CEPT ECC
Electronic Communications Committee

ECC Report 181

IMPROVING SPECTRUM EFFICIENCY IN THE SRD BANDS

Approved September 2012

0 EXECUTIVE SUMMARY

Considering the development of SRDs applications, the development of new technologies and the experience gained toward the deployment of SRDs equipment, this ECC Report investigates possible ways of improving spectrum efficiency in the frequency bands used by Short Range Devices (SRDs).

It is important to distinguish between spectrum occupancy and spectrum efficiency. The value of using a particular part of spectrum comes from the utility it provides to users, which is not necessarily the same as the data traffic. A distinction should be made between the concepts of Single system Absolute spectrum Efficiency (SAE), which is based on the raw data transmitted, and Group Spectrum Efficiency (GSE), which is closer to the broader utility or service provided.

One conclusion is that some SRDs operating in “exclusive” bands might indeed benefit if those bands were to be low occupancy so that devices relying on access by duty cycle (DC) limits alone can operate effectively.

At the same time, it would be wasteful and inefficient to operate all the spectrum identified for SRDs in this way. In other sub-bands, whenever there is demand, occupancy and throughput levels will have to rise. Regulators and industry will have to devise means of achieving this. Since basic DC is only effective as a sharing mechanism up to relatively low levels of occupancy and throughput, this may require the introduction of more advanced sharing mechanisms.

A second conclusion is that different sub-bands should be optimised for different communication needs. Users of the SRD bands have a variety of needs and different criteria for a successful service, and this should be recognised in the management of the spectrum identified for SRDs.

Access mechanisms and spectrum management should be based on sound technical foundations – the equivalent of “evidence based” rule making. This report initiated some work relating to the derivation of the technical parameters and spectrum management for a given SRD sub-band. This work should be continued and extended.

In addition, the following was concluded:

- **Spectrum occupancy** is the parameter most visible to observers and monitors of the spectrum. Section 3.11 shows the relationship between occupancy levels and access techniques. For monitoring purposes on an application level the distinction between spectrum occupancy and channel occupancy needs to be made. In most general cases this is not necessary but when investigations are made in specific sub-bands, especially when considering application of new spectrum efficiency metrics proposed in this report and some advanced mitigation mechanism, this may be relevant.
- There is a need to optimise some of the SRD spectrum to achieve **high reliability use.** The amount of spectrum required for such usage might be relatively small.
- The aims of this report are entirely consistent with the principle of **application neutrality** set out in CEPT Report 14 (section 2.7 [19]). For instance, it may be better to designate a sub-band not for inherently safety related critical alarm systems, but instead as a sub-band where high reliability, low latency, low duty signalling is always possible. This is a clearer path for regulators to follow in order to provide a better service both to alarm manufacturers and to alarms users while remaining application neutral, thus not preventing further innovation in the given sub-band.

  The principle of application neutrality means the end of segregation by application – whereby sub–bands were designated exclusively to a particular application, primarily within the European SRDs generic frequency ranges. In order to preserve technical efficiency, a suitable replacement could be partitioning of the bands based on technical objectives – e.g., sub-bands for high reliability, for low latency, for high throughput. However, this may lead to more detailed definition being needed in describing the technical requirements and this may lead to a reduction in technology neutrality if not performed properly.

  At the same time it is worth noting that sometimes an SRD application may have very clear specific technical characteristics that may employ opportunistic sharing techniques to enable it to politely operate

within spectrum allocated to radiocommunication services that otherwise would be interfered by generic SRDs. This may represent higher spectrum use efficiency, beneficial to both uses.

- The principle of **technology neutrality** is more difficult to realise and therefore may not always be realised by regulation without sacrificing spectrum use efficiency. It should be still possible to frame regulations so that, for instance, either analogue or digital modulation is allowed or a range of bandwidths is possible. In most cases, however, it is necessary to set specific technical conditions to allow successful sharing, so technology neutrality is at odds with spectrum efficiency. There may be a case for a "sandpit" area, akin to the concept of bands identified for ISM, where technology neutrality is applied as far as possible, to assist the emergence of new technologies.
- **Listen Before Talk (LBT)** is well known mitigation technique in the SRD field whereby the transceiver performs sensing of the channel before each packet transmission. This report carried out an extensive modelling with the aim of quantifying the precise benefits of LBT in various sharing scenarios. It was shown that the LBT is not a "silver-bullet" in that it has its limitations and shortcomings, most notably as described by the "hidden/exposed node" problems.

  The report considered the benefits of two related concepts, namely those of **Carrier Sensing (CS) and Collision Detection (CD)**, known as part of so called Aloha channel access protocol. CD is the detection of a collision after the event. This happens in all systems that work at the higher levels of the OSI model, such as analysis of message success rates. CS operates before the transmission with the aim of preventing collisions. It thus closely resembles LBT and sensing elements of more advance frequency agility mechanisms such as DAA, DFS and AFA. The notable conclusion of this report is that LBT and CS/CD require further studies in anticipation that some kind of hybrid mechanisms, involving both CD and CS aspects, would be necessary if wanting to achieve high levels of throughput and spectrum use efficiency in high channel occupancy scenarios.
- The traditional generic **FHSS** may be only truly effective in scenarios with lower levels of *band* occupancy; basically it spreads the traffic over a wide spectrum to reduce the per-channel traffic to low levels. **Hybrid or adaptive FHSS** need further study to see how effectively it overcomes the limitations of generic FHSS and what other types of spectrum access mechanisms it can most optimally share with.

  Noting the nature of FHSS as band-level, not channel-level access mechanism, it may be suggested that regulations should not make special provisions for FHSS, but should instead apply per-channel access rules taking into account the correlation of channel transmissions in the spatial domain.
- **Advanced technologies (FDMA, CR…)** for spectrum access may have received less attention and analysis to date in the SRD community than time domain techniques, mostly because of the higher involved complexities, including some of them needing central controlling entity with degree of "intelligence" etc., but they should be studied further as potential techniques for high occupancy, high traffic sub bands.
- It may be possible to achieve spectrum use efficiency gains and overall spectrum capacity increase by **combining longer- and shorter-range deployment scales** (still in the overall context of limited SRD range). This would resemble the principle of combined deployment of umbrella macro-cells and pico cells in the same area, even on the same channel. Systems operating with differing operating powers within sensible limits and ranges are able to effectively co-exist, thereby significantly increasing medium utilisation. The success of this mechanism depends on the typical usage scenarios and user expectation for the applications vying for co-existence. The achieved group spectrum efficiency depends on the choice of spectrum access parameters.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Explanation
AFA	Adaptive Frequency Agility
ACK	Acknowledgement
APC	Adaptive Power Control
CD	Collision Detection
CEPT	European Conference of Postal and Telecommunications Administrations
CS	Carrier Sensing
CSMA	Carrier Sensing Multiple Access
CR	Cognitive Radio
DAA	Detect and Avoid
DC	Duty Cycle
DFS	Dynamic Frequency Selection
DP	Combination of Data volume and Power consumption
dRSS	desired Received Signal Strength
DS	Combination of Data volume and size issues
DSI	Digital Spectrum Investigation
DSSS	Direct Sequence Spread Spectrum
EC	European Commission
ECC	Electronic Communications Committee
e.i.r.p.	Equivalent Isotropic Radiated Power
ERP	Effective Radiated Power
FHSS	Frequency Hopping Spread Spectrum
RFID	Radio Frequency Identification
GSE	Group spectrum efficiency
ITU-R	International Telecommunication Union-Recommendation
LBT	Listen Before Talk
LC	Combination of Latency and Cost issues
LDC	Low Duty Cycle
LO	Local Oscillator
LP	Level Probing
LS	Combination of Latency and Size issues
M2M	Machine to Machine
MCL	Minimum Coupling Loss
OSI	Open Systems Interconnection
RAKE	Radio Activated Key Entry
RC	Combination of Reliability and cost issues
RKE	Remote Keyless Entry
RP	Combination of Reliability and Power consumption
RSPG	Radio Spectrum Policy Group
RS	Combination of Reliability and Size issues
SAE	Single system Absolute Efficiency
SAW	Surface Acoustic Wave
SGRE	Single system in a Group Relative Efficiency
SNR	Signal to Noise Ration

SRD	Short Range Device
SRE	Single system Relative Efficiency
SUE	Spectrum usage Efficiency
TPMS	Tyre Pressure Monitoring System
TRP	Total radiated Power
TCXOs	Temperature Compensated Oscillators
ULP	Ultra Low Power
UWB	Ultra Wide Band
VLDC	Very Low Duty Cycle
WGFM	Working Group Frequency Management
WT	Wanted Transmitter

1 INTRODUCTION

The 433 MHz band was in use by SRDs prior to the allocation of the 868-870 MHz band (subsequently enlarged to 863-870 MHz). A number of stakeholders and manufacturers reported difficulties with using the 433 MHz band because of the open access nature and the presence of other high power, non SRD, devices. A strong preference was expressed for moving certain applications to 868 MHz where duty cycle limits were imposed. Similar sentiments have been expressed about the open access nature of the 2.4 GHz band. The conclusion is that significant parts of industry prefer to see spectrum access control methods in place.

ECC Report 37 [1] considered the potential to expand the use of SRDs within the band 863-870 MHz as originally proposed in the DSI Phase III Consultation and the CEPT Strategic Band Plan for this specific frequency band. Particular attention has been given to the use of new techniques, which could increase the number of users able simultaneously to operate within this band such as LBT and the effect of introducing spread spectrum techniques (DSSS (Direct Sequence Spread Spectrum) and FHSS (Frequency Hopping Spread Spectrum)). ECC Report 37 [1] provided the technical background for the regulatory framework in the frequency range 863-870 MHz as given in Annex 1 to ERC/REC 70-03 [2].

The SRD bands accommodate a wide variety of different users and applications. They provide a valuable economic service, and the use of these bands is expected to grow. Industry fears congestion in some bands and difficulties to obtain required quality, capacity and reliability levels of spectrum access for their SRD applications and is therefore requesting more spectrum and/or optimisation of existing spectrum bands identified for SRDs. However this is also offset by an observation (particularly after monitoring campaigns [3] as described in section 5.1), that the actual use of existing bands is not homogenously distributed in that many localised measurements show occupancy levels well below 100%. These measurements were performed in chosen hot spots. Where the occupancy is high, it is almost always because a single application/user is dominating the channel. It is also worth noting here that the occupancy itself is somewhat lopsided term, defined as the overall number of active transmissions observed in a given channel over certain time. It therefore may not by itself represent the actual number of SRD devices that are understood to be “using” this channel. This is because by nature of its operation many SRD devices spend a lot of time in dormant state, while requiring nearly instant access to radio channel when activated (consider all kinds of alarms here, car keys etc.).

In that respect it is worth noting that industry generally indicated the existing core SRDs spectrum 863-870 MHz is not overcrowded at the moment, but that high growth from several market sectors, such as Smart Metering/Smart Grid, RFID, home automation, industrial control (machine-to-machine) and alarms, is expected. CEPT ECC WG FM (May and October 2011 meetings) endorsed the conclusions of April 2011 Workshop dedicated to SRD developments and agreed a Roadmap on discussing possible future UHF spectrum needs for SRD use.

The above described situation of low occupancy observed in SRD bands has variously been interpreted as either a lack of demand, inefficient use of the spectrum, or a failing of intra-SRD sharing mechanisms. This report seeks to identify the conditions under which these assumptions are true or not. An objective way to do that is to define spectrum efficiency not as a function of occupancy, but as a function of mutual coexistence between various types of SRD applications/devices.

SRD are usually used under a general authorisation regime, where no individual permit or licence is required for operation. Therefore any amendment to SRD regulations must be mindful of the pre-existing SRDs that might be already deployed in the field subject to previous authorisation conditions of use. This means that any regulatory change intended to improve spectrum efficiency must be made in a way that allows the existing users of spectrum to continue enjoying the QoS they have been accustomed to, while allowing wider access to spectrum. Finally, it is also essential to allow manufacturers to evolve production and supply products built to comply with new/amended regulations over a reasonable time period.

A way to achieve those conflicting objectives is by creating licence exempt spectrum access rules with minimum and appropriate technological restrictions, in such a way that the QoS for all existing and predicted future applications can be achieved. Application neutrality is therefore a desirable overall aim.

It is very clear from SRD industry comments [4] that there is continuing widespread support for sub-bands with limited operational restrictions such as duty cycle limits and/or other defined SRDs spectrum access techniques, and that such restrictions are preferred to a general and open single designation without any defined spectrum access techniques.

However, there is an expressed preference by the EC, RSPG and administrations in CEPT for spectrum access regulations to be application neutral and technology neutral supporting the continuous process of development and innovation going on in the area of SRDs. This is for example expressed in European Commission Directive 2009/140/EC [5], albeit not taking into account the effect of this on spectrum efficiency. A balance between spectrum efficiency and technology neutrality needs to be established and is generally already taken care for in standards. Therefore, this report tried to have a close look at what essential minimum technical parameters need to be expressed in regulation and standards, noting that a too simple regulatory framework could create regulatory uncertainty and have a negative impact on spectrum efficiency.

The report attempts to consider the definition of “the terms spectrum occupancy” and “spectrum efficiency” in this context and to analyse ways in which spectrum efficiency might be measured or calculated.

This report is concerned with SRDs operating in a public shared environment. Controlled environments are outside the scope of the current work, although some of the same considerations of spectrum efficiency and sharing techniques will apply there too. The report provides an overview of the various sharing technologies which could be used in the SRDs bands.

2 DEFINITION OF SPECTRUM EFFICIENCY AND OTHER BASICS

2.1 MEANING OF SPECTRUM EFFICIENCY

There is a general "common sense" understanding of spectrum efficiency, and Article 3.2 of the R&TTE Directive [6] makes it a requirement without actually defining it.

Most radio professionals will recognise and agree on "inefficiency" when they see it. For instance, carriers left on without modulation to preserve a channel, transmitters or receivers with excessive bandwidth, and large amounts of dead airtime are all seen as inefficient use of the spectrum. In many of these cases it is possible to construct technical and/or economic arguments for doing it that way, and to claim that, taking other factors into account, the alternatives are worse. These arguments cannot be dismissed out of hand, but neither do they change the fact that these are situations where improvement is desirable.

Spotting inefficiency is one thing; defining efficiency is quite another, and a lot depends on the context and point of view.

Article 3.2 of the R&TTE Directive looks at the situation from the point of view of one piece of equipment and requires that it "uses the spectrum efficiently". The intent of this could be expressed as:

> "use no more of the resource than is reasonably necessary."

It should be understood that the resource in question is not simply bandwidth, but a complex combination of factors such as bandwidth, time and geographic footprint.

The idea of using the minimum amount of resource is useful, but the point of view of this study is not a single piece of equipment, but rather the spectrum access regulations and how they can be optimised so as to allow many users to coexist.

Spectrum efficiency from a purely technical point of view can be derived from spectrum utilisation which is well defined in Recommendation ITU-R SM.1046-2 [7]. This is discussed further in section 2.5 below.

Recommendation ITU-R SM.1046-2 [7] makes the point that calculations using specific definitions of efficiency, throughput etc, should only be used to compare similar systems. This can make it difficult to apply the concept directly to the SRD bands, where a variety of different applications share the same spectrum.

2.2 THE IMPORTANCE OF CONTEXT

Besides spectrum utilisation we need also to consider the useful effect obtained with the communication system in question. For example in the case of sending a large stream of data from point to point, a measure of this useful effect would be:

> "bits/sec/Hz"

In broadcasting, suitable measures might be:

> "(bits/sec/Hz) x (number of listeners)", or,
> "(bits/sec/Hz) x (area covered)".

When the traffic is short bursts rather than continuous data, it may be more appropriate to work in terms of messages sent rather than data rate. This suggests a measure such as:

> "(messages/minute/Hz) x (number of users per km^2)".

These measures could then be further adjusted for factors such as power and cost.

What the examples above show is that the definition of spectrum efficiency will be different in different contexts.

The SRD bands accommodate a variety of applications and technologies. To define spectrum efficiency in terms of only one application would be unfair. It would not even be correct to define it in terms of a weighted combination of measures for each user or application. It is shown later that different applications have such different requirements that the measures would not be equivalent.

2.3 THE UNDERLYING AIM

While a definition of spectrum efficiency itself in the context of the SRD bands is difficult, the underlying aim of improving spectrum efficiency is more easily defined. The underlying aim can be stated as:

> The aim of improving spectrum efficiency in the SRD bands is to minimise the adverse effects and maximise overall throughput when large numbers of different types of user share the same frequency space. The often used term frequency space may be considered as the combination of coverage, usage in time and usage in frequency of a device. This is not a simple multiplication of these factors but the interaction of these usage patterns with the usage pattern of another device. Section 2.5 covers this in more detail.

And this can be expressed as two complementary aims:

> To minimise the spectrum allocation needed to satisfactorily accommodate large numbers of different types of user, or,
>
> To maximise the number and variety of users that can be satisfactorily accommodated in a given spectrum allocation.

The key word in these sentences is “satisfactorily”; the exercise must be accomplished to the reasonable satisfaction of all, and in an equitable fashion. For instance in ECC Report 37 [1], it was argued that in cases of extreme congestion, it was better if all users experienced graceful degradation than some users were arbitrarily excluded. The idea is to satisfy as much of the demand as possible in as fair a way as possible.

The effects of extreme congestion may be:

1. Catastrophic failure, or gridlock, in which all users lose most or all service.
2. Exclusion of some users, or lockout, in which the spectrum is still used but while some users receive normal service, others receive none.
3. Graceful degradation, in which all users receive a reduced service.

Sharing techniques which lead to effect 1 should be avoided, or at least applied with care. The distinction between 2 and 3 may seem small for systems claiming binary behaviour[1], until it is realised that binary means the service is either above or below a threshold of acceptability. Outcome 2 then means that as congestion increases there is an increasing probability of no service. Outcome 3 means the service level falls gradually until it reaches the threshold. Outcome 3 is generally considered to be the preferred objective of spectrum access regulation.

A further point is made that when a band is not congested, the users should be able to use the resource and not be constrained by limits designed for the congested case.

2.4 GENERAL REQUIREMENTS APPLYING TO SRDs

Investigations about spectrum efficiency need to consider the requirements of available and planned SRD applications. For example, applications are anticipated in, remote control, metering, distributed sensor networks, control loops for energy and alarms, voice/audio, RFID, healthcare applications and automotive applications. On the other hand, some industrial control systems are not considered as a specific application in this report. This list therefore does not contain all possible categories where developments are expected but gives an indication of the tremendous expected growth of SRD applications in general.

1 Many systems claim a binary distinction between “working” and “not working”. The meaning of this is actually that there is an acceptable threshold in a continuum (of latency, BER, etc.) and the service is either above or below the threshold.

The following main operational requirements should be considered:

- reliability
- latency
- data volume.

In addition to these there are also constraints on the means to achieve these requirements or make the product successful. The following list is not exhaustive, but gives a good impression of the main restrictions on an SRD:

- power consumption
- appropriate cost level
- size of the device

Using these categories a matrix may be constructed pointing out the combination of properties. Each combination reflects a set of technical parameters. The matrix in the next figure is a simplified example describing a particular application.

Table 1: Example of objectives vs. constraints in design of an SRD application

Application requirements	Constraints applying to this application			
	Reliability (R)	**RP**	RC	**RS**
	Latency (L)	LP	LC	LS
	Data volume (D)	DP	DC	DS
		Power consumption (P)	**C**ost level (C)	**S**ize of the device (S)

E.g., RP in the table means the combination of Reliability and Power consumption, RS the combination of Reliability and Size, and so on.

