer to the maximum or minimum values of the DUR strip.

Figure 61 shows the width of the DUR strip as a function of the distance between the DTV and LTE systems for an LTE bandwidth of 1.4 MHz (similar trends were observed for other bandwidths). As it can be observed, the width of the DUR strip increases as the separation between the systems increases and is higher for low loads. This is an aspect to be accounted for when determining the minimum distance between the DTV and LTE systems based on the DUR requirements.

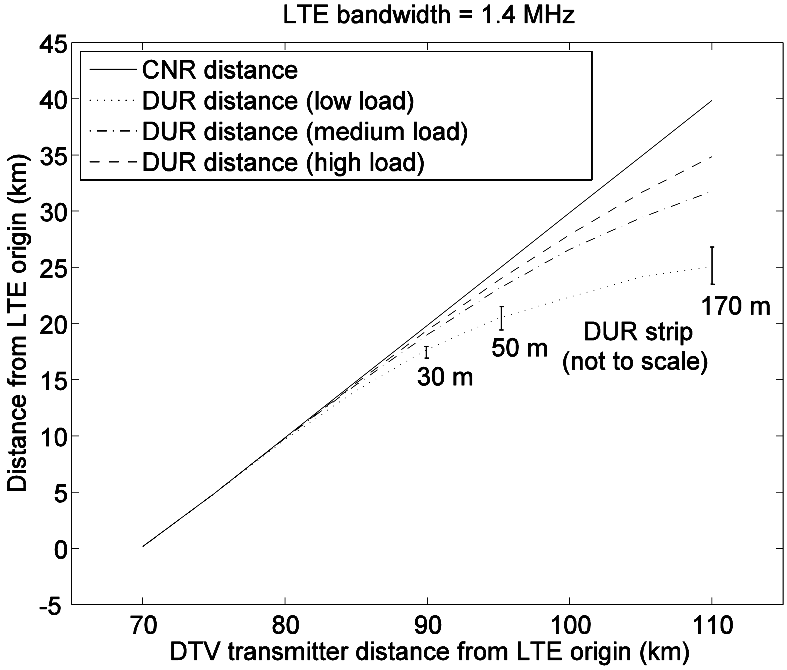


Figure 60. Illustration of the DUR strip effect (LTE bandwidth of 1.4 MHz).

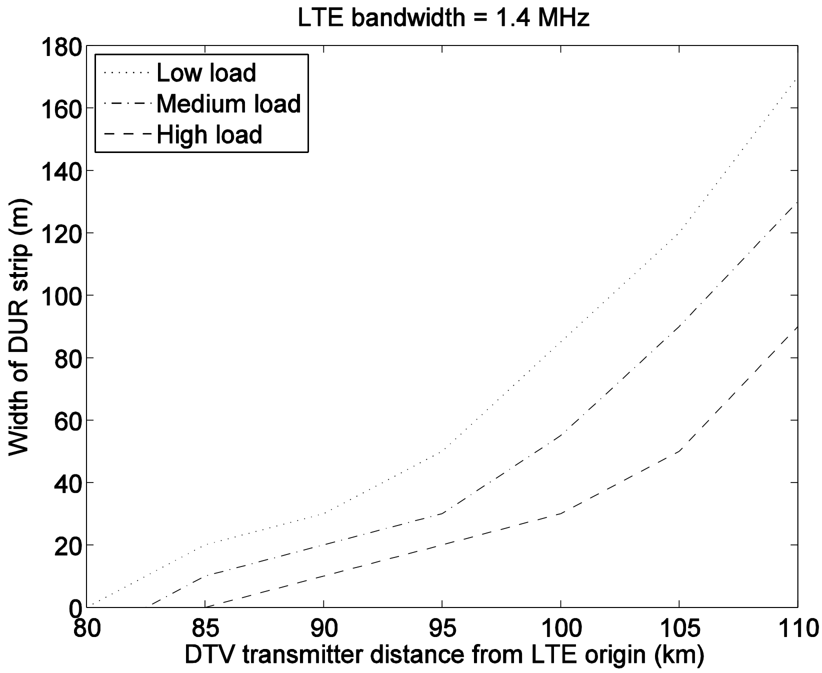


Figure 61. Width of DUR strip (LTE bandwidth of 1.4 MHz).