It can be seen that a particular considered SRD application has requirements **RP** and **RS**, reliability is required and there are power consumption and size issues. Another application with the same requirements RP+RS may share the same spectrum since the same technical solutions probably can be used for both applications.

A matrix like this, possibly expanded with more restrictions such as receiver capabilities, typical link budget, modulation scheme, etc. could be used to devise sharing schemes and assigning frequency bands on an application neutral basis.

Note that this scheme only depicts technical parameters and does not take production volume of devices and the possible typical geographical separation of application categories into account. The latter could provide a sharing possibility even if the scheme shows incompatibility.

2.5 MEASUREMENT OF SPECTRUM EFFICIENCY

Spectrum utilisation is defined as the product of the frequency bandwidth, the geometric (geographic) space, and the time denied to other potential users:

$$U = B \cdot S \cdot T$$

where:

B: frequency bandwidth
S: geometric space (usually area) and
T: time.

It may be noted that T is not equal to the transmit time of the device but equal to the time restrictions a device is imposing on all other users. Similar arguments are true for the frequency bandwidth and geometrical space factors. Since all mitigation techniques limit one or more of the three parameters B, S or

T to allow others to use the spectrum a mitigation technique can be therefore considered a spectrum utilisation limiting technique.

Such a mitigation technique may be primitive, simply restricting its spectrum utilisation by a fixed amount and in a fixed manner. It could also be more advanced and include a form of sensing, inducing some sort of dynamic “social behaviour”, often referred to as a politeness protocol.

When a more complex system of sensing and social behaviour is prescribed for a group of devices, we call such a mitigation technique a “Spectrum Access Mechanism”, not to be confused with a “spectrum access method” which is just describing the behaviour of a single device. The social behaviour may include dynamic changes in nominal frequency, power or timing, or in the amount of frequency space, geometric space or time space used.

E.g., LBT repositions the transmission in time, rather than stops it; AFA re-positions the transmission in frequency rather than stopping it.

When we project this on the definition of Spectrum Utilization Efficiency (SUE) expressed by the complex criterion:

$$SUE=\{M,U\}=\{M,B \cdot S \cdot T\}$$

where:

M: useful effect obtained with the aid of the communication system in question. The definition of this useful effect is up to the user, regulator or manufacturer

U: spectrum utilization factor for that system.

It can be concluded, also from experience, that some spectrum access or mitigation techniques are inherently “inefficient” because they limit the use of the spectrum while unused spectrum is still available and others are not because they allow the use of all available unused spectrum. It needs to be noted that there may be legitimate reasons for doing so, but it makes no difference for the calculation itself. Also it is not the intention to classify certain methods as better than others.

Considering these basic formulas one could get the impression that for a particular system all parameters in the utilisation formula are exchangeable. This is not always the case, the relation between B, S and T is not always linear and even if the parameters are exchangeable there are other boundaries caused by for example, physical receiver parameters.

However an approach like this provides a more flexible environment for SRD deployment than the current approach of giving each application its own reserved frequency space.

Recommendation ITU-R SM.1046-2 [7] indicated that these calculations of U and SUE should only be used to compare similar systems. This makes it difficult to apply this concept directly to the SRD bands, where a variety of different applications share the same spectrum. The move to application neutrality (see section 2.7.1) will make it even more difficult to apply the procedures in Recommendation ITU-R SM.1046-2 [7].

2.5.1 Observations and definitions based on the General Approach given in Recommendation ITU-R SM.1046-2

Spectrum efficiency can be described in different ways but the general consensus is that for a system to be efficient some useful information needs to be transmitted. The nature of this information can be very diverse. A standard time or frequency transmitter only sends its identification at regular intervals and a sound broadcasting transmitter sends it information for 100% in time but both can be considered spectrum efficient. For SRD’s that are usually operating in a group the situation is a little more complex. The following spectrum efficiency case definitions are based on common different identifiable scenarios. The definitions used are newly introduced for the purpose and context of this report and are a way of expressing these scenarios so they can be referred to in other sections of this report.

Single system Absolute efficiency (SAE)

This is the efficiency of a single system in free space under ideal circumstances:

$SAE=SUE$

It is difficult to measure because its efficiency depends on the perception/definition of a person, user or manufacturer. The application requirements dictate the spectrum utilisation in relation to the amount of useful information to be transmitted. For example redundancy or low latency is required for safety critical applications which means the application needs to utilise the spectrum more than needed or it needs to impose restrictions on other users. Both scenarios could be explained as spectrum efficient for that particular application and in the perception of that particular user/application but this is not necessarily the case for other devices/applications.

Single system Relative efficiency (SRE)
This form of efficiency is easy to recognise and even measure:

$SRE=SUE_1/SUE_{ref}$

When for example two transmitters transmit exactly the same information to the same amount of receivers with the same quality of service using different modulation schemes, bandwidth or different power levels, the relative efficiency can be calculated using the spectrum utilisation formula.
This form of efficiency calculation and measurement is easy but not very useful because it assumes an ideal clean and interference free environment.

Single system in a group relative efficiency (SGRE)
This form of efficiency is a logical result of the previous two forms and can be measured by taking into account the variation of certain environmental parameters:

$SGRE=SUE_1(condition\ x)/SUE_{ref}\ (condition\ x)$ *Under various environmental conditions*

Some modulation schemes are robust and keep working while others fail in heavy interference or bad propagation situations. A relatively spectrum efficient system can cope with interference while maintaining the same operational parameters as the relatively spectrum inefficient system that in turn fails under these interference conditions. The whole digital versus analogue debate falls for example under this category of efficiency.

2.5.2 Modified Approach in Recommendation ITU-R SM.1046-2

Group spectrum efficiency or multiple systems in a group (GSE)
This type of efficiency is calculated as a hybrid of the above methods. The contribution of a single device to the whole group of devices of different nature needs to be determined. How do the other devices react and how is the total spectrum utilised when a single new device is added to the group. The absolute efficiency of a single device cannot be calculated or measured in a meaningful way but the efficient use of the whole environment in which the device operates can be analysed to conclude something about the efficiency of a device. The interesting part is that both the susceptibility of a device to interference from the group and the interference contribution to the group is taken into account.

For each individual SRD the quality of information or quality of service is regulatory irrelevant but the quality of service of the typical SRD taking all SRDs in a particular environment into account is an issue.

$GSE=SUE_{total}/SUE_{total\ after\ adding\ new\ device}$

GSE appears an interesting way to define and measure spectrum efficiency because the policy for SRD's is that the functioning of an individual device cannot be guaranteed but it may be possible to do this for the average or typical device in a group[2]. This also leads to an average efficiency for that group. For each device SRE can be calculated but after adding a new device to the group the GSE can also be recalculated for each existing device. The SRD environment becomes dynamic, spectrum efficient technologies may be reassessed and even become inefficient based on technological progress. Grouping or clustering certain technologies or deployment schemes could also lead to overall better GSE.

[2] It is also common practice in ECC and ETSI to investigate the impact of a new spectrum user on the existing users. The definition of GSE formalises this common practice.

The GSE approach would, however, require input from new system metrics as described later in the report.

2.6 THE OSI LAYER MODEL

The OSI model is a theoretical layered model of any information system, which is useful to explain the different functions of such a system. Real life SRD systems mostly do not have all layers implemented or use a combination of these theoretical layers. The model in this case is only used to explain some basic principles.

In the SRD bands the choices of modulation systems, error correction protocols and link establishment choices for robustness and latency and the application are all made by the manufacturer.

- Interference management for SRDs is therefore completely different from other planned/licensed systems [5]: For planned (licensed) radio systems interference management is performed employing detailed studies of interference sensitivity of one specific system in the vicinity of another specific system. The interference sensitivity is related to degradation of the payload of the interfered system. This is at the application layer of the OSI model

- For SRD's the upper 5 or 6 layers of the OSI model can be used freely by the manufacturer of a system. All decisions about this influence the robustness of the application Interference can therefore not be measured at the level of payload. Interference management takes place but only at the medium itself. This is in the Physical Layer, and with a Spectrum Access Mechanism also in the lower part of the Data Link Layer.

Describing a well-defined spectrum access mechanism is from a regulatory point of view the easiest and fairest way to manage interference for SRDs giving manufacturers maximum innovative freedom. Usually a set of spectrum access mechanisms and spectrum access methods that work well together is chosen for a particular frequency band or frequency segment.

The OSI model is a theoretical model leaving the actual boundary between interference management and industry implementation a little flexible. The actual minimal regulatory limits however needed for interference management can be found in the EC SRD decision and the recommendation ERC/REC 70-03 [2].

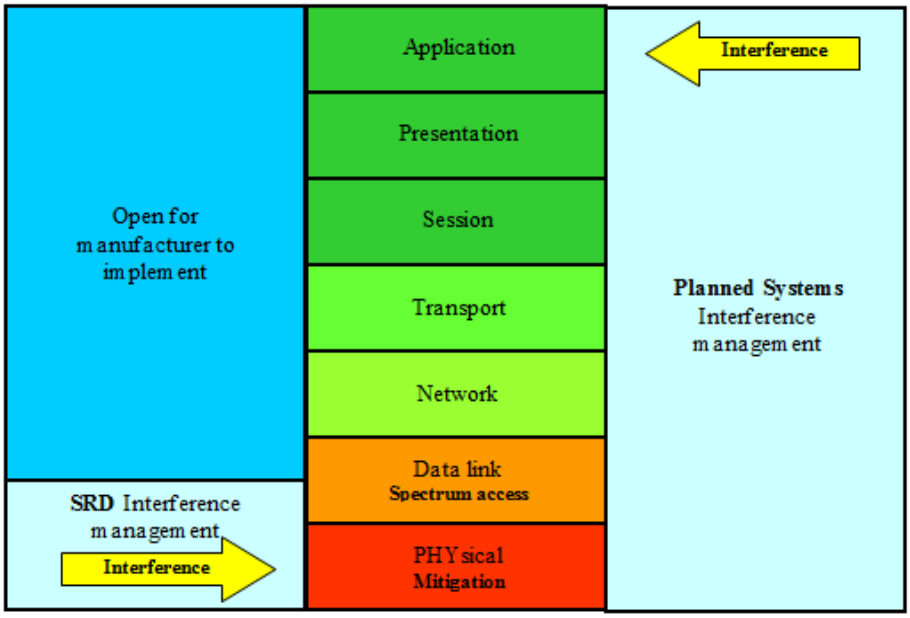


Figure 1: OSI Layers [8]

2.7 NEUTRALITY PRINCIPLES

There is an expressed preference by the EC, RSPG and CEPT's national administrations for spectrum access regulations to be application neutral and technology neutral, with the objective of supporting the continuous process of development and innovation going on in the area of SRDs.

This is only made possible if the technical layout of complete radio systems can be chosen with maximum freedom. The choice of modulation, error correction protocols and link establishment choices for robustness and latency and the type of application served are all left to the choice of the manufacturer.

It is likely that for the same reason of technology neutrality there will be a trend towards grouping users not by application but more by the type of signal transmitted. E.g., access to a frequency sub band will depend on a combination of parameters such as power, duty cycle, length of transmission, spectrum access method. This section is a discussion of some of the issues arising from this preferred neutrality principle.

2.7.1 Application neutrality

One immediate point to make is that the expectation and requirements of different SRD users vary widely.

Consider as example the following applications, each of which generates short data bursts. In each case the application data content is only one or two bits, but the message or packet is built up to some 50 to 100 raw data bits consisting of overhead and security needs. The actual transmissions are very similar, and possibly indistinguishable without a priori knowledge.

1. Remote control, lighting control: the user expects the message to be delivered and acted upon within a very short time, of the order of 100 ms. A noticeable delay or a manual retry is unacceptable to the user.
2. RAKE (Radio Activated Key Entry) car systems. Garage door opener: the user has the same expectation of almost instant response, but is conditioned to make a retry in the event of failure.
3. Building security systems; intruder detection, social alarms: a delay of the order of 5 seconds may be acceptable. While some intruder systems may have 90 second delays for verification, social alarm and fire alarms would expect a response in a few seconds.
4. Heating, ventilation, air conditioning control; building management: the acceptable delay could be of the order of minutes.

Although the data bursts belonging to the applications above may be almost identical in form to an external observer, the applications they belong to have very different criteria for success, and therefore different needs in terms of spectrum access. Or to express it more formally:

The relationship between spectrum access and perceived functionality is different for different applications, even though the signal parameters are identical

The key issue that differentiates the examples chosen is Latency – the time within which the message must be transferred and acted upon as well as the possibility or not of re-transmission in case of communication failure – in other words, the importance of reliability (probability of success) of individual transmissions. Latency is also an issue for moderate sized data bursts. For instance point-of-sale equipment or GPS location data may require latency of no more than a few seconds, but certain telemetry or status reporting could accept much more.

Even with large or continuous data streams, the same variation occurs. Voice, for instance, requires very low latency, but audio streaming can tolerate a few seconds delay and some applications, such as file transfer, can tolerate longer delays. **Application neutrality therefore can only be achieved if the proper technology, in terms of latency, reliability or data bandwidth is described for all application types in the same environment.**

This is obviously not always possible in shared spectrum. A segment of shared spectrum does therefore not always support all applications. This means that in some cases (for instance very demanding applications) full application neutrality may not be an achievable objective.

2.7.2 Technology Neutrality

Technology neutrality has different definitions in different areas of technology and is in electronic communications usually described as “the rules should neither require nor assume a particular technology” [9]. As one can see this reads in two parts *require* as in regulation and *assume* as in (harmonized) standards.

Technology neutrality is a desirable aim, but similarly, is only truly achievable when applications have equal access and equal requirements.

From section 2.5 we can conclude that a maximum group spectrum efficiency (GSE) is achieved when the used technology is of the highest achievable mitigation level for that particular application. Mitigation level in this context means the effectiveness with which the spectrum may be equally divided between a fixed number of users/applications/devices allowing at the same time all users/applications/devices to fulfil their operational requirements. Only the addition of systems with equal mitigation levels relative to the original systems may be added to keep the same GSE level. Systems with better mitigation levels may be added as long as their mitigation levels are equally polite to the existing systems as to systems of their own kind. This may be explained with two different examples, the first example adds a more sophisticated system to a group of relatively spectrum inefficient devices, the other example describes the opposite and adds a less spectrum efficient and less polite device to a group.

1. Adding devices with a basic spectrum access method, such as LBT+DC limiting, to a DC only band may increase the GSE in some cases but adding a more adaptive system with high SRE to the DC only group will destroy the GSE, it is therefore sometimes advisable to allow LBT+DC but prohibit the use of devices that increase their DC value dynamically above a level that makes the DC only devices inoperable.

   This is explained in more detail in the section on mitigation.

2. Another example is the 2.4 GHz band often referred to as the WIFI band because only one dominant access method out of 5 allowed is used for almost all applications. These access methods are matched in terms of mitigation. Deviating from those mechanisms creates an unreliable situation for the whole group of applications. Usually these more complex access methods rely on a strictly defined network structure, a device not belonging to that network structure degrades the functioning of the whole network.

From the above it can be concluded that group spectrum efficiency (GSE) and technology neutrality are in direct conflict with each other if no mandatory technical border conditions for all devices in a certain environment are defined.

These border conditions are the technical boundaries between which a signal parameter value such as power, bandwidth, duty cycle etc. may vary; it is of course not the intention to describe one mandatory technical solution.

2.8 PATTERNS OF INTERFERENCE

After the discussion on application neutrality above, it would seem obvious to state that a given pattern of interference will have different effects on different users. But an equally important point to be made is that the pattern of interference is often not defined. The need to simplify a complex situation to a simple metric inevitably loses important detail.

In WGSE compatibility studies it is common to calculate a statistical probability of interference. I.e., to predict (using SEAMCAT for instance) the probability that at a particular place, time and frequency there is already someone else using the channel. The probability of interference is found by taking many snapshots and seeing in what proportion interference occurs. If the events in the snapshots are truly random in place, time and frequency (as may be the case with mobile systems), then, this is a valid approach and a single number for probability of interference has a meaning.

But in many cases, such as when the considered radio system is (quasi-)stationary, it may be subject to some locally present source of interference and the underlying events and processes are not random. What might a 10% probability of interference mean when the application is not defined and the interference environment is not entirely random?

1. 10% of the people who buy a unit will never get it to work in their houses, or
2. One packet in every ten is lost
3. The system is unavailable for 6 minutes each hour
4. There is a regular 100 ms pulse of interference every second.

These scenarios might seem artificial, but they do demonstrate cases where interference cannot be treated as a simple number.

Scenario 1 is what happens when a permanently working stationary interferer, such as powerful broadcasting station sterilises a fixed geographic area. In fact, this situation is commonly reported by interference investigators.

Scenario 2 is what could happen when a simple frequency hopping system overlaps with a fixed frequency system. Each system will experience a regular pattern of collisions.

Scenario 3 is a hypothetical example of a networked system gathering data once an hour, or of a voice system (these are generally accepted to fit within the existing 10% over one hour duty cycle rule).

Scenario 4 is a known instance of a system that takes a continuous data stream and compresses it into a regular train of bursts in order to comply with a 10% duty cycle limit.

These four examples are different, but it can be seen that a simple analysis would ignore the pattern and measure each one as being a random 10% probability of interference.

The other point to bear in mind is that the effect of these different patterns will depend on the circumstances.

Scenario 1 is surely unacceptable in any case as a policy aim – either by regulators or by manufacturers.

Scenario 2 depends on the application. If there is a manual operator, such as with a car key fob, he will just push the button again. But for an automated, unattended system, such as an alarm, the consequences could be more serious.

Scenario 3 again shows how it depends on the application. For some systems this could be acceptable but not for low latency requirements. In building management, for instance, a heating control system could accept this pattern of interference, but a lighting control or security system could not.

Scenario 4 on the list is an example showing how it depends on the equipment. That pattern of interference would be fatal to an analogue cordless audio system, for instance, but a digital system with error correction could take it in its stride.

These examples highlight two significant risks when working with a simple probability of interference. Firstly, the assumption that events and processes are random may not be always correct. The interference may actually have a pattern, whether in time, frequency or space. Secondly, even when the pattern of interference is known, the effect on the victim is not application neutral. In fact the pattern can even be exploited in some cases to mitigate the effects of the interference.

2.9 LIMITATIONS OF CONVENTIONAL COMPATIBILITY STUDIES

The Minimum Coupling Loss (MCL) method (see ERC Report 101 [10]) calculates the isolation required between interferer and victim to ensure that there is no interference. The method is simple to use and does not require a computer for implementation. The result of an MCL calculation is an isolation figure which, can then subsequently be converted into a physical separation having chosen an appropriate path loss model. The primary drawback of the MCL method is that it is a worst case analysis and produces a spectrally inefficient result for scenarios of a non-determined nature.

A Monte Carlo (MC) simulation (see ERC Report 101 [10]) is a statistical technique based upon the consideration of many independent instants in time and locations in space. For each instant, or simulation trial/snapshot, a scenario is built up using a number of different random variables i.e. where the interferers are with respect to the victim, how strong the victim's wanted signal strength is, which channels the victim and interferer are using etc. If a sufficient number of simulation trials are considered then the probability of a certain event occurring can be evaluated with a high level of accuracy.

The MCL method calculates whether interference could or could not occur in a one on one situation; the MC method tries to estimate the probability or rate of occurrence in a real world situation. Each method goes further than just looking at collisions in time domain; the test is whether the collision is harmful from the victim’s point of view.

Each method, however, has the drawback that it only considers one snapshot at a time, and then only considers whether the PHY layer is disrupted in that snapshot. No account is taken of the time domain system dynamics, such as possibility of repeated re-transmission in case of collision, incurred latency, etc. The importance of this consideration is discussed in 2.8 above.

Thus when more complex mitigation and spectrum access techniques are used (for instance those that rely more heavily on time domain dynamics) problems with conventional studies may arise. The problem is often not the methodology itself but the parameters used. The definition of a parameter such as DC is not just a static value to be used in the simulation but a complex timing sequence with an interference potential based on the scheme and the demodulator of the victim. Unfortunately, as of today there are no known methods that would allow reliably evaluating system dynamics and resulting interference potential in a generally defined inter-system scenarios on required macro scale of complex operational environments of spectrum-space. Therefore evaluation is usually done through variously approximated simulations by general tools such as MC or MCL, while verification of system dynamics aspects (in cases of doubt) may be ensured through implementing complementary real-life tests.

Some general requirements to any successful interference modelling include the following:

- A simulation should be to the maximum extent possible based on the real properties of the mitigation and spectrum access techniques used and, whenever possible and practical backed up with appropriate measurements.
- The proper translation should be ensured between simulation parameters, the definition in the regulation and the definition in the harmonised standard

- Simulations concerning critical parameters should be well documented and have references in the documentation to the compatibility conditions determined during studies or otherwise.

The difficulty then is of course that it is that much harder to create a standard simulation. One solution may be to use a MC simulation for the statistical processes and a separate device simulator for analysis of system dynamics and combine the results.

Later in this report, several methods of analysis in the time domain by means of spreadsheets are presented. It is expected that these new tools can help to close this gap.

2.10 MITIGATION FACTORS

For mitigation a number of definitions exist, most of them are related to a particular technology. A general definition of mitigation could be as follows.

> Mitigation is the ability of a radio transmitter or transceiver system to coexist and share frequency space in time, bandwidth and geographical space with other radio systems causing no or a defined quantifiable amount of interference to each other.

The level of mitigation depends on the technology and radio interface used and is often a combination of technical requirements and operational conditions. In the most ideal case the number of devices and types of devices present has no influence on the level of mitigation. A consequence of a high level of mitigation is that it can lead to low data rates when many devices are using the frequency space at the same time.

If, on the other, hand a mitigation technique is used offering less than perfect mitigation a progressively increased probability of interference for an increased unit density occurs. Spectrum efficiency is highly reduced for high unit densities, which is undesirable in cases of frequency scarceness. Duty cycle for example is not enough to ensure an efficient use of the spectrum in most cases. The following are examples of techniques offering mitigation. Keep in mind that many techniques can be and are used in combination:

Table 2: Examples of techniques offering mitigation

Mitigation		Description
Mitigation in time		Duty Cycle (DC) Duty cycle is a spectrum access technique but, where duty cycle limits are set below a value required for a victim system's operation, mitigation occurs
		Low Duty Cycle (LDC) LDC is a variety of DC with a low DC value compared to the DC of potential victims and specific timing considerations such as a defined TX_{on} and TX_{off} time
		Listen Before Talk (LBT)
Mitigation in the frequency domain	Multi frequency	Frequency Hopping spread Spectrum (FHSS)
	Frequency spreading (reduces power spectral density and thus power into a narrow band Rx)	Direct Sequence Spread Spectrum (DSSS) often combined with CDMA where spreading and multiple access are two complementary functions.
		Ultra Wide, with spreading as the primary function according the UWB definition.
		Time hopping a method superseded by UWB, here mentioned for completeness
	Frequency selection or avoidance, also called adaptive frequency agility (AFA)	Detect and avoid (DAA), avoid an occupied channel permanently of based on specific compatibility rules. or change to another frequency permanently
		Dynamic frequency selection (DFS), avoid an occupied channel temporarily or change to another frequency temporarily
Mitigation in time and frequency domains together		Listen before talk (LBT) with detect and avoid (DAA), or with dynamic frequency selection (DFS)

Mitigation	Description
Mitigation based on geographical space and radiated power (footprint reduction)	Antenna pattern (Effect of beam width, main-beam and side-lobes)
	Total radiated power (TRP), for groups of devices where antenna patters of individual devices are averaged out to a fixed level in the spatial domain. The whole concept of TRP is that it only works for a group
	Ultra low power (ULP) communication, different from UWB this means both narrowband or wideband devices operating under a low power value harmless to all other devices in the same frequency space. Ultra low power is the only mitigation technique and devices typically may have a range of cm's. Applicable levels are currently under study in the CG for the revision of the EC SRD decision
	Adaptive Power Control

It is clear that mitigation cannot simply be derived from spectrum occupancy. It is convenient to discuss mitigation techniques in 4 types because mitigation levels cannot be simply expressed in a number based on technical parameters of a device. Mitigation levels, just as efficiency, are based on the behaviour of a device in relation to other devices. An attempt to describe these levels is given below. The types are based on the assumption that a particular technology is able to protect another technology to some extent. Of course there are grey areas because types may be combined and secondary effects such as the influence of the environment are not taken into account. The examples need to be seen in this light.

TYPE 1 self-limiting, non sensing

a) This type is a simple combination of, for example, duty cycle or FHSS which may be considered as a DC on a number of parallel channels and environmental parameters. Also physical parameters like the antenna pattern, in combination with environmental parameters, falls within this definition. There is a mutual protection between devices of the same kind based on the acceptance of a number of collisions.

b) Within this type a higher level of mitigation may be obtained with a spectrum access technique such as DSSS-CDMA or TDMA. These techniques offer mutual protection between devices of the same kind in the same communication chain without the need for carrier sensing. It does not offer the same level of protection to other types or devices outside the communication chain since there is no central control. An example is TDMA that looks like DC for those systems that are not part of the TDMA communication chain. Or DSSS-CDMA that looks like an increase in the noise floor to non CDMA systems.

TYPE 2 self-limiting, sensing

This type is based on single sensing LBT or repeated sensing LBT with duty cycle in combination with more advanced techniques such as DFS, AFA, DAA or any other agreed interference limiting behaviour with comparable performance to the ones mentioned. These techniques can offer mutual protection between devices of the same kind in the same communication chain and it can also protect devices of type 1 and even devices falling outside the four described types such as non limiting devices.

This type can be divided into:

a) destructive (non-polite) sensing systems that use for example use an ACK to retransmit without a new sensing action if interfered.
and
b) non destructive (polite) sensing systems that listen before performing each transmission.

TYPE 3 Group optimised

Mitigation is based on one optimized Spectrum Access Mechanism for the whole group, yet everything beyond that spectrum access mechanism remains free for the manufacturer to choose. In case of congestion, the Spectrum Access Method ensures equal access to the spectrum (and hence a gradual degradation of service to all users).
A group means here all SRD devices that are in each other's working environment. It needs to be noted that group optimisation is difficult to obtain in shared spectrum

TYPE 4 centrally organised
It is a system based on Type 1b or type 2 in combination with a central control system to expand the benefits of the simple type 1b and 2 mechanisms to a collective/group of devices. This type offers mutual protection between devices within the group and it can protect devices of type 1 and 2 outside the collective. The level of mitigation towards systems that are not in the collective is based on the type 1b or 2 mechanism chosen. A centrally organised system needs to have parameters to change in order to apply the organization. A type 1 mechanism such as simple DC for example causes collisions that create a situation where mitigation is not under full central control. DC under central control is not excluded. However, it is defined as a type 2 under central control.

The choice of one of these types is sometimes based on specific use like the type 3 used in the 2.4 GHz band for wideband data system, or on cost like the type 1 used for alarms and social alarms. Spectrum efficiency and, more importantly, technology neutrality and flexibility, outside the scope of a particular application or range of applications, has never been the primary goal for choosing mitigation a strategy based on one of the types.

The following table gives a few examples of the different types of mitigation techniques.

Table 3: Examples of techniques offering mitigation

Type	Mitigation		
Type 1	Duty Cycle based FHSS systems	RAKE systems, alarms, meter reading	Costs, battery lifetime, size and simplicity
Type 2	DC + LBT or LBT+AFA	(some) Home automation systems	Guaranteed throughput and / or better reliability in an environment shared with DC devices
Type 3	Duty cycle sequenced		
Type 4	TDMA with controlling base station	DECT	Guaranteed interoperability and safety

Type 3 and type 4 offer the best possible level of mitigation but are not effective when a situation is sought to accommodate as much as possible diverse types of SRDs in the same spectrum. Spectrum efficiency and diversity are therefore in conflict.

The harm caused by a particular device is based on the level of interference acceptance within the group. If a group for example consists of DC devices, collisions are part of the normal operating conditions of the devices. For an environment with only LBT devices, collisions are much less common.

Mitigation is not the same as GSE, for GSE the total data throughput of a group in the PHY layer is the criterion. For mitigation the mutual difference in data throughput in the PHY layer is the criterion.

2.11 NEW METRICS

As discussed in section 2.5 spectrum efficiencies can be defined in terms of the GSE in an environment where devices of different and similar nature are present.

As a principle we can use an equal division of frequency space, in terms of medium utilisation. Further we can use an equal division of possible/available data throughput in the group as a measure of spectrum efficient behaviour of that group. In practice it may be possible to realise this by choosing the technical parameters from a pool of possible combinations of power, bandwidth, geographical distribution, mitigation techniques and spectrum access methods.

Each parameter needs to be controlled and limited in such a way that it is not the dominant factor in the spectrum utilisation or data throughput calculation. If we do not do that, we are promoting a type 4 or type 3 system, basically the best way from an efficiency point of view, but not acceptable when we want to allocate shared spectrum. We may also create for example a balance of power problem.

An example is the time period over which DC is defined or the LBT threshold that is based on signal level instead of predefined signal or data properties.

In section 2.7.2 we conclude that GSE and technology neutrality are in direct conflict with each other if no mandatory technical border conditions for all devices in a certain environment are defined.

If we really want to achieve maximum spectrum efficiency, then each parameter needs to be defined based on the minimum application requirements of all devices in the group by defining border condition for each parameter and as such sacrificing some technology neutrality.

In short, the realisation of this minimum application requirement is much more relevant than the realisation of a certain probability of interference figure. A zero interference figure may not be obtained so the situation of ideal spectrum efficiency is always accompanied by a reference maximum interference figure.

The following section describes delay as one of these requirements. Each relevant parameter should be analysed in a similar way.

2.11.1 Probability Distribution of Delay

This section discusses delay, the time spent waiting on a shared channel until a message can be sent. Although this delay cannot generally be expressed as a single number, it can be analysed in probability terms. In the diagram below graphs A and B show the situation in a clear channel, where no delay is expected. Graph A is the probability distribution function (pdf) showing the likelihood that a message will be delivered at a given time. Graph B is the same information as a cumulative probability. This shows the probability that a message will be delivered by (i.e. at or earlier than) a given time. The cumulative probability is found by integrating the probability distribution function. The left hand edge of the plot in A is actually a delta function but is shown expanded for clarity.

Graphs C and D show the effects expected in the presence of other users. Diagram D, the “Cumulative Probability of Delay” is particularly useful.

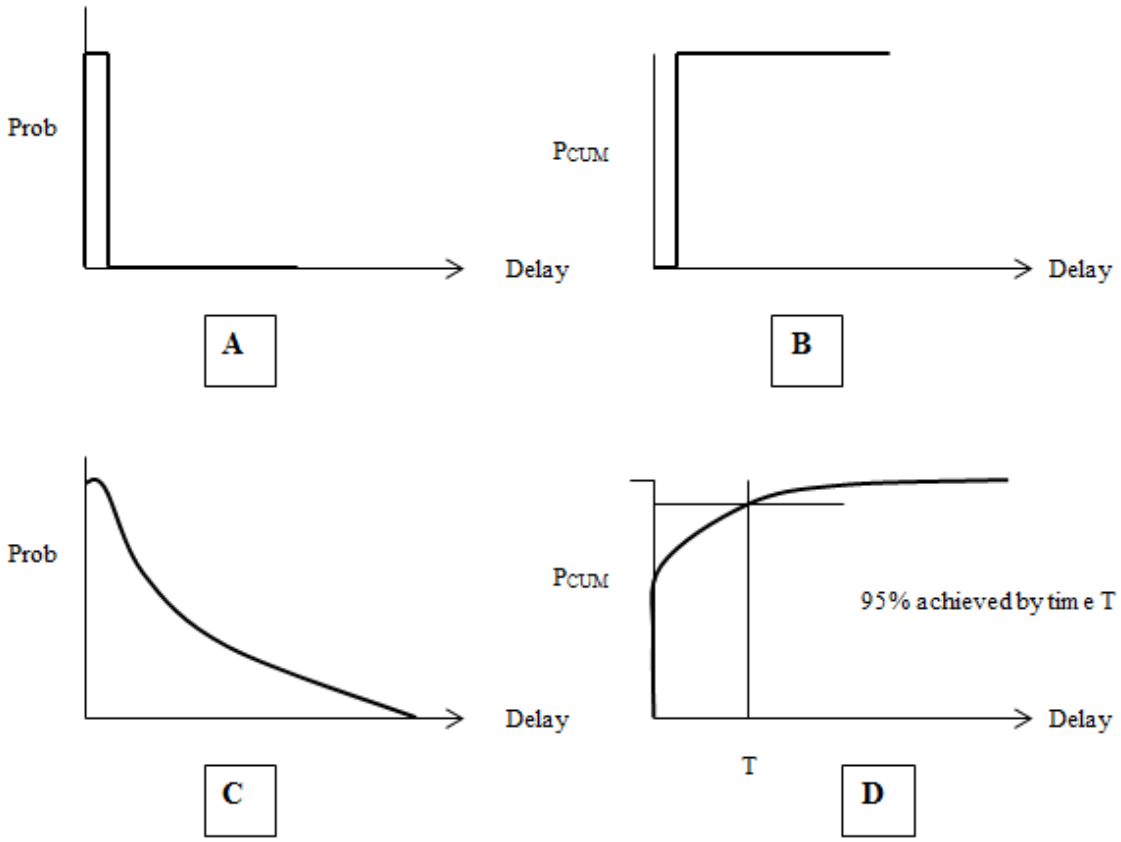


Figure 2: Probabilities of delay without (A&B) and with (C&D) shared channel use

When it comes to measures such as latency and reliability, the expectation of a user is often expressed as "X% of messages must be delivered within a time d" and this is easily read from the diagram, whether X is, for instance, 90, 95% or 99% as required by the application.

2.11.2 Calculating Probability of Delay

In some cases, constructing a diagram such as the Cumulative Probability of Delay one may require complex analysis. It might be possible to model this in a centrally managed telecommunications system as TDMA (GSM) or Ethernet line, etc. It should be noted that there is a considerable body of work in the fields of telecommunications and networks that can be drawn on, although care must be taken in applying it to wireless systems. But it is unlikely that this probability could be modeled as a general objective for deployment of dispersed non-homogeneous systems like SRDs in shared bands.

In some cases, however, it is relatively simple to generate a Cumulative Probability of Delay diagram.

Consider the case of a user wishing to send a short message when there is already one other existing user. The existing user sends transmissions of duration T, at random times with an average frequency of F. In other words, the duty cycle τ is

$$\tau = TF$$

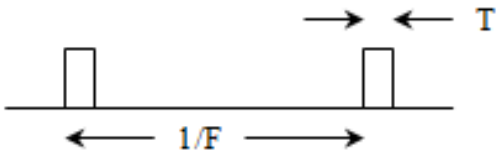


Figure 3: Random transmission of competing signals

The important parameter is the wait time, or delay, until the channel is free. Both the pdf and the cumulative probability of this can be found by inspection.

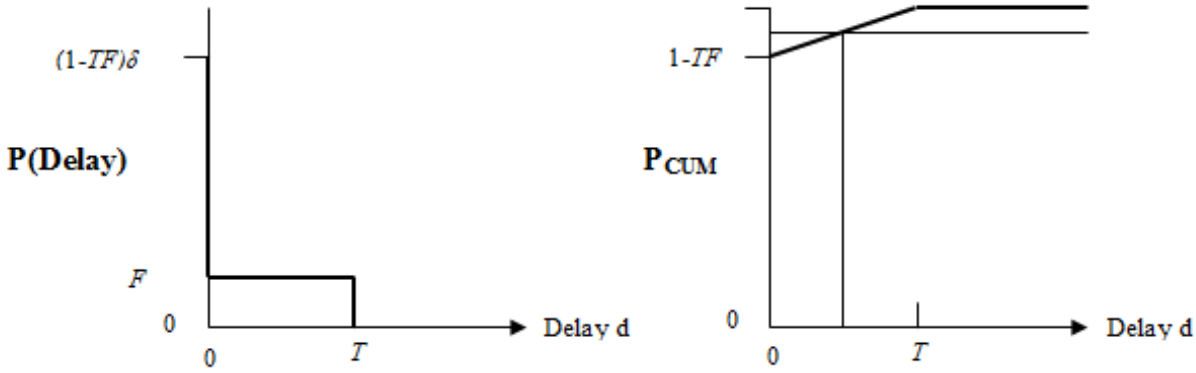


Figure 4: Probability of delay in case of competing signals/users

Suppose for instance that the transmissions are 1 sec every 10 secs, so that

$$T = 1\,\text{s} \quad , \qquad F = 0.1\text{Hz} \qquad \text{and duty cycle } \tau = TF = 0.1$$

The success times for various probabilities are then easily found:

- 90% achieved by d=0 sec
- 95% achieved by d=0.5 sec
- 99% achieved by d=0.9 sec
- 100% achieved by d=1 sec

Consider next the case where the competing user is still at 10% duty cycle, but with transmissions of 10 sec duration every 100 sec.

$$T = 10\,\text{s}, \quad F = 0.01\text{Hz} \qquad \text{and duty cycle } \tau = TF = 0.1$$

The success times are then:

- 90% achieved by d=0 sec
- 95% achieved by d=5 sec
- 99% achieved by d=9 sec
- 100% achieved by d=10 sec

It can be seen that the delay times for a given probability of success are increased by a factor of 10.

This is an important result. In both cases the competing transmission is the same duty cycle; a simple analysis based on probability of interference will come up with the same result. But the cumulative

probability of delay shows that from the point of view of a victim, the harm done by one is 10 times greater than the other.

In the case of N such identical interferers, the cumulative probability curve will be of the form below.

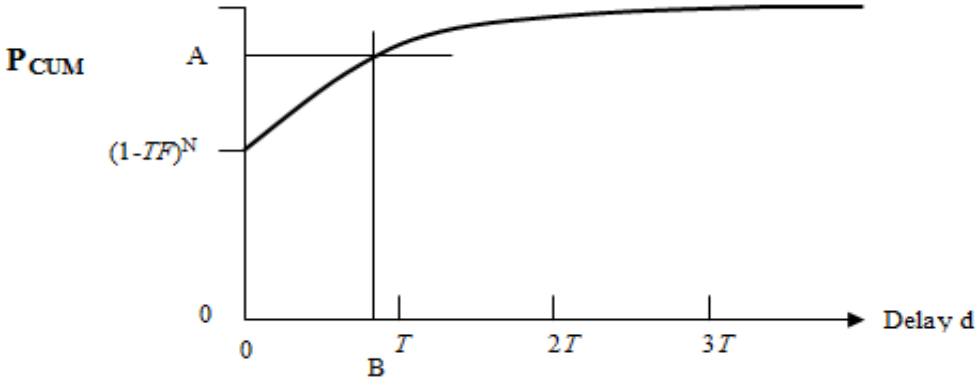


Figure 5: Cumulative probability of delay in channel with N competing users

The horizontal axis is entirely in terms of T, the duration of the transmissions, rather than the duty cycle TF. Therefore it can be seen that, in any given situation, the delay B at which a given success rate A is achieved is directly proportional to the duration of the interfering transmissions, as their duty cycle is held constant.