Summary of protection results

When deciding on the selection of DTV bands based on the protection of the primary DTV system it is important to take into account that:

- The minimum distance at which the DTV spectrum can be reused depends on the total aggregated interference of the UEs on the DTV receivers in the worst case, which increases:
    - When the number of simultaneous UE increases.
    - When the selected bandwidth increases, as this increases the maximum number of simultaneously supported UEs.
- The bandwidth of DTV spectrum that can be selected for secondary reuse is:
    - Lower-bounded by the minimum load level to be supported by the LTE system.
    - Upper-bounded by the maximum aggregated interference that can be tolerated by the DTV system, which in turn depends on the number of supported UEs and distance.

6.4.3 Performance of the secondary LTE system

After having analysed the protection of the primary DTV system in Section 6.4.2.1, this section analyses the performance of the secondary LTE system following a similar structure and based on the same aspects, namely the distance between the DTV and LTE systems, the selected bandwidth and the supported load of LTE users. The performance of the LTE system is mainly analysed in terms of throughput, although transmission error rates (e.g., BLER) are considered as well.

Impact of the distance between the DTV and LTE systems

The results obtained in Section 6.4.2.1 indicated that the protection of the primary DTV system is highly dependent on the distance between the DTV and LTE systems. The performance of the secondary LTE system, however, was observed to be rather unaffected by the distance between both systems. This is illustrated in Figure 62 and Figure 63 in terms of the user and sector throughputs (i.e., the average throughput experienced per UE and the average throughput aggregated over the UEs within the same sector, respectively). The results are shown for an LTE bandwidth of 20 MHz but similar trends were observed for other LTE bandwidths as well as for the rest of analysed performance metrics.

As it can be appreciated, the LTE system performance is not noticeably affected by the operation distance with respect to the DTV system. This can be explained by the fact that the protection of the primary system requires a minimum separation distance that results in low interference levels from the DTV system, thus leading to similar performance results at various DTV-LTE distances. As it can be noted in Figures A-1 and A-2 of [D6.5], the most relevant interference from the LTE to the DTV system occurs from transmitting UEs close to the border of the LTE coverage area to DTV receivers close to the border of the DTV coverage area. Therefore, the LTE-to-DTV interference occurs at the minimum separation distance between both systems, which explains the high impact of such distance on the protection of the primary system. On the other hand, the interference from the DTV system to the LTE system occurs from the DTV transmitter, which is far from the border of the DTV coverage area, to the LTE receivers (i.e., the eNodeBs), which are far from the LTE cell border as well. Therefore, the DTV-to-LTE interference occurs at much larger distances than the LTE-to-DTV interference, which explains the negligible impact of such distance on the observed LTE performance.

Impact of the selected bandwidth and supported load

While the operation distance between the DTV and LTE system cannot be considered as a crucial aspect in the uplink LTE performance, the selected bandwidth and supported load have a more relevant impact on the uplink performance. This is illustrated in

Table 9 and

Table 10 in terms of the user and sector throughputs respectively. For a constant operation bandwidth, higher loads result in a lower amount of available resources per user and therefore in a lower user throughput as expected. While in theory the overall sector throughput should remain constant for different load levels, in practice higher loads result in lower sector throughputs as well since the higher number of UEs leads to higher levels of interference in the system. In any case, the supported load level has a higher impact on the user throughput than in the sector throughput. On the other hand, increasing the selected bandwidth means increasing the amount of available resources for the LTE system and therefore in a clearly appreciable increase of the overall sector throughput and thus the overall uplink capacity. For a similar ratio of UEs per available bandwidth (i.e., load level), the user throughput should be similar for different operation bandwidths. However, higher operation bandwidths lead to a higher number of supported users thus leading to higher interference levels in the system, which explains the user throughput degradation as the selected bandwidth increases.