2.11.3 Expected Delay

The analysis above shows the general pattern of the delay probability, but does not give us a quantitative result, except for the probability of zero delay.

Queuing theory, however, can be used to make a simple model. Suppose that a number of users are sending packets on a channel, where T is the duration of the packets and F is the overall frequency (rate of sending summed across all the users).

We can equate F to the rate of arrival of objects in a queue and 1/T to the rate of clearing. The Expected Delay D until a clear slot is then equal to the expected waiting time in the queue.

$$D = \frac{T.F}{1/T - F}$$

This is not a perfect model of queuing with, for instance Aloha or LBT, which are discussed in section 3. Nevertheless it is a useful indicative result for the expected wait when using an access mechanism in a shared channel. This is the formula used for predicting wait times with LBT in the simulator described in Annex 4.

The equation above can be re-arranged to show the effect of holding constant *TF*, the aggregate duty cycle

$$D = \frac{T.(T.F)}{1 - T.F}$$

In this model, therefore, the expected delay is directly proportional to the transmission duration, which is the same result as derived from the model in section 2.11.2.

2.11.4 Metrics for Latency and Reliability

The traditional metric used in compatibility studies is probability of interference. In many cases discussed above this is not adequate, as it will not completely reflect the harm done to various types of victims by different types of interference. In particular it does not take direct account of the need for metrics such as

low latency or high link reliability (probability of successful transmission, i.e. including re-transmissions) by many users of the spectrum.

It should be considered that latency and reliability are related. The requirement of the user can be expressed as X% probability of success within maximum delay D. For instance, a user stressing low latency might require 90% within 200 ms; one stressing high reliability might require 99.9% within 3 seconds. In these circumstances, Cumulative Probability of Delay and Expected Delay are useful concepts, although they may be difficult to accurately quantify in actual circumstances.

The above is a simple analysis of a complex mechanism. It assumes that the transmission that the “victim” is waiting to make is short compared to the “interferer’s” transmissions. It also assumes that the victim has a way of knowing when it is possible to make the transmission. This is the case if it employs LBT, but similar results will be obtained e.g. if it makes trial transmissions and listens for an acknowledgement. The difference lies in the potential for interference back to the existing user.

The similarity with telecommunications traffic theory and the Erlang distribution and (expressed by Erlang equation) should be also noted, though care should be taken as it is not directly applicable. There is more than one variation of the Erlang equation, and there are a number of differences to be considered. Chief among these is the difference between wireless and wired systems that not all nodes can necessarily hear each other.

Thus it may be concluded that latency and reliability are useful new metrics that may add value to traditional interference analysis, whenever the considered wireless systems and interference scenarios allow some meaningful deterministic analysis of these phenomena.

2.12 SUMMARY

When moving towards defining an authorisation framework for systems described based on technical parameters rather than application, it must be recognised that it is not only the technical parameters of the radio signal and the resulting link budget that are important. The modern adaptable packet-switched systems have complex operational patterns through involving not only the physical layer but also higher OSI levels into the picture for overall maintaining of communications stream. Therefore ideally the system designers as well as spectrum managers should endeavour to consider those more sophisticated aspects in order to determine and establish the balance between the levels of operational resilience of considered systems.

One of the most important operational parameters of this category is the latency requirement. This is the maximum acceptable delay in transferring the packet/message and cannot generally be inferred alone from the technical consideration of the useful link budget vis-à-vis the interference instance. Therefore the latency as well as other similar parameters/metrics may need to be considered when pursuing application neutral spectrum planning.

Another conclusion is that when different applications are mixed, an analysis based on a simple probability of interference does not reveal the full story. Therefore, compatibility analysis in an application neutral environment will require more extensive analysis in the lowest two layers of the OSI model, mainly in the time domain, than is currently done in situations where the applications are defined.

3 BASIC SPECTRUM SHARING TECHNIQUES

3.1 DUTY CYCLE

Spectrum sharing by means of duty cycle limits is a simple and well established technique whereby every user has his transmission time restricted.

Consider the case of N users on a channel, each sending a series of transmissions.

Following the procedure used in ECC Report 37 [1], consider the situation from the point of view of user number N, who arrives at a channel that is already being used by N-1 users.

Suppose user 1 sends transmissions of duration T_1, at a rate of F_1. User 2 sends transmissions of T_2 at F_2, etc. Therefore their duty cycles are

$$\tau_1 = T_1 F_1 \qquad \tau_2 = T_2 F_2 \qquad \text{etc}$$

The relative timings between users is random. User N sends a transmission of duration T_N. The probability that this collides with a transmission from user m is

$$P_{COLLm} = (T_m + T_N) F_m \text{ for } (T_m + T_N) F_m \le 1 \text{ otherwise } P_{COLLm} = 1$$

For the case of $P_{COLLm}<1$, the probability that it does not collide with any transmissions can be written

$$P_{MISS} = \{1-(T_1+T_N)F_1\}\{1-(T_2+T_N)F_2\}.....\{1-(T_{N-1}+T_N)F_{N-1}\} \quad \text{or,}$$

$$P_{MISS} = \prod_{m=1}^{m=N-1} \{1-(T_m+T_N)F_m\}$$

and the probability of that individual transmission suffering a collision is

$$P_{COLL} = 1 - \prod_{m=1}^{m=N-1} \{1-(T_m+T_N)F_m\}$$

This is the general case. If it is then assumed that all the transmissions are similar, ie

$$T_m = T_N = T \quad \text{and} \quad F_m = F \text{, then}$$

$$P_{COLL} = 1-(1-2TF)^{N-1}$$

This function is readily plotted and is shown below for a range of duty cycles.

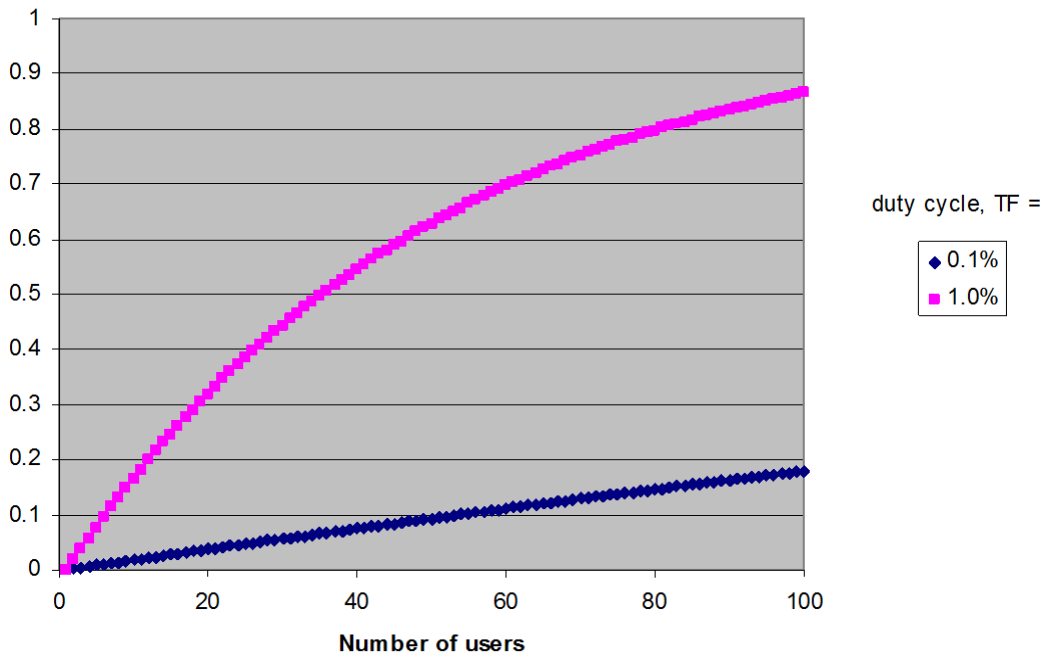


Figure 6: Probability of collision for individual transmission

The figure above shows the collision probability with users operating at 0.1% and 1% duty cycle. Note that for the 1% curve (purple) the X-axis equates to the normalised traffic loading as a percentage. I.e. 100% represents the theoretical maximum traffic capacity.

The figure below shows the collision probabilities with lower numbers of users, and also the effect of a 10% duty cycle.

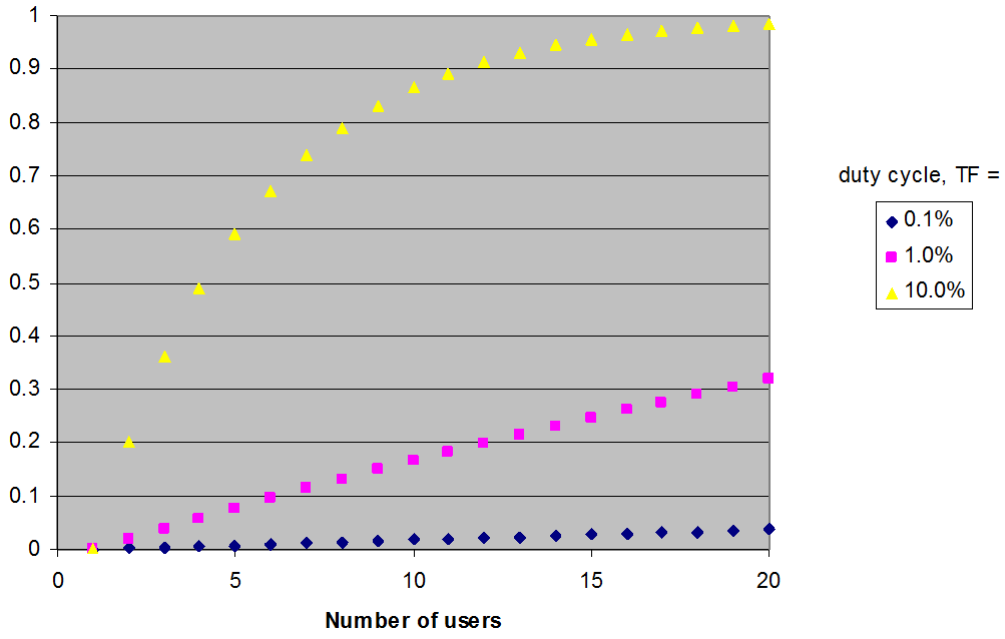


Figure 7: Probability of collision for individual transmission

The probability of collision that is acceptable is a matter for debate, and will in any case vary with application. A system that detects collisions and resends the transmission could in theory cope with high probabilities. For such a system, keeping the collision probability below 20% might be a reasonable target.

For systems that do not detect collisions, e.g., simplex links, lower collision probabilities are required. There is a standard argument that if the collision probability is 10%, then sending the message 3 times results in a success rate of 999/1000. This should be treated with caution as the theory requires that the process is perfectly random, when it may not necessarily be so in the real world. Systems that require 999/1000 ("three nines") performance from their radio links need a more elaborate analysis than can be given here.

Note: Repeating transmissions creates two effects. It means the distribution in time is no longer random, so the simple analysis is no longer necessarily valid. It also increases the number of transmissions, which increases the probability of collisions.

Nevertheless, without collision detection, it is necessary to keep the probability of collision low, and a figure of 5% is suggested as the level above which problems occur.

It can be seen that with 10% duty cycle, there is trouble immediately. As soon as there are two users the collision probability is 20%.

With 1% duty cycle only 3 users can be accommodated before the 5% probability is breached (4 users gives Pcoll = 5.9%).

With 0.1% duty cycle the situation is better; 26 users can be present before 5% collision probability is reached.

This is of course an idealised case, and the analysis considers only collisions in the time domain – it assumes that any such collision results in the loss of the message. But it does lead to some conclusions about the effect of duty cycle as a channel access technique and the strategies for using it.

3.1.1 Strategies for users in duty cycle limited channels

At 10%, a duty cycle limit alone is not effective as an access technique and also offers a useful level of mitigation in only a few specific cases. As soon as 2 users are present, the collision probability if they both operate randomly is uncomfortably high. Even if an array of techniques such as LBT and Aloha are used, throughput on the channel may still be affected negatively.

The choices available in the access layers are to accept delays when other users are present, or to use frequency agility to access other channels.

Another option is to solve the problem of data loss in the higher layers of the OSI model, which is discussed below in 3.6.

A 1% duty cycle limit is less effective than might appear. The likely probability of collisions in many circumstances with typical SRD user densities would be 10 to 20% (not the 1% that might be imagined by an intuitive analysis). In a 1% duty cycle limited channel, blind transmissions may suffer significant collisions with other users. It would be wise to only use this duty cycle limit in conjunction with collision detection.

But at 0.1% duty cycle limit, the situation is better. A significant number of users may be present before the 5% collision probability is reached. In a 0.1% duty cycle limited channel, good results may be obtained just with blind transmissions.

In many cases however, there may be only one user of a channel at a time, so sharing by duty cycle is not relevant. Indeed in these cases, a duty cycle limit may not be appropriate as the sole means of access control, since all it achieves is to limit the use that a legitimate user may make of the channel.

Summary of sharing with Duty Cycle limits:

- Very Low Duty Cycle works well for many systems
- Low duty cycle limits work well for some systems, but the effectiveness will be largely dependent on the density of spectrum use.
- Higher values of Duty Cycle are unlikely to serve as an effective mechanism for good spectrum efficiency other than in some specific systems with a very low density of users.
- Duty Cycle specified over short cycle times will change the impact on others, with respect to the one hour cycle time, but the change in impact will vary between victim services.
- Duty Cycle combined with other techniques may improve spectrum efficiency beyond that achieved by Duty Cycle alone. E.g., carriers sensing and/or avoidance.
- Duty Cycle is the only option for unidirectional systems. This is a simple form of spectrum sharing with minimum hardware requirements and the benefits of this should not be ignored when assessing overall utility.

3.1.2 Implications for Regulators and Manufacturers

It is shown above that if it is desired to keep the collision probability below 5%, then, in the scenario where all devices are in reception range of each other, this is only possible when the aggregated channel occupancy is below a limit of 2.5% to 3% (e.g., 3 users at 1% each or 26 users at 0.1% each).

Where Duty Cycle is the only access mechanism for interference mitigation, it will only be effective in sub-bands of low occupancy. The distinction between occupancy and congestion is important. In this situation, congestion starts to occur at occupancy of 2.5 to 3%. It follows therefore that:

- If provision is to be made for devices using only Duty Cycle as an access mechanism, then regulators must accept that some of the respective sub bands should be arranged to be low occupancy. For these sub bands, the spectrum efficiency should not be assessed in terms of occupancy or data throughput, but in terms of what utility is provided to how many users. An obvious example of the utility outweighing the throughput is the case of a large number of alarm systems, all tuned to the same channel, but all quiet most of the time.
- Conversely, manufacturers must accept that, in the interests of efficiency, not all spectrum can be arranged this way. Duty Cycle as a sole access mechanism can only be relied upon in certain low occupancy sub bands. In other sub bands, occupancy levels above 3% may well be encountered and in these additional techniques will be required.

However, it should be finally noted that the above considerations are valid for scenarios where shared band users/systems have direct interaction (“hear” each other). In some other cases, e.g. with sufficient geographic spacing or other kind of effective shielding/decoupling between sharing peer systems, the timing considerations may be irrelevant.

3.2 ALOHA

The name Aloha comes from a wireless network run by the University of Hawaii in the 1970s. It is significant as one of the first such systems and the trigger for much of the theoretical analysis of packet data networks, both wired and wireless.

In the same way that Duty Cycle could be described as a mitigation technique rather than a spectrum access mechanism, Aloha is not strictly an access, mechanism. It does not attempt to manage or avoid collisions, but rather it is a technique for detecting them after the event and dealing with their effects.

From section 3.1 above it can be seen that the probability of one device out of N experiencing a clear slot for transmission is

$$P_{CLEAR} = (1 - 2TF)^{N-1}$$

Suppose that instead of N discrete devices sending messages at rate F, the same traffic originates from an arbitrary number of devices n, sending at rate f, such that

$$G = Tfn \quad \text{and} \quad P_{CLEAR} = \left(1 - \frac{2G}{n}\right)^{n-1}$$

G is the traffic on the channel[3], normalised so that G=1 represents the maximum theoretical throughput if all the messages were somehow ordered in perfect sequence rather than sent at random times.

Since n is arbitrary, we can consider the case as it tends to infinity, and noting that

$$\lim_{n \to \infty} \left(1 - \frac{\lambda}{n}\right)^{n} = e^{-\lambda} \quad \text{then,}$$

$$P_{CLEAR} = \left(1 - \frac{2G}{n}\right)^{n-1} = \left(1 - \frac{2G}{n}\right)^{n} \left(1 - \frac{2G}{n}\right)^{-1} = e^{-2G}$$

Note the same result can also be derived from consideration of the Poisson distribution, which states that if the average rate of occurrences of a random event in a given interval is λ then the probability of *k* occurrences is:

$$f(k;\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

by setting *k*=0 and λ=2G

Aloha is a system in which unsuccessful messages are retransmitted. In the simplest variation, Pure Aloha, messages are transmitted blind at random time, messages that suffer collisions are retransmitted, also at random times.

If S is the throughput, or the rate of successful messages as a fraction of the theoretical capacity, then S is the rate of attempts multiplied by the probability of an attempt being successful.

$$S = G.e^{-2G}$$

The relationship between S and G is illustrated in the diagram below:

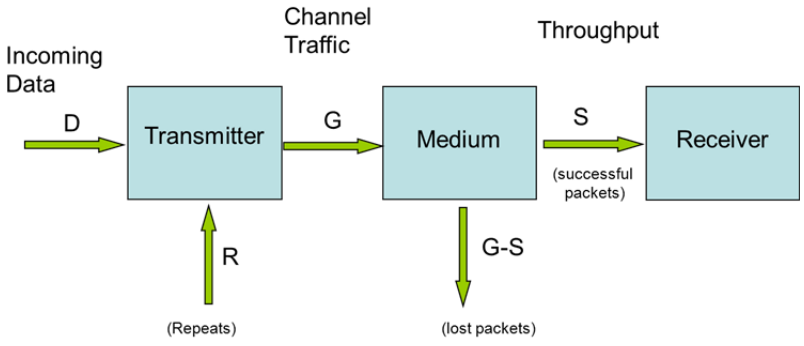


Diagram showing relationship of G and S
Transmitted traffic G=D+R.
In steady state, S=D and R=G-S

Figure 8: Data transmission chain: throughput (S) and channel traffic (G)

3 The term Channel Traffic is, in other literature, sometimes referred to as 'Offered Load'. In the context of this report this term 'Offered Load', is avoided because of possible confusion with the term 'Offered data'.

S is plotted below as a function of G.

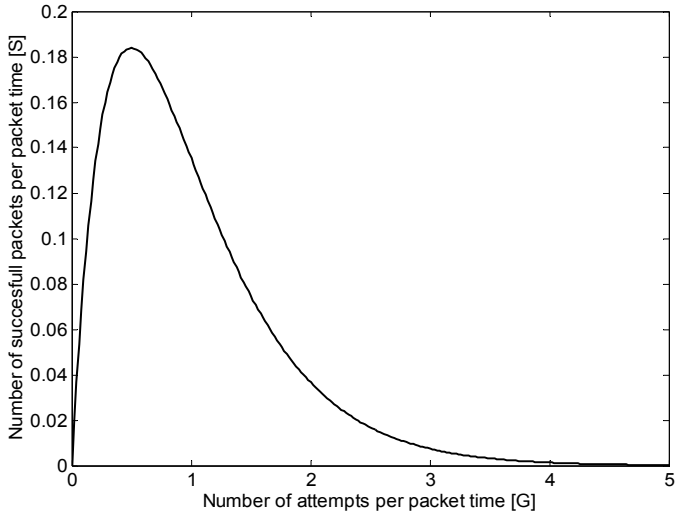


Figure 9: Plot of throughput (S) versus channel traffic (G)

and again, showing the region up to 100% traffic loading in more details.

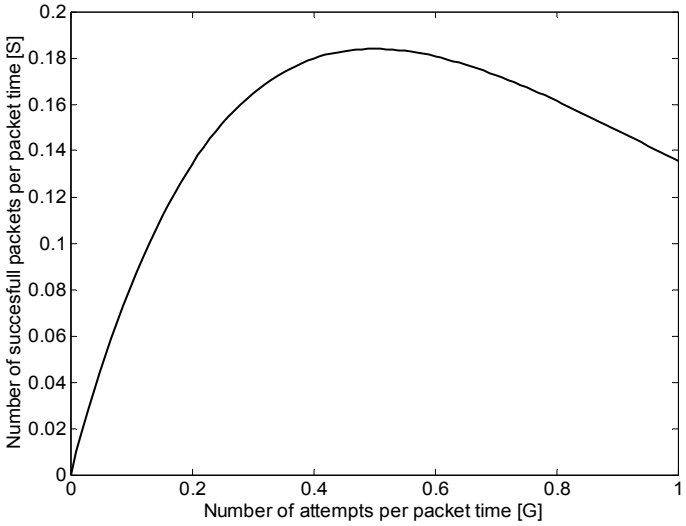


Figure 10: Close-up on the function of throughput (S) versus channel traffic (G)

This function reaches a maximum at G=0.5, when S=0.184. This is the often quoted result of maximum throughput for Aloha of 18.4% of channel capacity. Note, however, that S is the rate of successful messages, whereas G is the rate of attempted messages, therefore the rate of unsuccessful messages is G-S.

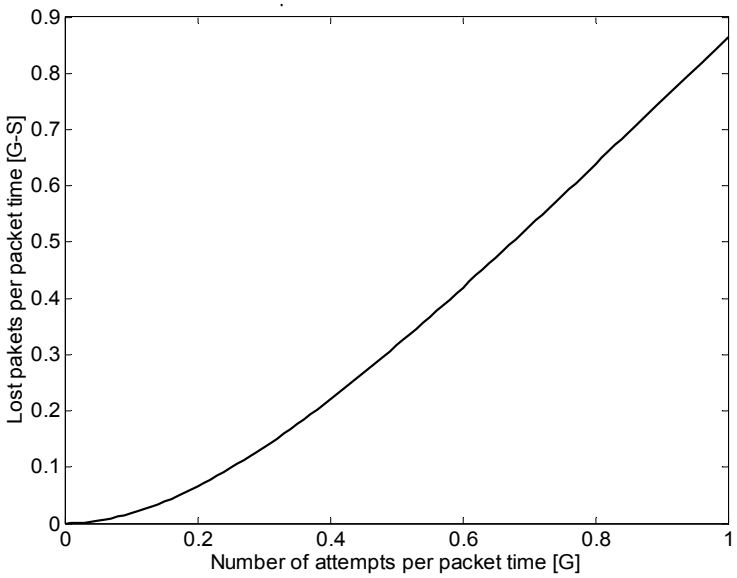


Figure 11: Lost packets (G-S) versus channel traffic (G)

What can be seen from the G-S curve is that an individual packet only has a reasonable chance of success at low channel traffics. At high channel traffics nearly all the packets are lost and the net throughput suffers accordingly.

It is important to understand that the maximum throughput is only realised if all the users have some means of detecting that a transmission was not successful and repeating when necessary and only when necessary. Maximum throughput of 18.4% is accompanied by a rate of 31.6% of unsuccessful messages. Ie., for every message successfully sent, nearly two are lost to collisions. The success rate per message is 36.8%. Many users would consider this unacceptable and would see the channel as congested at much lower values of S and G.

3.2.1 Comparing Aloha and Duty Cycle Limiting

The figure below plots the probabilities of success or failure of an individual message against the channel traffic. These curves are derived using the classic Aloha analysis. The collision probability curve can be compared with the 1% duty cycle curve in Figures 6 and 7. The X-axis is the same in each case. In the diagram below G is the channel traffics, and can be equated with the number of users multiplied by the duty cycle of each user in Figure 12.

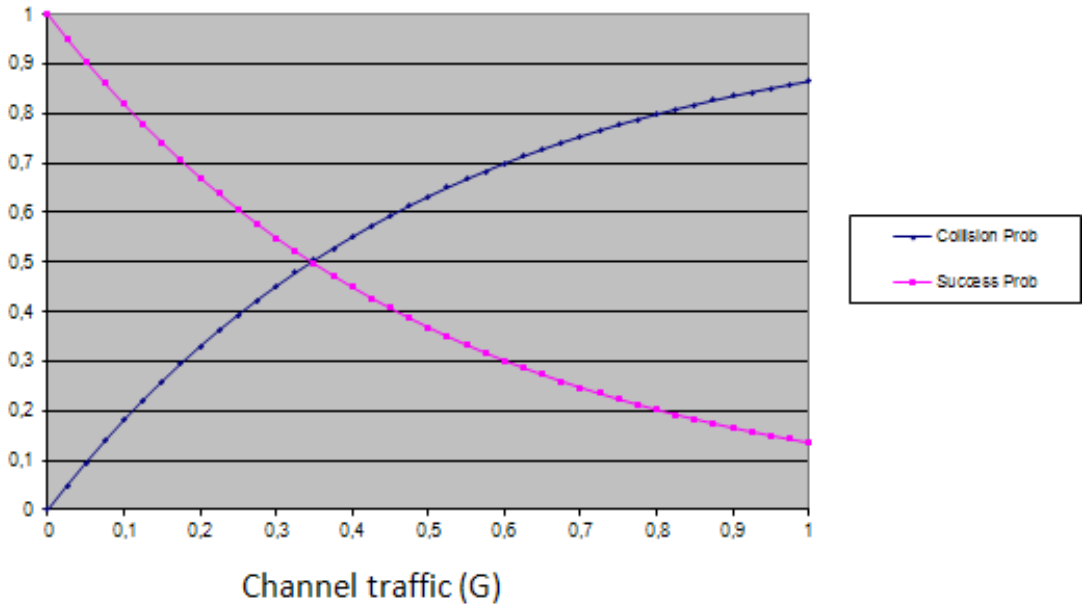


Figure 12: Collision/success probabilities as function of channel loading for classic Aloha

At high traffic loadings the Aloha analysis based on traffic levels and the analysis in 3.1.1 based on discrete users produce the same results for collision probability.

At low channel traffic levels the results differ slightly. This is because in a population of N users the discrete approach considers what happens to one of the N users; the Aloha model considers what happens when one extra user arrives. If N is large the same results are obtained, but if N is small, the discrete approach is recommended. Whichever approach is used it can be seen that, for unidirectional systems without the possibility of collision detection, problems of congestion and collision arise at traffic loadings (or total duty cycle) of 3%.

3.2.2 Variations on Aloha

The analysis above relies on several assumptions, one is that a collision between messages is fatal to both messages, another is that the sending device knows whether or not a message is successful. In the original Aloha system this was achieved by a separate return channel. In wired networks it is done simply by monitoring the line. In wireless systems without a return channel it can also be done by means of a return acknowledgement signal (ACK). If the ACK is very short in comparison to the forward message the analysis is still valid.

In Slotted Aloha the timing of the messages is not completely arbitrary but randomly distributed into slots. If the slots are spaced to match the length of the messages and overheads to control the timing are ignored, then the maximum throughput can be doubled. This is a useful technique for a network of similar devices with a central controller, but it is clearly not applicable to general use in an SRD band.

There is a number of variations of Aloha, according to the manner in which collisions are detected, and the action taken. Collisions may in some systems be detected during the event and the transmission halted; the action taken may, for instance, be an immediate retry or backing off for a fixed or random time. These variations lead to slightly different versions of the throughput formula, and to the many variations of Erlang's equation.

Not all of the variations are applicable to wireless systems. One that is, however, is Carrier Sensing Multiple Access (CSMA). In this, the sending device checks first to see if another device is using the channel; the equivalent in wireless terms is Listen Before Talk (section 3.6 below).

3.2.3 Aloha behaviour with high traffic loading

A factor to note in the basic Aloha analysis is that the “back side” of the throughput curve does not represent a stable situation. If the desired throughput S is not reached, the Aloha system’s response is to send more packets, i.e. to increase G. On the back side of the curve, this leads to lowering of S and therefore to a sort of packet runaway. The operating point moves to the right, with each transmitter in the system sending packets at the maximum rate.

This is illustrated by plotting the behaviour against S (throughput) instead of G (channel traffic).

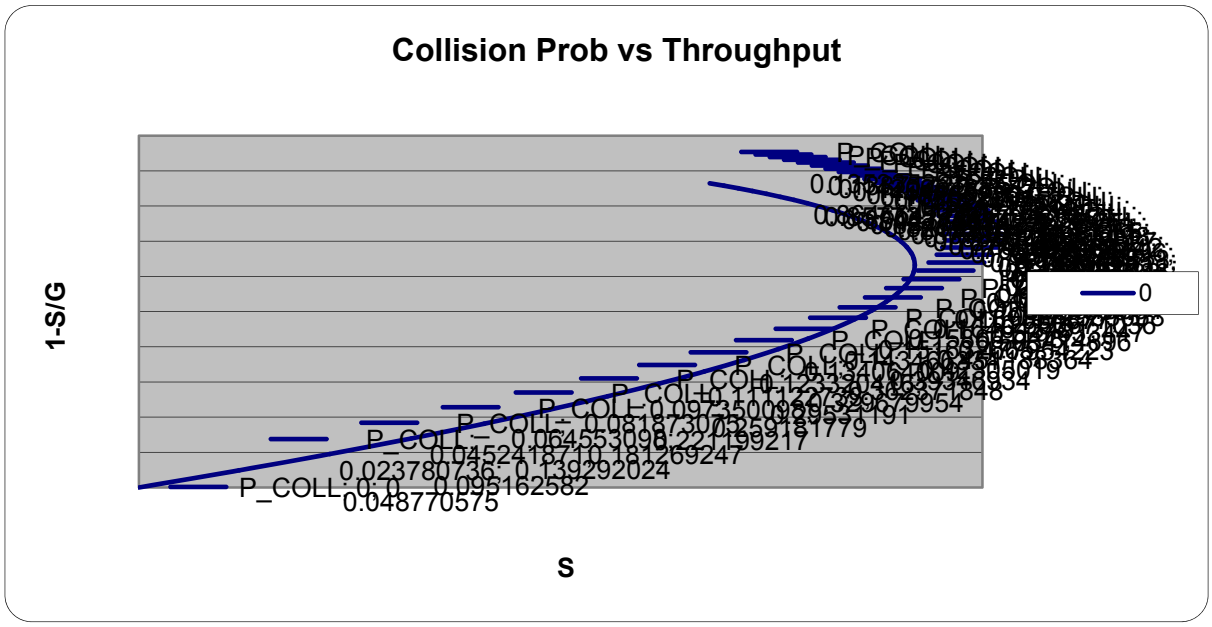


Figure 13: Collision probability vs. throughput

The diagram above shows the collision probability for an individual packet against the system throughput or data traffic. For any value of S up to the maximum, there are two solutions, but only one is stable.

The collision probability increases rapidly as the maximum throughput is approached. When the maximum is reached the system could become unstable. Therefore when presented with too much traffic, collision detection mechanisms as used by Aloha systems can exhibit catastrophic failure rather than graceful degradation (see section 2.3).

To prevent this, it is necessary to ensure that the system always stays on the front side of the curve. Methods for this might include setting a limit on the packet rate for each transmitter in the system, or some form of channel monitoring, or application layer control. It follows therefore that the maximum throughput is not actually achievable; it is necessary to keep the system some margin short of it.

3.2.4 Aloha under Stress

The above analysis assumes that the Aloha system is operating in isolation and that the only difficulty it experiences is collision with its own packets.

In practice, other factors should be taken into account, such as the probability of less than perfect acknowledgements, interference, etc. For instance, if the same channel is used there is a probability that acknowledgements are lost to collisions. The response to any kind of stress is always to transmit more packets, this in turn may aggravate the situation. Care must be taken that this does not lead to the system approaching its capacity limit as described above.

For example, suppose two systems operating Aloha find themselves sharing a channel. Each will react to the other's packets by increasing its rate of sending packets. If the two systems have similar packet lengths, the situation can be analysed by treating them as one larger system with the sum of the throughputs.

For instance two systems each with a throughput of 8% (S=0.08), would each in isolation operate with a value of G = 0.1. Put together the target throughput is 16%, which requires a traffic loading of 28% (G=0.28). This is uncomfortably close to the point of maximum capacity.

3.3 LISTEN BEFORE TALK WITHOUT AFA TECHNIQUES

Listen Before Talk (LBT) is a technique in which a device checks that the channel is unoccupied before transmitting. It requires therefore that the device contains some sort of receiver as well as a transmitter. This imposes a cost penalty, but the reward is hopefully a lower rate of collisions with other users. The receiver can also allow other benefits, such as the use of acknowledgements and return data.

Sections 3.1 and 3.2 above analysed the collision probability when users sharing one channel made blind (i.e., without LBT) transmissions. This section analyses first the effects in the time domain of introducing LBT for users sharing one channel (i.e. without AFA) and then in section 3.3.9 the hidden node and the exposed node problems are analysed.

3.3.1 LBT Analysis in the Time Domain

When an LBT device attempts to transmit a message, there are three possibilities. The transmission is stopped because another signal is detected, or it suffers a collision or it gets through.

$$P_{STOP} + P_{COLL} + P_{THRU} = 1$$

The proportion of actual transmissions that are successful is

$$P_{SUCCESS} = \frac{P_{THRU}}{P_{COLL} + P_{THRU}} = \frac{P_{THRU}}{1 - P_{STOP}}$$

This section is a discussion of the collision probability. It shows the factors that prevent P_{COLL} being driven to zero. Further discussion and the derivations on P_{STOP} and $P_{SUCCESS}$ can be found in Annex 5.

An LBT system may be described by the following parameters:

Listen time T_L
Minimum response time T_R
(T_R is the time taken to detect another signal)
Changeover or dead time T_D
Talk or transmit duration T_{LBT}
Average repetition rate F_{LBT}

Figure 14: LBT device timings

A normal sequence for a transmission is shown above. The transmission T_{LBT} is only made if no signal is detected. For a signal to be detected it must be present during the listen period T_L for a minimum time of T_R.

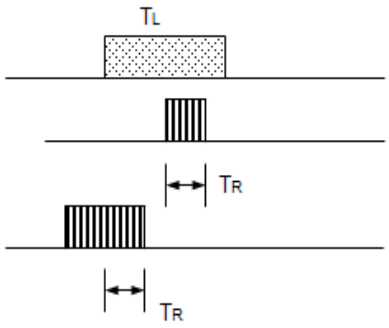


Figure 15: LBT Listen timings

3.3.2 LBT and Duty Cycle

Consider the case of an LBT system occupying the same channel as a system transmitting blind but with a duty cycle limit.

The parameters of the duty cycle (DC) limited systems are:

- Transmit duration T_{DC}
- Average repetition rate F_{DC}

Suppose an individual LBT transmission and an individual DC transmission is related as shown below.

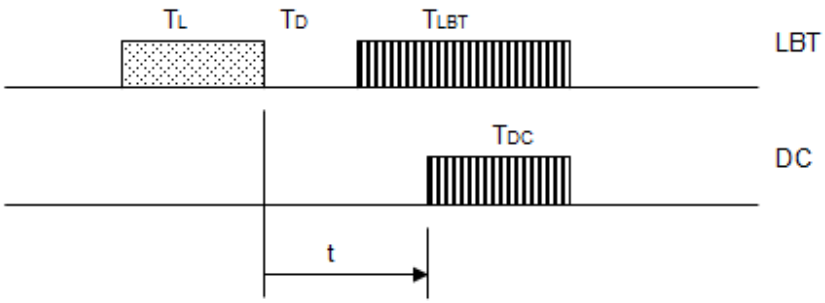


Figure 16: LBT and DC transmissions coinciding

The LBT system detects the DC transmission if it falls in a certain window

$$-\left(T_{DC}+T_L-T_R\right)<t<-T_R$$

Ie, there is a detection window of

$$T_{DC}+T_L-2T_R$$

Detection, however, does not ensure collision prevention.

Assume that $T_D+T_R<T_{DC}$, which will almost certainly be true. (T_{DC} is the message length, whereas T_D and T_R will be similar in practice to bit lengths.)

There is then a collision that is not prevented by the LBT process if the following conditions are met

$$-T_R<t<T_D+T_{LBT}$$

This is equivalent to there being a danger window in the relative timing of size

$$(T_D + T_R + T_{LBT})$$

The probability of a collision is then given by the size of the danger window and the relevant rate of transmissions. The situation is the same whichever party is considered the victim or interferer, since it is assumed that the collision destroys both messages.

Therefore:

Prob of an LBT transmission suffering collision $P_{DC-LBT} = (T_D + T_R + T_{LBT})F_{DC}$ DC on LBT case

Prob of a DC transmission suffering collision $P_{LBT-DC} = (T_D + T_R + T_{LBT})F_{LBT}$ LBT on DC case

Note that the term T_{DC} does not appear in either result. I.e., the duration of the non-LBT transmission does not matter, provided it is longer than $T_D + T_R$; it is the repetition rate rather than the duty cycle that is important.

Note the similarity to the earlier equation for probability in the purely random case (section 3.1.1).

The two diagrams below show the probability of a collision between two users of a channel. One user sends a transmission without using LBT (the DC transmission). The collision probability is plotted against the duty cycle of the other user, according to whether he uses LBT or not.