In terms of BLER, it was observed that the mean BLER averaged over all the UEs lies within the interval 0.15-0.17 regardless of the operation bandwidth and supported load level, which is slightly higher than the target BLER of 0.1 (see Table 5). The analysis of the obtained results indicated different causes for the various considered load levels.

For low and medium loads, it was observed that a low fraction of users were subject to strongly unfavourable radio propagation conditions (i.e., extremely low SNR values at the cell border) that resulted in erroneous transmissions despite using the lowest modulation and coding scheme and the highest transmission power. Although these users represent a small fraction (around 7% most of the times and never higher than 10%), the exceptionally high error rates of these users (above 90%) lead to the slightly higher average BLER of 0.15-0.17.

On the other hand, for high loads, it was observed the a larger fraction of users (around 60-70 %) experienced a wide variety of BLER values above the target value of 0.1 as a result of the increased interference levels in the system, thus leading a similar average BLER of 0.15-0.17.

Therefore, it is important to take into account that the supported load has not a significant impact on the average BLER but it does on the BLER distribution over the UEs.

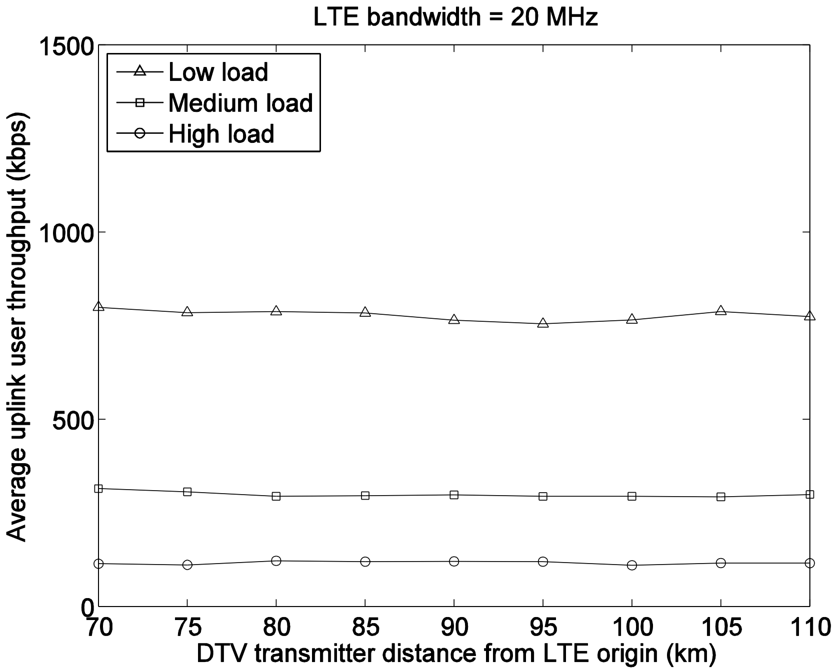


Figure 62. Performance of the secondary LTE system (LTE bandwidth of 20 MHz).