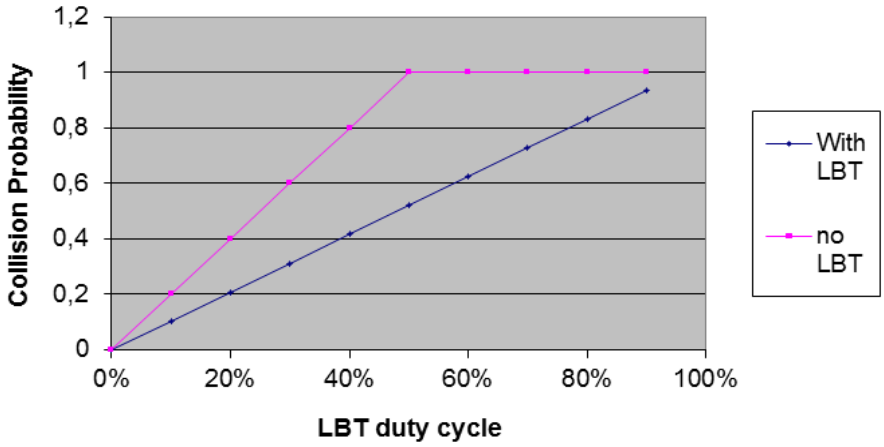


Figure 17: Probability of DC transmission suffering collision (LBT and DC, same duration transmissions)

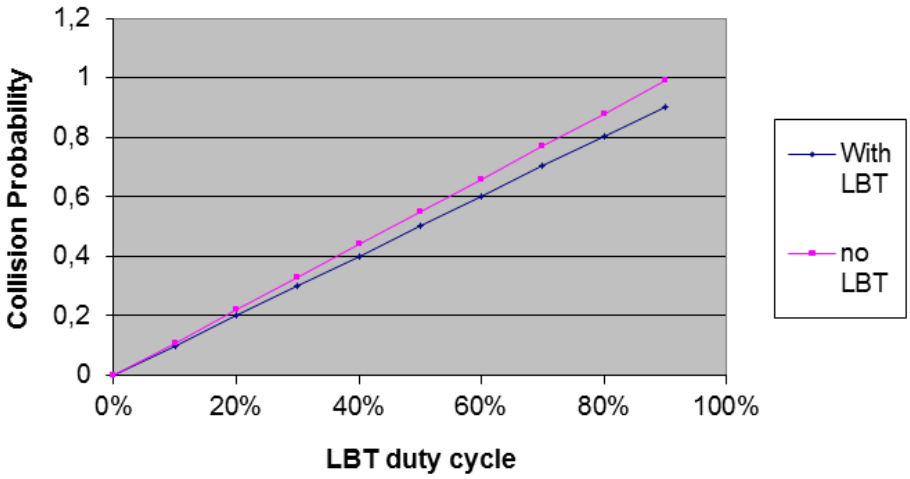


Figure 18: Probability of DC transmission suffering collision (LBT duration 1 sec, DC duration 100ms)

In the first diagram the transmission times are equal and $T_D + T_R$ is small compared to them. In this case the use of LBT by one party reduces the collisions to approx. half of that without LBT.

In the second diagram the message durations are mismatched, the LBT transmission is much longer. In this case the use of LBT has little benefit. A similar effect occurs if $T_D + T_R$ is not small compared to the transmission times.

3.3.3 LBT and LBT

Consider the case of two systems, each operating LBT, that attempt to make transmissions with the relative timing shown below.

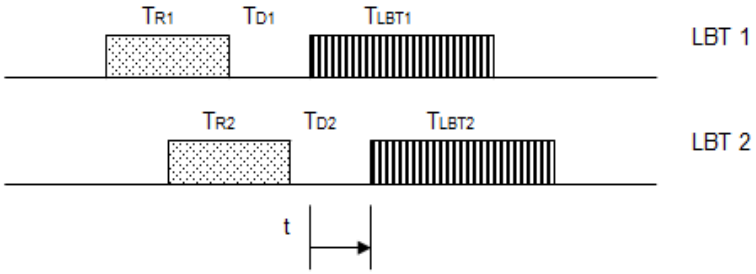


Figure 19: Two LBT transmissions coinciding

A collision occurs and is not prevented by either LBT process if the following conditions are met

$-T_{LBT2} < t < T_{LBT1}$ condition that the transmissions would collide

$-T_{D1} - T_{R1} < t < T_{D2} + T_{R2}$ condition that they would collide and do not hear each other

Assuming Tlbt1 > Td2+Tr2 then only the second condition is important, and the size of the danger window is

$$T_{D1} + T_{R1} + T_{D2} + T_{R2}$$

Therefore:

Probability of LBT 1 suffering interference $P_{2to1} = (T_{D1} + T_{D2} + T_{R1} + T_{R2})F_{LBT2}$

Probability of LBT 2 suffering interference $P_{1to2} = (T_{D1} + T_{D2} + T_{R1} + T_{R2})F_{LBT1}$

Intuitively we might expect that two LBT systems would never suffer collisions with each other. These results show that while the collision probability is very much reduced, it is not driven completely to zero because of the dead times and the reactions times.

The diagram below shows the collision probability between two users running at 5% duty cycle, according to whether each uses LBT, and plotted against the dead time.

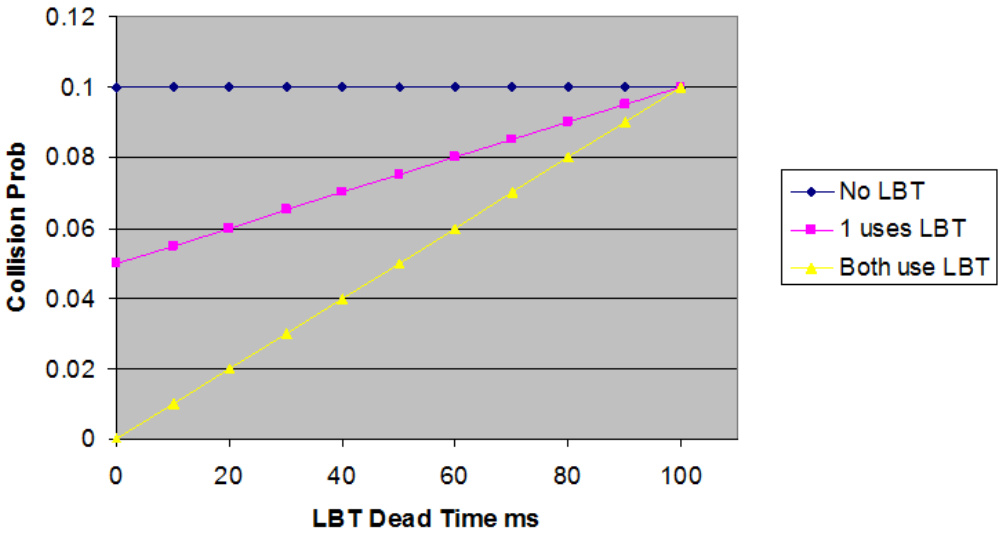


Figure 20: 2 Systems – 100 ms Tx at 5% DC

If the dead time is zero, then the use of LBT by one party reduces the collisions by half, as seen in the previous section. If both use LBT then the collision probability is reduced to zero. But as the dead time increases, the effectiveness of LBT is progressively reduced until it disappears altogether.

EN 300 220 [11] has a provision for spectrum access conditional on the use of LBT. A minimum T_L is specified that varies between 5 and 10 ms and a minimum off time of 100 ms between transmissions is specified. Previously there was no specification for T_D, but as a result of liaison with ETSI during the preparation of this report, an upper limit of 5 ms has been set in v2.3.1 of EN 300 220.

3.3.4 Summary of 2 device analysis

The results for the various scenarios analysed above are brought together in the Table below.
It is assumed that the dead time is small compared to the transmission times; therefore each case the probability of a collision is directly proportional to the dominant factor listed.

Table 4: Summary of analyzing interference probabilities in two-devices scenarios

Interfere	Victim	Dominant factor	Comment
DC	DC	$T_{DC}F_{DC}$	Duty cycle of DC transmission (assuming duty cycle is low)
DC	LBT	$T_{LBT}F_{DC}$	Duration of LBT transmission, and Rate of DC transmission
LBT	DC	$T_{LBT}F_{LBT}$	Duty cycle of LBT transmission
LBT	LBT	F_{LBT}	Rate of LBT transmissions P_{COLL} is low in most circumstances

Note that this analysis is based on the probability of collision when both interferer and victim can hear each other.

Although a collision works both ways, DC-LBT interference has two rows in the table. This is because the probability is expressed in terms of the risk to an individual transmission from one party.
Also, no examination has been made of the consequences of a collision. It should be noted that the probability of collision is not necessarily the best or only measure of the harm done to the victim. It may also be noted that not all collisions are fatal to both parties, although in these circumstances a large proportion may be expected to be.

Nevertheless, some features of the Table are worth highlighting.

- In terms of collision risk to a DC user, the important factor is the duty cycle of the interferer.
- But in terms of collision risk to an LBT user this is not so. The important factors are then:
  - Duration of LBT transmissions
  - Rate of LBT transmissions
  - Rate of DC transmissions

3.3.5 Multiple devices

Extending the above analysis to multiple devices – both populations of common access methods and mixed populations – is not always tractable, and so it is necessary to either use numerical methods or analytical methods for a restricted set of cases.

Two techniques have been developed as part of the preparation of this work: a quasi-Monte Carlo analysis; and a probability analysis intended to analyse the general case.

The associated spreadsheets with numerical simulations may still require validation. In particular it has to be noted that different definitions of efficiency and throughput may have been used than those introduced at the beginning of this report. The applicability of the spreadsheets to particular circumstances also needs to be established.

Quasi Monte Carlo analysis

This tool uses statistical techniques to calculate, numerically, the way in which both multiple LDC systems and LBT and LDC interact/interfere with one another. The tool predicts:

- LBT Temporal Spectrum Use Efficiency
- LBT Throughput
- LBT P99.9% back off delay
- DC Temporal Spectrum Use Efficiency
- DC Throughput
- DC P99.9% back off delay
- Sensitivity TPR = Victim Protection Rate.

Current limitations of the technique are that all devices are located in the same domain, each device can "hear" all others, all devices have the same parameters for frequency, bandwidth, power, timing, there are no effects from outside the domain, and packets that require re-transmission are discarded, leading to both false positive and false negative events. The impact of these limitations is calculated and based on that the percentage uncertainty in the results is estimated.

Further details of the technique are shown in Annex 3.

Probability Analysis

This tool extends the two-LBT collisions probability analysis to multiple systems by modelling re-transmissions as statistically independent transmissions. The tool predicts for various sharing scenarios:

- Probability of collisions for single attempts
- Probability of collisions for multiple attempts
- Impact of transmit times (Ton) on probability of collisions
- LBT wait times.

Current limitations of the technique are that the analysis is applicable to the non-persistent LBT retry mechanism only, and where the retry times are, on average, longer than the arrival times of other packets.

Further details of the technique are shown in Annex 4.

The overall effect when multiple devices are present also depends on the strategy or protocol followed when a competing transmission is detected. This is discussed below in section 3.3.7.

3.3.6 Simulation of non-persistent LBT operation

The most important difference between DC and LBT is the behaviour in the time domain. Complementary to the analytical approach in previous chapters a numerical simulation of time domain behaviour has been carried out.

The basic idea of this simulation is to calculate the number of recognised or not recognised collisions for a set of devices using the same RF parameters and a common spatial range. Hence, mutual communication between all devices in the simulation is possible and no propagation effects are taken into account. The mitigation methods compared in the simulation are duty cycle and LBT without AFA.

The mathematical approach is based on a Monte Carlo Simulation. For simulation of collisions, a set of random numbers is mapped to individual transmit times for each device within a common transmit interval. Every device will only transmit once per transmission interval. For LBT devices detecting a signal, the transmission is suspended, as the devices are assumed to be energy limited. The LBT parameters used in simulation are listen time, dead time, recognition time, transmit time and duty cycle. A detailed description of this simulation, which is performed in a spread sheet, can be found in Annex 3. The results carried out by the simulation are obtained from signal detection theory to assess the ability of the LBT spectrum access method to detect potential collisions. Besides some statistical figures the receiver operating characteristic is used to visualise the diagnostic capability of LBT as a test method. This diagram shows the true positive rate vs. false positive rate of recognition which is in case of guessing a relation of 1:1 and in case of perfect recognition a false positive rate of zero or a true positive rate of 100%.

In example, three representative cases are considered. The LBT parameters used for these simulations are based on the targets set out in the current version of the relevant standard EN 300 220-1 [11].

3.3.6.1 *Very short transmission, low duty cycle*

Transmit Time	5 ms
Duty Cycle	0,1%
LBT Listen time	7,5ms
LBT Dead Time	1 ms
LBT Sample Time	0,1 ms

These parameters are, for example, typical for systems using short burst transmissions like (sub-) metering devices.

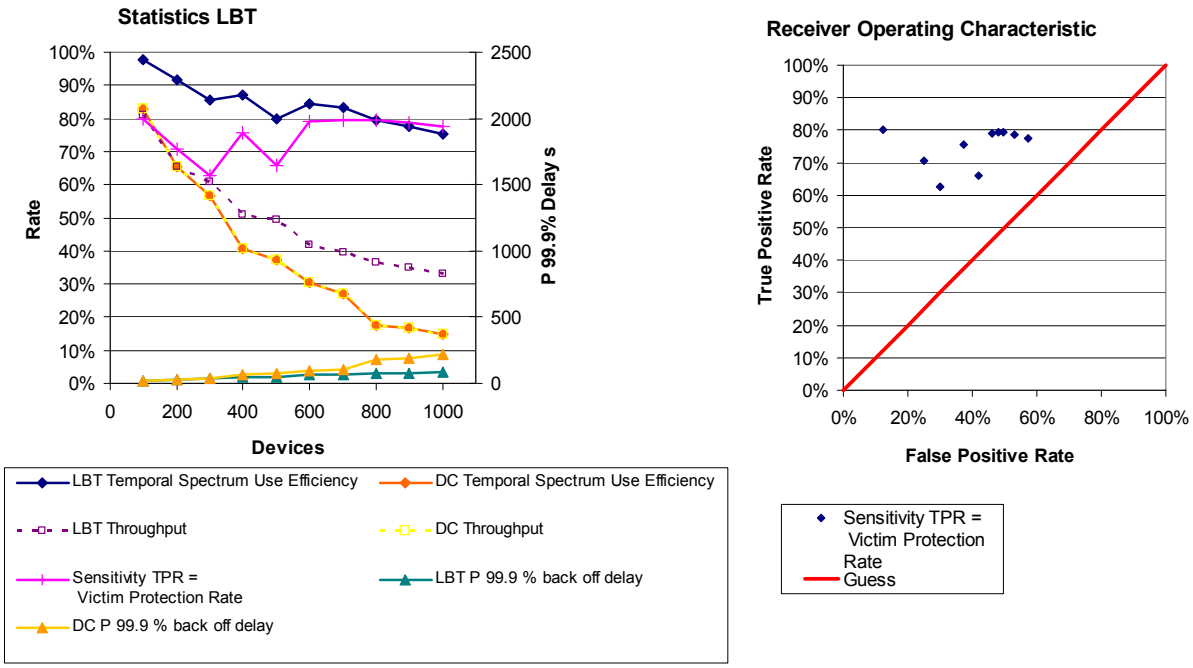


Figure 21: Results of simulations for very short transmission, low duty cycle

Up to a population of about 300 devices there is no significant difference between LBT devices and DC devices. The throughput and the back off delay, i.e. the time which is needed to get 99.9% of the transmissions successfully sent, and which results from the multiple repetitions of lost and retained transmissions, are very similar.

The receiver operating characteristic of the listen mechanism is rather close to a guess. This comes from the relation between listening and dead time and the duration of the transmission. Numerous LBT transmissions are retained although there would be no collision if they would be transmitted. Expressed in statistical terms this behaviour is represented in a relatively high false positive rate.

When the occupancy of the channel increases the advantage of LBT devices over DC devices becomes more visible. But the performance gain is not based on the ability to detect potential collisions. In fact the LBT device detects a signal of any other device and retains its own transmission. Thereby a gap is created which can accidently be filled in by another device.

On the other hand there remains a certain quantity of collisions which cannot be detected.

The limited functionality of LBT in this case is also reflected by the relatively low sensitivity (also known as True Positive Rate), which is a measure for the ability to protect a potential victim from interference.

3.3.6.2 *Short transmission, very low duty cycle*

Transmit Time	25 ms
Duty Cycle	0,02%
LBT Listen time	7,5ms
LBT Dead Time	1 ms
LBT Sample Time	0,1 ms

These parameters are typical for systems using short transmissions in combination with rare transmissions, e.g. battery operated metering systems.

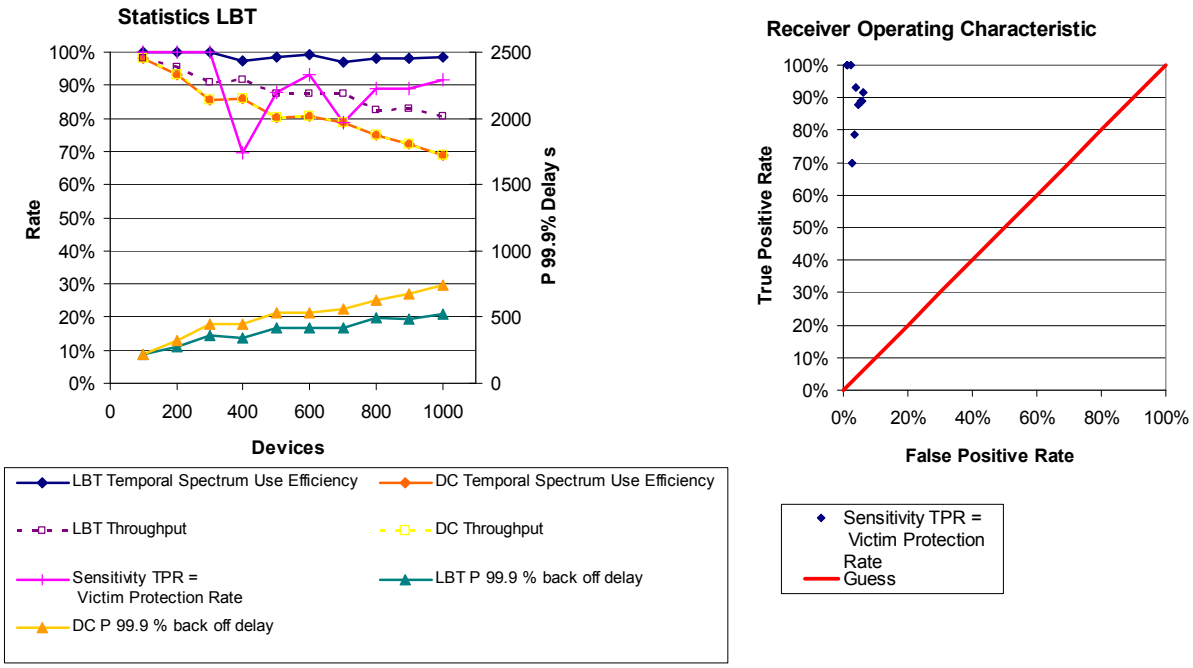


Figure 22: Results of simulations for short transmission, very low duty cycle

Similar to the previous case there is no significant difference between LBT devices and DC devices up to a population of 300 devices. The throughput and the back off delay are very similar.
However the receiver operating characteristic reveals a good ability of the listen mechanism to detect a potential collision whereas the ability to detect all relevant collisions is limited to a certain extent, because the duration of the transmission is significantly longer than the listen time. Nevertheless due to the very low duty cycle the advantage of LBT is limited as long as the occupancy of the channel is low.

3.3.6.3 *Medium duration of transmission, medium duty cycle*

Transmit Time	50 ms
Duty Cycle	1%
LBT Listen time	7,5ms
LBT Dead Time	1 ms
LBT Sample Time	0,1 ms

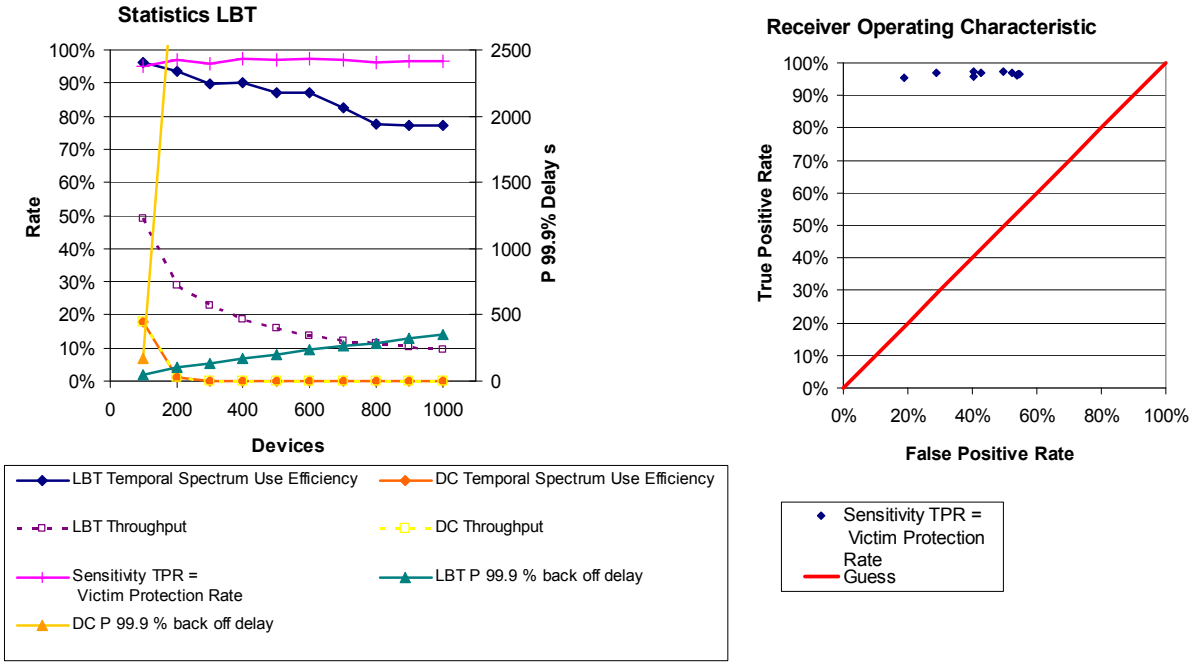


Figure 23: Results of simulations for medium duration of transmission, medium duty cycle

LBT works very well in these cases. The throughput of the LBT devices is significantly higher than the throughput of the DC devices, and the 99.9% back off delay of the LBT devices is low compared to the DC devices, which even are not able to achieve 99.9% successful transmissions when the number of devices exceeds 100.

From the receiver operating characteristic it can be seen that the ability of the listen mechanism to detect potential collisions is really good.

Yet it should be noted that in case of a duty cycle of 1% the channel is overloaded, when the number of devices exceeds 100.

3.3.6.4 Preliminary Conclusions

In many real SRD systems it may not be feasible for supply power limited (i.e. battery-driven) devices to re-transmit. The calculations given in this section were developed for such devices. It implies that “no re-transmission” was implemented if the LBT (without AFA) device detects another device and it abandons the transmission. This is not representative of either of 1-persistent or non-persistent LBT.

Under those assumptions, in cases of very short transmissions or very low duty cycles there is no significant advantage of systems using LBT over systems using DC as long as the occupancy of the channel does not exceed a certain fraction of the total capacity. LBT brings a benefit to systems which have to transmit longer packets of data than typical low duty cycle systems or are operating in high occupancy channel.
The tool was not developed in order to model non-persistent LBT. The implementation of non-persistent LBT in the tool is still under consideration and should be further considered since there are diverging views on its implementation within the tool. SE24 considered the implementation of non-persistent LBT and no conclusions were drawn yet. This could be further considered toward another work item.

The LBT parameters set out in the current version of EN 300 220-1 [11] seem not to be suitable for an optimized performance of the LBT mechanism. Particularly the listen time should be shortened to reduce the false positive rate of the listen mechanism. This may need to be considered further, in particular in the framework of ETSI.

3.3.7 Throughput with Carrier Sensing

Aloha, discussed above in section 3.2, is a scheme relying on Collision Detection (CD). The loss of a packet is only discovered after the event. LBT provides a means of Carrier Sensing (CS), in which the potential clash is discovered before the event. The use of CD and CS together leads to a class of protocols known as Carrier Sensing Multiple Access-Collision Detection (CSMA-CD).

The throughput with LBT depends on the strategy or protocol followed when a device detects another user. One of the simplest protocols is to keep checking the channel and transmit as soon as it is free. This is known as 1-persistent CSMA because the device transmits with probability 1 when the channel is free. The difficulty with this protocol in a high traffic environment is that if two devices are waiting, they may both start transmitting simultaneously. There are therefore variants of this protocol employed with different probabilities.

Another protocol is non-persistent CSMA, in which, after detecting another user, a device will back off for a random time before retrying. The other protocols, known generally as p-persistent, apply only to slotted systems. Therefore in the diagram only Pure Aloha, 1-persistent CSMA and non-persistent CSMA are relevant to SRDs without a central controller.

The diagram below shows the performance of various schemes in a wired environment or where all devices can hear each other.

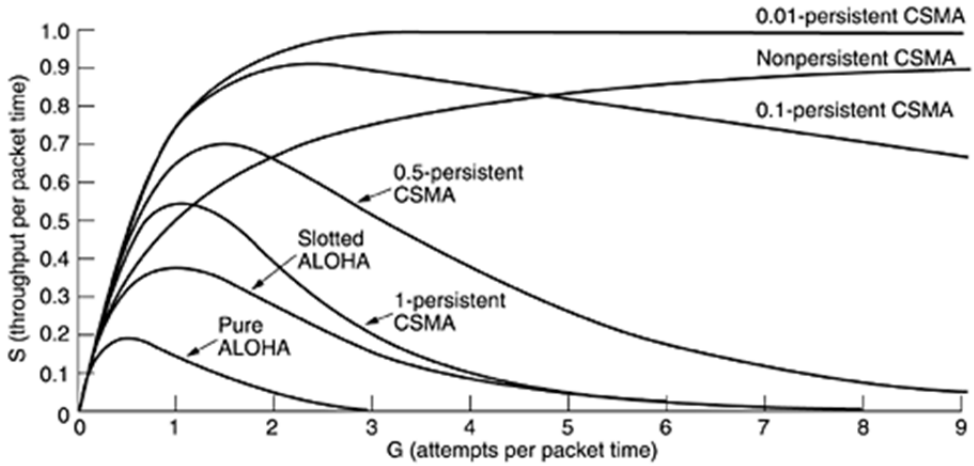


Figure 24: Throughputs of various CSMA protocols. [12 – figure 4.4]

Although the use of LBT can be mandated in the regulations for an SRD band, it is more difficult to go further and apply back off protocols. Such protocols do exist in specifications such as IEEE 802.11 [13] but it is felt they would be difficult to apply in the case of general purpose SRDs. Key difficulties are that the length of transmissions is undefined and that probabilities would be extremely difficult to test. The choice is effectively between 1-persistent CSMA and non-persistent CSMA, and it would be difficult to enforce that in regulations or standards for SRDs. In small networks and for single devices, the cost of waiting or backing off is borne by the individual device but the benefit accrues to everyone, so it might be expected that most devices would choose 1-persistent. The use of any other would be voluntary, and could be expected to be adopted only in large networks of common devices.

It is worth noting that 1-persistent CSMA shows the same fold back in the curve as Aloha and therefore would, unless prevented, show catastrophic failure at high traffic, whereas non-persistent CSMA would show graceful degradation (at least up to the traffic levels shown).

Another thing that the diagram shows is that, to achieve high levels of throughput, CS in addition to CD is required. CD alone means there is always an underlying level of collisions that sets a limit to the overall performance. In extreme cases, however, other strategies such as changing channel may be more useful.

3.3.8 Summary of LBT timing issues

Because of the various timing issues analysed above, LBT is not 100% effective at avoiding collisions. With a device population consisting entirely of LBT devices there is still a residual probability of collisions because of the receiver response time and the dead time in the changeover from listening to transmit. Yet the use of LBT enables operation at higher occupancy and throughputs than either DC or Aloha.

The probability of collision does however increase with an increase in channel occupancy. The exact upper limit depends on the detail of the system and the protocol chosen, but is likely to be of the order of 50% occupancy or 50% throughput.

When LBT and non-LBT devices share a channel, the LBT operation reduces the collisions suffered by both devices. I.e., LBT provides a benefit to both the device using it, and to others on the channel. However, the benefit to each party is not as great as when both use LBT.

Therefore, LBT and DC devices can successfully co-exist, as long as the LBT devices operate at the same duty cycle patterns. The utilisation in this case is always better than with DC only.

3.3.9 Hidden and Exposed Nodes

This section discusses the so called “hidden/exposed node problem” and does not consider effects in the time domain; those timing effects are analysed in previous sections.

The diagram below represents an idealised space in which a victim receiver VR is receiving messages from the wanted transmitter WT over a distance Rsig. A potential interferer IT (which has receiving capabilities for LBT) is randomly placed.

For simplicity, it is assumed there are no polarisation, propagation or antenna pattern effects, so signal strengths are related to distance.

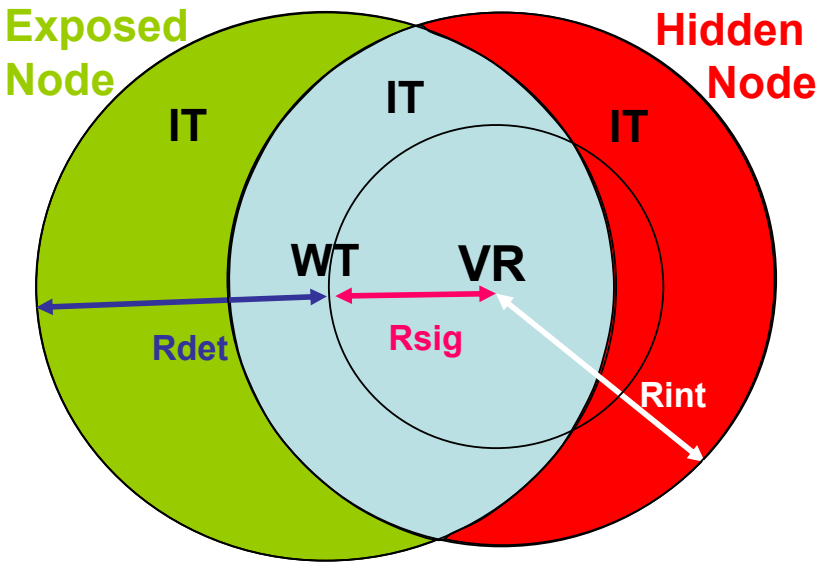


Figure 25: Graphical depiction of LBT sensing (exposed node) vs dead (hidden node) zones

Within a radius of Rint around the VR the IT can exceed the protection objective of the VR (e.g. C/I). Within a radius of Rdet around the WT the IT can detect the WT.

In the light blue area in Figure 25 LBT is working effectively. The red area is the so called “hidden node”, where the IT is not able to detect the WT. The green area is the so called “exposed node”, where the IT detects unnecessarily the WT. The scales of the circles in Figure 25 are arbitrary.

There are 3 main parameters which mainly impact the hidden node issue: the LBT threshold, the Tx power of the victim and interfering link and the SNR for the victim link.

For a balanced Tx power situation, a realistic threshold value of -87 dBm and victim links having a high SNR (>35dB) the hidden node probability is very low, while with a low SNR (<15dB) the hidden node probability is very high. The next figure shows the detailed results for a specific set of parameters and shows also the dependency on the propagation model (exp 2= Free space loss).

These results are extracted from the analytical study in Annex 1. Further material is also provided in Annex 2 (SEAMCAT simulation).

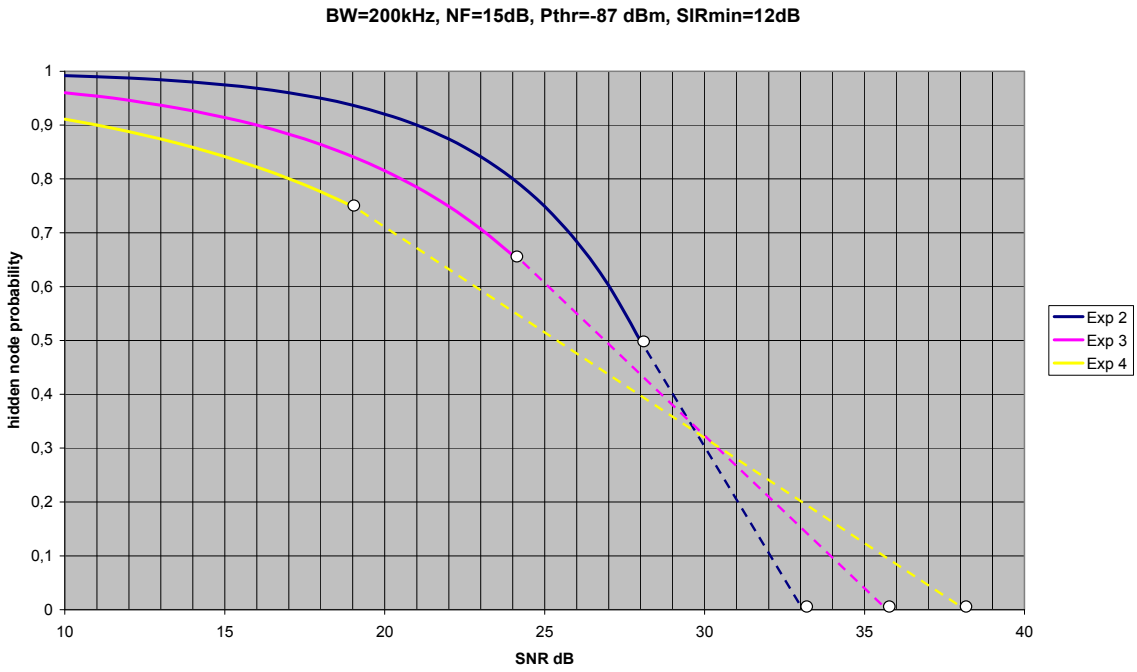


Figure 26: Probability of hidden node occurrence as function of SNR in victim link

The unbalanced Tx power situation is even worse; here the hidden node probability for victim links having for example 20dB less TX power than the interfering link is close to 100%, even with a high SNR.

It is important to consider what the changing SNR means in a real system context. The SNR will normally vary according to the distance of the wanted link. If the receiver is close to the transmitter, the SNR will be high and the hidden node probability correspondingly low. On the other hand, if VR is far from its serving WT, the SNR will be lower and the danger of a hidden node effect increases.

If the victim link distance is held constant, there is a similar effect if WT applies adaptive power control or APC (see section 4.6.), ie., it reduces its power if there is excess signal strength at VR in order to reduce its own interference footprint.

In this case the relationship between r_{int} and r_{sig} is broken. As WT reduces power, r_{det} goes down and r_{int} goes up. This effectively guarantees a hidden node problem. In an attempt to be neighbourly, the wanted system ends up undermining its own operation.

APC, however, can only be used by a bi-directional system. If reducing the power results in interference, the system will detect that and increase the power.

The next figures give an illustration of how the results depend on the SNR ratio for a propagation exponent 4. In the light blue area LBT is working effectively; the red area is the hidden node and the green area the exposed node. The diagrams show how, as the signal strength increases, there is a shift from hidden nodes

to exposed node which will reduce the absolute spectral efficiency that can be achieved by a factor of about two by preventing, unnecessarily, approaching half of the nodes from transmitting.

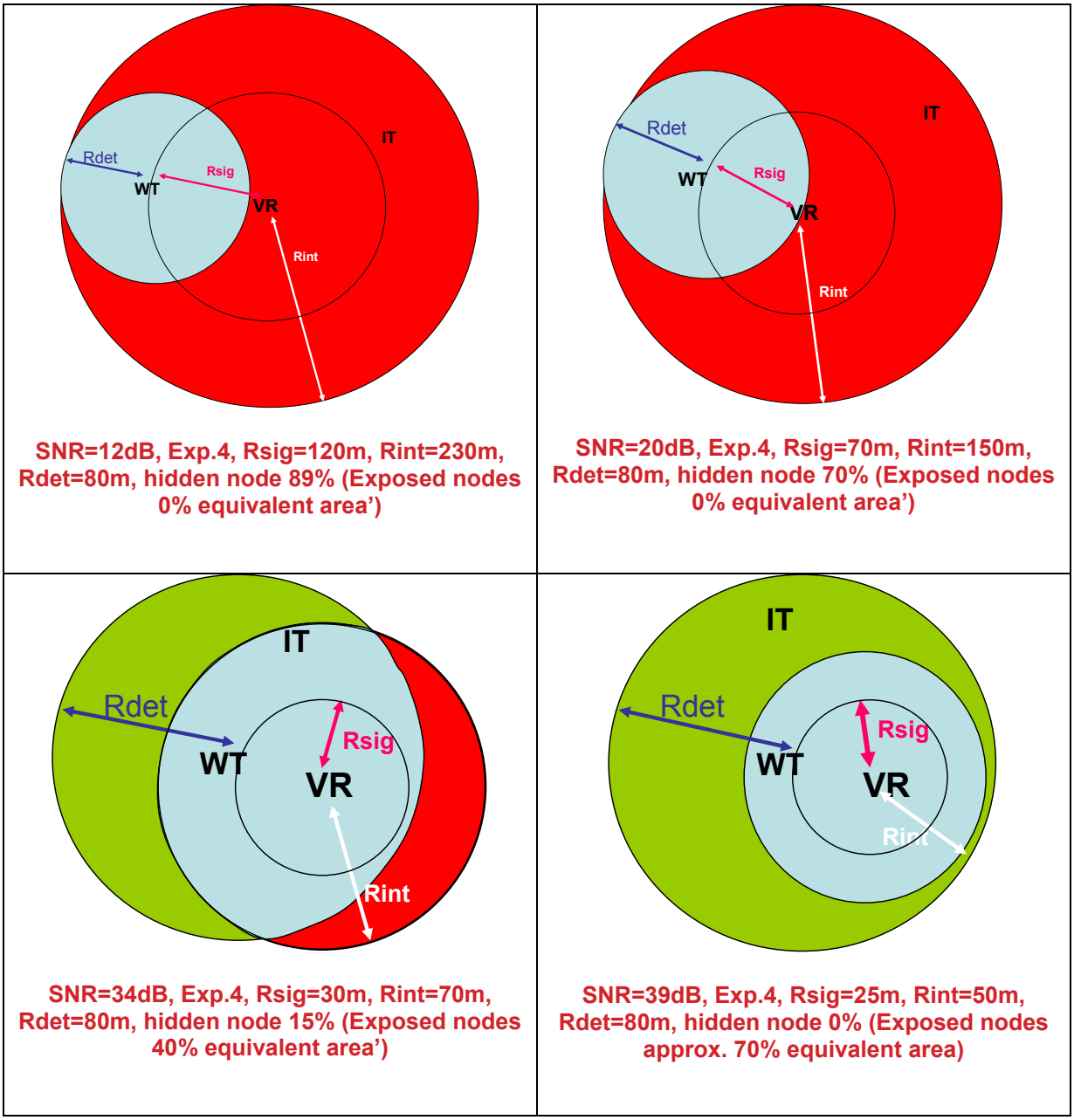


Figure 27: Illustration of the exposed node problem

The hidden node problem means that LBT is not a complete solution to interference avoidance, i.e. it is not as effective as it would otherwise be at preventing collisions. In effect, the occurrence of hidden node means that LBT has failed to do any good, but has not done any harm provided that the LBT parameters are appropriately set. Further work is needed to derive the appropriate LBT parameters that would minimise the probability of hidden node occurrence, such as sensing threshold.

The corollary of the above, the exposed node problem means that a transmission is prevented when it would otherwise have gone ahead without any problem. How serious the effect of this is on internal performance of the LBT system depends on the circumstances. If it reduces the throughput of the considered LBT system, then it is significant. If, on the other hand, it simply displaces transmissions in time, then the harm done may be very little.