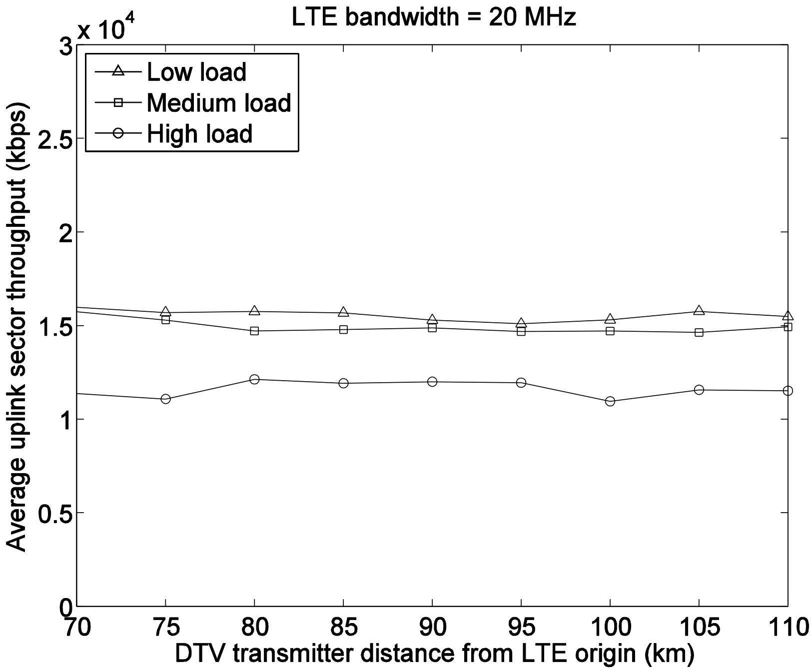


Figure 63. Performance of the secondary LTE system (LTE bandwidth of 20 MHz).

Table 9. Average LTE uplink user throughput.

		Load		
		Low	**Medium**	**High**
Bandwidth	**1.4 MHz**	941.67 kbps	318.74 kbps	132.32 kbps
	5 MHz	781.79 kbps	312.86 kbps	121.76 kbps
	20 MHz	777.68 kbps	298.44 kbps	116.02 kbps

Table 10. Average LTE uplink sector throughput.

		Load		
		Low	**Medium**	**High**
Bandwidth	**1.4 MHz**	0.97 Mbps	0.94 Mbps	0.79 Mbps
	5 MHz	3.91 Mbps	3.82 Mbps	3.04 Mbps
	20 MHz	15.5 Mbps	14.92 Mbps	11.60 Mbps

Summary of performance results

When deciding on the selection of DTV bands based on its impact on the LTE system performance it is important to take into account that:

- The distance between the DTV and LTE systems has not a relevant impact on the uplink performance of the LTE system. However, the operation bandwidth and supported load have an important impact of the LTE uplink performance.
- In terms of user throughput and overall system capacity:
    - Higher load levels for the same LTE bandwidth result in a degradation of both the user throughput and the overall system capacity (sector throughput), but the impact is much more significant on the former.
    - Larger LTE bandwidths for a similar ratio of UEs to available bandwidth result in an increased overall system capacity (sector throughput) but leads to some degradation of the throughput experienced by individual UEs (user throughput).
- In terms of BLER:
    - The mean BLER averaged over all UEs is not significantly affected by the operation bandwidth and supported load.
    - However, the supported load level has a relevant impact on the BLER distribution over the UEs:
        - For low and medium load levels, there is a low fraction of UEs that experience very high error rates as a result of unfavourable radio propagation conditions in particular cases (e.g., at the cell border).
        - For high load levels, there is a much larger fraction of UEs that experience a wide variety of error rates above the desired target value.

6.4.4 Efficiency of spectrum utilisation

This section discusses how efficiently the spectrum can be exploited by the LTE system depending on the distance between the DTV and LTE systems, the selected bandwidth and the supported load of

LTE users. Efficiency is analysed in terms of the bandwidth utilisation parameter and the spectral efficiency parameter, both defined in Section A.1.3 of [D6.5], which are respectively shown in Table 11 and Table 12. Similarly to the LTE performance metrics considered in Section 6.4.3, the efficiency metrics considered here do not show a relevant difference for the considered operation distance between the DTV and LTE systems, but show certain variations depending on the operation bandwidth and supported user load. In general, it can appreciated that when the number of simultaneous UEs increases (i.e., for higher loads and/or operation bandwidths) the overall efficiency decreases, for both metrics, as a result of the resulting increased interference levels. Notice that the convenient choice in terms of spectral efficiency would be the selection of narrower operation bandwidths. However, this would not be a feasible alternative when the secondary user demands require larger bandwidths.

Table 11. Efficiency of spectrum utilisation (bandwidth utilisation).

		Load		
		Low	**Medium**	**High**
Bandwidth	**1.4 MHz**	0.87	0.85	0.77
	5 MHz	0.86	0.84	0.76
	20 MHz	0.78	0.77	0.69

Table 12. Efficiency of spectrum utilisation (spectral efficiency).

		Load		
		Low	**Medium**	**High**
Bandwidth	**1.4 MHz**	0.78 bit/sec/Hz	0.76 bit/sec/Hz	0.61 bit/sec/Hz
	5 MHz	0.77 bit/sec/Hz	0.75 bit/sec/Hz	0.58 bit/sec/Hz
	20 MHz	0.69 bit/sec/Hz	0.67 bit/sec/Hz	0.57 bit/sec/Hz

6.4.5 Concluding remarks

The adequacy of selecting a DTV band for secondary reutilisation by LTE uplink users depends on whether the secondary operation results in acceptable interference levels on the primary band and the resulting LTE performance can satisfy the demands of the secondary users. This section has presented a simulation study aimed at analysing and determining the conditions under which the coexistence of both systems in the DTV band is feasible along with the resulting technical implications.

The study has been based on three variables, namely the distance between the DTV and LTE systems, the opportunistically reused bandwidth and the supported load of secondary users, which constitute a simple set of design parameters. It is important to notice that not all these variables are under the control of the CM-SM and therefore cannot be modified by the CM-SM decisions, such as the distance between the DTV and LTE systems or the supported load, which is the responsibility of the CM-RM and the employed RRM techniques (admission control, load control, power control, scheduling, etc.). Nevertheless, there are two important variables the CM-SM can act over: a) the binary decision on selecting the DTV band for reutilisation by the LTE system, and b) the amount of primary bandwidth that can be reused by the LTE system. The decision can be made based on the variables considered in this study. For instance, the selection of DTV band may depend on the particular distance of the LTE system, the current (or predicted) load and the capability of the expected performance to satisfy the user demands. If the DTV band is selected, the width of the spectrum block to be reused can be tuned as well based on the DTV-LTE intersystem distance and readjusted based on load variations of the LTE system load in order to meet the user performance demands while satisfying the interference requirements.

In this context, this study has identified the qualitative trends of relevant parameters and has provided quantitative results for the considered metrics and system aspects that could be taken as a reference point for system design and evaluation. These quantitative/qualitative reference results can be employed to help decision entities such as the SSE to make decisions on the adequacy of selecting DTV bands for extension of the LTE uplink system and predict the resulting consequences of such decisions in terms of the protection of the primary system, the performance of the secondary system and the overall efficiency of spectrum utilisation.

7 References

7.1 QoSMOS deliverables

[D2.1] "Initial description of system architecture options for the QoSMOS system", FP7-ICT-2009-4/248454 QoSMOS project deliverable D2.1, Apr. 2010.

[D2.2] "System architecture options for the QoSMOS system", FP7-ICT-2009-4/248454 QoSMOS project deliverable D2.2, Dec. 2010.

[D2.3] "System specification and evaluation criteria", FP7-ICT-2009-4/248454 QoSMOS project deliverable D2.3, Nov. 2011.

[D5.1] "Initial description of framework for supporting QoS and handling mobility", FP7-ICT-2009-4/248454 QoSMOS project deliverable D5.1, Jul. 2010.

[D5.2] "Final framework description, preliminary cognitive manager structure and first mechanisms for QoS support", FP7-ICT-2009-4/248454 QoSMOS project deliverable D5.2, Jul. 2011.

[D5.3] "Final description and specification of cognitive manager and corresponding QoS support mechanisms with performance evaluation", FP7-ICT-2009-4/248454 QoSMOS project deliverable D5.3, Jul. 2012.

[D6.1] "Initial description of spectrum management framework, requirements analysis and approach selected", FP7-ICT-2009-4/248454 QoSMOS project deliverable D6.1, Jun. 2010.

[D6.2] "Description of context filtering, aggregation and communication functions of the spectrum management framework", FP7-ICT-2009-4/248454 QoSMOS project deliverable D6.2, Mar. 2011.

[D6.3] "Initial description of cognitive and opportunistic functions of the spectrum management framework", FP7-ICT-2009-4/248454 QoSMOS project deliverable D6.3, Aug. 2011.

[D6.5] "Specification of cognitive and opportunistic functions of the spectrum management framework", FP7-ICT-2009-4/248454 QoSMOS project deliverable D6.5, Aug. 2012.

7.2 Other references

[ARIBB21] ARIB, "Receiver for digital broadcasting", ARIB STD-B21, Ver. 4.7, December 2008 (in Japanese).

[Blu01] D. Blumenfeld, *Operations research calculations handbook*, CRC Press, 2001.

[BOYD04] S. Boyd and L. Vandenberghe, Convex optimization, Cambridge University Press, 2004.

[Cai03] X. Cai and G. Giannakis, "A two-dimensional channel simulation model for shadowing processes", IEEE Transactions on Vehicular Technology, vol. 52, no. 6, November 2003, pp. 1558-1567.

[CelEtal2011] U. Celentano, B. Bochow, C. Lange, F. Noack, J. Herrero, B. Cendón, O. Grøndalen, V. Mérat, C. Rosik, "Flexible architecture for spectrum and resource management in the whitespace", *Proc. Int. Symp. Wireless Personal Multimedia Commun.* (WPMC 2011), Brest, France, 3-7 Oct 2011

[Cic98] D. J. Cichon and T. Kürner, “Propagation prediction models (Digital mobile radio towards future generation systems, chapter 4),” COST231, Tech. Rep., Nov. 1998.

[Cla05] H. Claussen, “Efficient modelling of channel maps with correlated shadow fading in mobile radio systems”, in proceedings of the IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2005), vol. 1, 2005, pp. 512-516.

[Clarke68] R. H. Clarke, “A statistical theory of mobile-radio reception,” Bell System Technical Journal, vol. 47, pp. 957–1000, 1968.

[Cost231] “COST 231. digital mobile radio towards future generation systems. final report,” E. Demosso, Ed. Brussels, Belgium, EU: European Commission, 1999. [Online]. Available: http://www.lx.it.pt/cost231

[ECC11] “Technical and operational requirements for the possible operation of cognitive radio systems in the ‘white spaces’ of the frequency band 470-790MHz,” European Communication Committee, Tech. Rep. ECC Report 159, Jan. 2011.

[ECC159] ECC Report 159, “Technical and operational requirements for the possible operation of cognitive radio systems in the ‘white spaces’ of the frequency band 470-790 MHz”, January 2011.

[EESM03] “Effective SIR computation for OFDM system-level simulations,” 3GPPTSG-RAN-1, Tech. Rep. RI-03-1370, 2003.

[Egl57] John J. Egli, “Radio propagation above 40 MC over irregular terrain”, Proceedings of the IRE (IEEE), vol. 45, no. 10, pp. 1383–1391, Oct. 1957.

[FAB07] F.J. Fabozzi, et al., Robust Portfolio Optimization and Management, Wiley Finance, 2007.

[FCC04186] FCC, “In the matter of unlicensed operation in the TV broadcast bands”, ET Docket No. 04-186, Notice of Proposed Rulemaking, May 2004.

[Golmie09] N. Golmie, R. Rouil, D. Doria, X. Guo, R. Iyengar, S. Kalyanaraman, S. Krishnaiyer, S. Mishra, B. Sikdar, R. Jain, R. Patneyand, C. So-In, and S. Parekh. (2009) WiMAX forum system level simulator NS-2 MAC+PHY add-on for WiMAX (IEEE 802.16). [Online]. Available: http://code.google.com/p/ns2-wimax-awg/

[Hat80] M. Hata, “Empirical formula for propagation loss in land mobile radio service”, IEEE Transactions on Vehicular Technology, vol. VT-29, no. 3, pp. 317-325, Aug. 1980.

[Huffman] http://www.its.bldrdoc.gov/resources/radio-propagation-software/itm/irregular-terrain-model-%28itm%29-%28longley-rice%29.aspx

[ieee80216] Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands and Corrigendum 1, IEEE Std. 802.16e, Apr. 2006.

[ieee80222] IEEE Standard for Information Technology–Telecommunications and information exchange between systems Wireless Regional Area Networks (WRAN)–Specific requirements Part 22: Cognitive Wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Policies and Procedures for Operation in the TV Bands, IEEE IEEE Std 802.22-2011, 2011.

[ieee802221] IEEE Draft Standard for Information technology Telecommunications and information exchange between systems Local and metropolitan area networks Specific requirements Part 22.1: Standard to Enhance Harmful Interference Protection for Low Power Licensed Devices Operating in TV Broadcast Bands, IEEE Std. P802.22.1 Draft 8, Jul. 2010.

[ITU1368] ITU-R BT.1368, "Planning criteria for digital terrestrial television services in the VHF/UHF bands".

[ITU1546] ITU-R P.1546-3, "Method for point-to-area predictions for terrestrial services in the frequency range 30 MHz to 3 000 MHz".

[ITUR97] "Guidelines for evaluations of radio transmission technologies for IMT-2000," ITU ITU-R, Tech. Rep. M.1225, 1997.

[Jeon08] W. S. Jeon, D. G. Jeong, J. A. Han, G. Ko, and M. S. Song, "An efficient quiet period management scheme for cognitive radio systems," IEEE Trans. Wireless Commun., vol. 7, no. 2, pp. 505–509, Feb. 2008.

[JOU03] W. Jouini, C. Moy, and J. Palicot, "Decision making for cognitive radio equipment: analysis of the first 10 years of exploration," EURASIP Journal on Wireless Commun. and Networking, January, 2012.

[KON05] A. Konak, D.W. Coit, and A.E. Smith, "Multi-objective optimization using genetic algorithms: A tutorial," Reliability Engineering and Safety, Elsevier, 2005, pp. 992-1007.

[Lee11]W.-Y. Lee and I. F. Akyildiz, "A spectrum decision framework for cognitive radio networks," IEEE Transactions on Mobile Computing, vol. 10, no. 2, pp. 161–174, February 2011.

[Lop11] M. López-Benítez and F. Casadevall, "Empirical time-dimension model of spectrum use based on discrete-time markov chain with deterministic and stochastic duty-cycle models," IEEE Transactions on Vehicular Technology, vol. 60, no. 6, pp. 2519–2533, July 2011.

[Lop13] M. López-Benítez, "Cognitive radio", chapter 13 in *Heterogeneous cellular networks: Theory, simulation and deployment*, Cambridge University Press, 2013 (in press).

[LU12] L. Lu, X. Zhou, U. Onunkwo, and G.Y. Li, "Ten years of research in spectrum sensing and sharing in cognitive radio," EURASIP Journal on Wireless Commun. and Networking, January 2012, Open Access.

[MIT93] J. Mitola, "Software radios: survey, critical evaluation and future directions", IEEE Aerosp. Electron. Syst. Mag. 8(4), 25–31 (1993).

[Ofcom09] "Digital dividend: Cognitive access, statement on licence-exempting cognitive devices using interleaved spectrum," Ofcom, Tech. Rep., Jul. 2009.

[Oku68] Y. Okumura, E. Ohmori, T. Kawano, and K. Fukuda, "Field strength and its variability in VHF and UHF land-mobile radio service," Review of the Electrical Communications Laboratory, vol. 16, no. 9-10, pp. 825–873, Sep. 1968.

[PAL06] R. Pal, D. Idris, K. Pasari, and N. Prasad, "Characterizing Reliability in cognitive radio networks," First Int. Symposium on Applied Sciences on Biomedical and Commun. Tech., 2008.

[Wan93] Y. H. Wang, "On the number of successes in independent trials," Statistica Sinica, vol. 3, no. 2, pp. 295–312, Jul. 1993.

[Wellens09] M. Wellens, J. Riihijärvi, and P. Mähönen, "Empirical time and frequency domain models of spectrum use," Elsevier Physical Communication Journal, vol. 2, no. 1–2, pp. 10–32, 2009.

[Wimax06] Mobile WiMAX–part 1: A technical overview and performance evaluation, Aug 2006. [Online]: http://www.wimaxforum.org/resources/documents/marketing/whitepapers

[WimaxFor] WiMAX forum website. [Online]. Available: http://www.wimaxforum.org/

[WIND112] WINNER Del. D1.1.2. Final link- and system-level channel models, http://www.ist-winner.org/.

[WIXSLS] WiMAX system level evaluation methodology, WiMAX Forum, V.0.0.1, 2006.