The overall conclusion is that the LBT is not ideal in that it may not always protect the peer/victim but also the LBT-equipped SRD may sometimes itself suffer by hindering its own operation unnecessarily. The key may therefore be found in considered application of this technique, including the possibility of combination with other mitigation techniques (e.g. ACK techniques…).

3.3.10 Cost and Benefits of utilising LBT

In order to guide regulators and manufacturers as to whether LBT inclusion is justified, it may be worth to address the question whether channel sensing provides potentially useful information and to ask in what circumstances it makes sense to not have that information. The answer is simply when the cost of collecting the information is more than its value.

In most cases the cost of collecting the information is low - a few milliseconds spent listening - and the value is high - it increases the probability of successful operation.

An instance where the value of the information might be seen to be low would be a low duty cycle device operating in a channel with very low occupancy. But the device does not know that without collecting the information.

Two instances where the cost of collection is not trivial should be considered. One is the case of transmit only devices. These form an important and well established class of SRD; they function well in low occupancy environments. Forcing or expecting them to use LBT would not be appropriate.

The other is battery powered, or energy limited devices. Listening, or channel sensing, consumes energy to power the receiver. Against this cost must be set the potential benefit that the number of transmissions could be reduced or a wake up period shortened. For such devices the cost-benefit balance would depend on the occupancy of the channel.

3.3.11 Summary LBT

Advantages of LBT:

- Reduced probability of interference through avoiding (to variable degree) collisions with neighbours' transmissions and, thus, overall positive impact on sharing scenarios, either through obligatory or voluntary implementation of LBT;
- Increased throughput at higher channel loads: when the channel occupancy exceeds about 10%, various forms of LBT are able to offer improved success rates and therefore higher throughput.
- When a DC device additionally operates LBT, the LBT itself may or may not do any good, but it generally does not do any harm to another user of the spectrum.

Disadvantages of LBT:

- Mechanism short comings: LBT is not able to avoid all collisions due to the limited listen time, the inability to receive and transmit at the same time. The results of time-domain analysis shows there is still a residual probability of collisions.
- Hidden nodes: sensing only at the transmitter and the non-ideal power threshold may lead to the hidden node problem, which means that victim receivers may not always be protected despite the interfering transmitter obeying the rules of LBT. Numerically, the probability of an individual victim receiver being a hidden node (and hence not being protected) can vary between 0 and 100% dependent on LBT threshold and power balance within victim system vs interfering system.
- Exposed nodes: LBT can also cause problems when it stops transmission when they would have succeeded without interfering with any others.

3.4 DIVISION BY FREQUENCY – CHANNELISATION

The previous discussions on duty cycle, Aloha and LBT are all examples of techniques for sharing in the time domain, and fall under the general title of Time Division Multiple Access (TDMA). The equivalent in the frequency domain is FDMA or Frequency Division Multiple Access.

In its simplest form FDMA just means users occupying different channels where they can operate completely independently. A user on one channel can run up to 100% duty cycle without affecting a user in another channel. There are, however, key aspects of isolation and organisation.

3.4.1 Isolation

In TDMA, isolation is all or nothing. Two packets that do not overlap in time are completely isolated. In FDMA, the isolation between frequency channels is generally finite and is set by transmitter and receiver performance.

The isolation by frequency seen in SRD systems varies widely. For instance, in sub bands with 25 kHz channels (considered as narrowband for the purposes of EN 300 220 [11]), there are restrictions on the transmitter adjacent channel power and high performance receivers are often used. This may result in a receiver being able to achieve a rejection of the transmitter signal in an adjacent channel of 70 dB. If the receiver requires a C/I ratio of 15 dB to operate, then the isolation achieved is 55 dB. This level of isolation allows a signal from a wanted transmitter to be received even when an unwanted transmitter is much closer. Even in a hotspot it would generally be possible to use all the channels.

In the non-narrowband sub bands, the situation is different. There is generally no channel structure to work to, and usually no set values for either bandwidth or channel spacing. Instead of an adjacent channel power specification, EN 300 220 [11] sets limits on the transmitter spectral density at the sub band edges.

In these circumstances, it is common to find relatively low levels of isolation between nominal channels. For instance a receiver on one channel may only achieve 40 dB rejection of a transmitter in the adjacent channel. If the required C/I is 20 dB, then the isolation is only 20 dB.

In some cases it can be even worse, depending on the definitions used for channel spacing and bandwidth. If channel spacing is set to close in relation to the bandwidth, or if the transmitter spectrum or receiver filtering is too wide, then systems with zero or negative isolation values are created. What happens then is that in a hotspot, the system can only use every 2^{nd} or every 3^{rd} channel.

Similar inefficiencies in spectrum use occur if the transmitter or receiver bandwidths are set excessively large in relation to the data rate, for instance to accommodate poor frequency stability or to reduce cost.

Whereas in TDMA, once the signals are separated in time, complete isolation is obtained, with FDMA there is a strong correlation between isolation and the cost of equipment.

3.4.2 Organisation

Selection of channels is also important. In some systems this may be pre-planned or dictated by a network controller, neither of which is appropriate for the SRD bands.

If each user chooses a frequency at random then, in an unstructured band, it only takes occupancy of 2.5% for each user to have a 5% probability of suffering a frequency overlap (it is the same mathematics as for random packets in the time domain). In practice, if no attempt at channel organisation is made, the users are not spread randomly but tend to congregate around given frequencies. An example is low cost devices in the 433 MHz band and the 868.0-868.6 MHz sub band all using SAW devices at the centre of the band. Even after the widespread use of synthesisers, devices still commonly target the centre of the band.

An early attempt at organised channel selection, popular with VHF SRD telemetry systems, was to conduct a site survey and then choose a fixed frequency based on the result.

Nowadays intelligent automated selection techniques are possible and are known by names such as Cognitive Radio (CR), Detect and Avoid (DAA), Dynamic Frequency Selection (DFS) and Adaptive Frequency Agility (AFA). These are all similar in that one or more devices monitor the band and choose an operating frequency on the basis of what they hear. The terms DFS and AFA are associated more with rapidly changing environments and CR and DAA with static or slowly changing ones.

Channel selection can be affected by the Hidden Node Problem and the Exposed Node Problem (see section 3.3.9). Channel selection differs from LBT, however, in that it does not operate on a packet by packet

basis. The Hidden Node Problem is less severe because a device only has to hear one side of a bi-directional exchange in order to move away. If two systems are each using channel selection then the Hidden Node Problem disappears – if neither system can detect the other then there is no possibility of interference.

Similarly, the Exposed Node Problem disappears in the right circumstances. If each device ignores the effect of its outgoing transmissions and only initiates a frequency change if it detects a signal that would interfere with its reception, then there is no unnecessary frequency changing.

3.4.3 FDMA Summary

FDMA is a very effective and well established sharing technique. Pre-planned FDMA is the basis on which all analogue sound and TV broadcasting works. Automatic or adaptive FDMA makes it possible to share traffic in the SRD bands the same way.

FDMA does not necessarily suffer the same throughput limitations as TDMA. Unless there is a central controller, all the TDMA techniques discussed above show severe quality of service issues at a fraction of the theoretical throughput (round about one-sixth in the case of Aloha). By contrast, FDMA has the capability to deliver high quality of service almost up to the theoretical throughput limit.

3.5 SPREAD SPECTRUM

Spread spectrum systems are conveniently divided into Frequency Hopping (FHSS) and Direct Sequence (DSSS). Traditionally a spread spectrum transmission is defined as one in which the bandwidth used is many times greater than the bandwidth required by the data rate. This, however, could also include transmissions with excessively wide modulation or poor frequency accuracy, both of which have been encountered at the low cost end of the SRD world. It is useful, therefore, to add a requirement that there is a corresponding processing gain in the receiver when the bandwidth is reduced.

3.5.1 Frequency Hopping

In FHSS the radios hop through a set of channels according to a pseudo-random sequence peculiar to each transmitter/receiver. For transmissions extending over a number of hops, this effectively mixes up the time and frequency domains. In terms of sharing studies, therefore, traffic on each channel is reduced by a factor equal to the total number of channels available. Note that as a result of this spreading of interference over time/frequency, any particular fixed channel transceiver working in the subject shared band should be normally suffering less interference from FHSS systems if DC rules for a single system are applied to the total occupancy figure of the FHSS system.

Within the FHSS category, there are two important sub-categories, namely “slow hopping” vs. “fast hopping” systems. In the former, the most classical example, the hopping rate is much lower than the rate of transmitted information. Typically, this could be described as a system hopping into particular channel, transmitting a burst/packet of information, and then hopping to another one and so on. In the fast hopping system, the rate of hopping is comparable to or higher than the information transmit rate. So e.g. the single bit of useful information may be transmitted while in transition between more than one hop. As a result of this, in terms of interference impact the fast hopping FHSS becomes similar to DSSS. Therefore, it is important to keep in mind that the following analysis mostly concerns classical slow-hopping FHSS systems.

When considering the operation and sharing potential of FHSS systems, it is important to distinguish two types of FHSS:

- Generic (plain) FHSS;
- Hybrid FHSS (i.e. FHSS combined with other mitigation techniques, such as DAA).

For more background on the origins, functioning and differences between various FHSS types as seen in the SRD field, please refer to Annex H of ECC Report 37 [1].

3.5.1.1 Generic FHSS

The generic FHSS is a most traditional and most commonly met type of system in which a single carrier is hopped among a number of discrete frequencies. In the Generic FHSS, the number/range of hopping

frequencies is pre-determined during the manufacturing or operational set up of the system. Then during the operation, the sole mechanism of interference mitigation is spreading of transmissions over the broadest possible range of hopped channels, without any selectivity associated.

In this case, for two FHSS systems sharing the same band (or for one FHSS and one fixed frequency) the collision probabilities are exactly as derived in section 3.1, after allowing for the sharing out of the traffic.

In the below figure the Frequency Hopping sequence example is shown with the channels spread across a spectrum and 'hopped' in a pseudo-random manner. In practice, the number of channels would usually be chosen to hop to every available channel once per sequence.

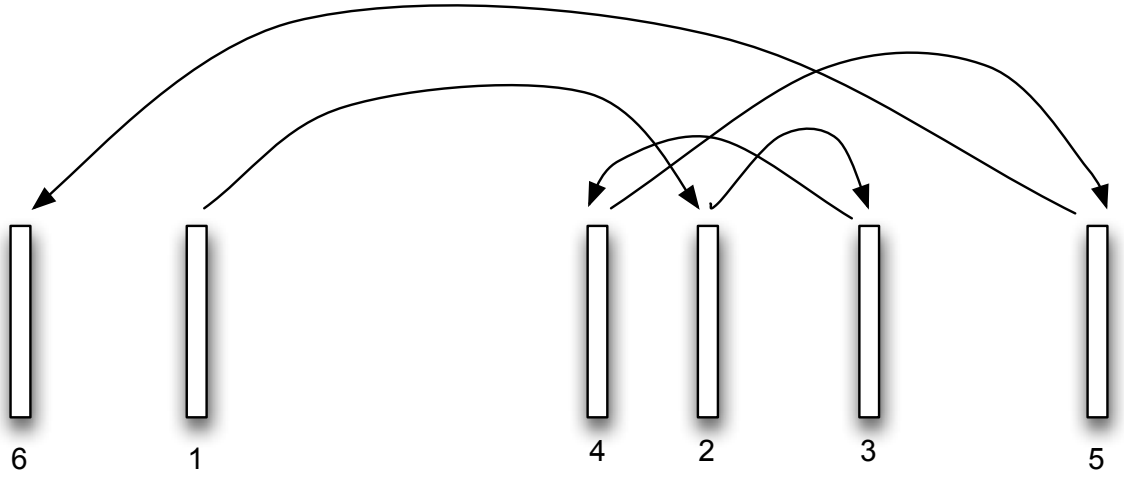


Figure 28: Frequency Hopping sequence example

Thus a generic FHSS uses all channels equally for communication. This provides fair and equitable spectrum sharing in cases of sharing with DC based systems with using the same DC rules. However when frequency band congestion increases, the operational efficiency (reliability of transmissions) of the generic FHSS may deteriorate due to the fact that more and more channels in the hopping set will be occupied by some other systems. Also, the generic FHSS may unnecessarily create constant levels of residual interference (collision instances) when sharing the band with some fixed channel systems, which in repetitive hopping sequences will lead to quasi-stationary pattern of interference on a given channel of victim fixed frequency system. In order to improve these aspects of FHSS performance, some more advanced types of FHSS may be deployed, namely making hybrid combinations of FHSS with some environment sensing/coordination functions.

The description above applies if there is a common channel plan and bandwidth. Where this is not the case, there are issues of mismatched bandwidth and overlapping channels to be considered.

3.5.1.2 Hybrid FHSS

The drawbacks of generic SRD end to end FHSS at higher levels of band occupancy may be addressed by providing the system with the ability to sense the environment and exclude the channels occupied by other systems from the hopping sequence. This may be done in individually managed transceivers through combination with AFA, in this case the Detect-And-Avoid (DAA) variety. For a centrally managed (multiple) FHSS system, this supervision could be carried out by alternative solutions implemented through a central coordinated entity (i.e. realising Type III or IV mitigation mechanism as described in section 2.10).

In such combined FHSS/DAA configuration, the Hybrid FHSS systems would perform better than static AFA/DFS type systems in periods of congestion. Because in this case the FHSS may avoid single/few congested channels by locating, using, and completing their exchange on channels that would otherwise be idle. Whereas the traditional (static) AFA/DFS might still continue working in the congested channels (for some time) while using remaining available albeit diminishing time slots. Hence 'FHSS system locating free channels' in this sense means keeping data on those channels that have historically (over a period of

seconds to minutes) been successful according to 'Layer 2' or 'Datagram Success' statistics[4]. This is a more successful technique than simply sensing for energy in the band because:

- The technique takes into account conditions at both the receiver and the transmitter
- The technique is sensitive to other mechanisms for the channels being unsuitable, such as destructive multipath

3.5.1.3 Summary of FHSS

Generic FHSS systems are basically operating as a number of parallel DC limited channels. For a single channel DC based system they can be treated as such. The difference between a generic FHSS system and a number of independently operating DC transmitters is that they are correlated in the spatial domain. They simply transmit from the same location.

The Hybrid FHSS may provide better sharing opportunities and, most importantly, the possibility of graceful degradation of operational performance proportional to increasing band congestion. As with any sensing system there is still the possibility of hidden node and exposed node problems and care should be taken in the design to minimise these effects.

For instance, if the generic FHSS were to share with e.g. LBT+AFA/DFS system there might be severe consequences since generic FHSS system does not perform an LBT/DAA operation on each individual channel. As another example, in Hybrid FHSS/DAA system a procedure to detect a constant carrier in a dynamic environment with multiple DC based systems may not be sufficient.

In the same way the number of channels on which any generic or hybrid FHSS system operates needs to be large enough to satisfy the DC rules in each channel. In other words FHSS systems need space to operate, especially in mixed SRD environments.

3.5.2 Direct Sequence

In DSSS a nominally narrow band signal is converted to a wideband signal by multiplying it with a fast pseudo random code. The same code is used in the receiver to recover the signal. It is possible to make the wideband signal look like white noise

A key factor in DSSS is the processing gain, which depends on the ratio of the available bandwidth to the equivalent narrow band signal. In certain circumstances DSSS is a useful sharing technique as it allows multiple signals to be overlaid in the same spectrum. This can be done when either the available bandwidth is sufficient to allow a high processing gain, or the signals are all the same amplitude at the receiver. An example of the latter is the GPS system.

DSSS is in use in the 2.4 GHz SRD band and in the North American 902-928 MHz band. It is permitted in the 863-870 MHz band, but the maximum width of 7 MHz means that only a low processing gain is achievable and duty cycle limits are required in addition [4].

In the 863-870 MHz band, DSSS is best analysed as a modulation method rather than as a sharing technique.

3.6 FREQUENCY AGILITY

Adaptive Frequency Agility (AFA) includes two derivatives. One is known as Detect and Avoid (DAA), which works by a radio transceiver sensing the environment and avoiding permanently those channels that are deemed occupied by other systems transmitting in the proximity. A second agility method is known as Dynamic Frequency Selection (DFS), which is constantly scanning the frequencies in order to avoid occupied channel(s) temporarily or to change to another frequency temporarily.

In other words, it could be said that DFS is realising the dynamic blacklisting of channels whereas DAA implements static blacklisting.

[4] Checking at Layer 2 or datagram success analysis is Collision Detection (after the event) as opposed to Carrier Sensing (before the event). There is therefore an underlying rate of collisions and limits to the throughput achievable.

When the band becomes congested, it may be observed that AFA is well placed to offer a desired solution of gradual degradation of service. Because as the number of users in reception range grows, the AFA would gradually blacklist more and more channels and thus reduce the operational freedom of the respective AFA-managed transceiver. So in that sense the AFA mechanism responds to the environment to create interference mitigation. However, it might suffer the same sensing shortcomings as the LBT method. So here again the ultimate efficacy of interference mitigation would depend on the particulars of the particular interference scenario, such as the number of different interacting applications, the physical scale of the EM coupling area, system dynamics, etc.

3.7 LBT+AFA

In EN 300 220 v2.2.1 [11] is a scheme for Listen Before Talk + Adaptive Frequency Agility for use in the 863-870 MHz band. A specific set of requirements for LBT are combined with rules for how long a device can transmit without changing frequency. In return for compliance with this scheme, devices were exempted from complying with duty cycle limits.

In the most recent revision of EN 300 220 (v2.3.1) [11] a duty cycle limit was imposed on LBT+AFA devices. This was to reduce potential conflict between LBT+AFA devices and low duty cycle devices sharing the same frequency. It was found that, without this new duty cycle limit, certain devices could exploit the LBT+AFA rules to achieve over 80% duty cycle on a single channel.

3.8 DIVISION BY APPLICATION

Dividing spectrum access according to application is extremely common, both outside and within the SRD bands. Outside the SRD bands we find, for instance, bands set aside for broadcasting, for mobile phones, for business use, etc. Within the SRD world, there are whole bands set aside for specific applications (e.g., 169 MHz) and also sub bands within bands (e.g., the alarm sub bands within 863-870 MHz). There are also cases where applications share frequencies but the access rules vary between them (e.g., cordless audio devices in 863-865 MHz are allowed higher duty cycle than generic devices.)

The trend, and the expressed preference, is away from such division, and to apply the principle of Application Neutrality. This is discussed above in section 2.7. It is expected that no further application dependent access regulations will be made in the SRD bands, and there is pressure to remove some of the existing ones.

3.9 CHANNELISATION

Channelization helps to eliminate destructive interference. In Figure below, two signals occupy adjacent channels and do not interfere.

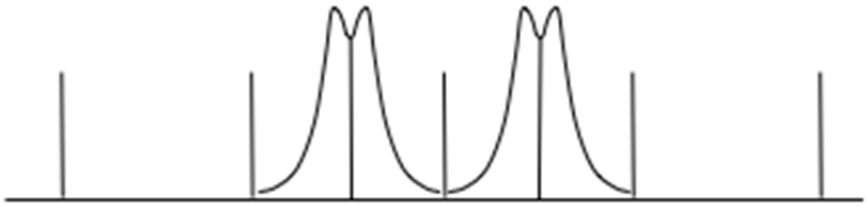


Figure 29: Uncontended channels

In the following Figure, we have the worst-case scenario where a single, weaker, signal disrupts *both* of the previously successful channels. Its location *between* channels puts sufficient interfering energy to disrupt both channels (red).

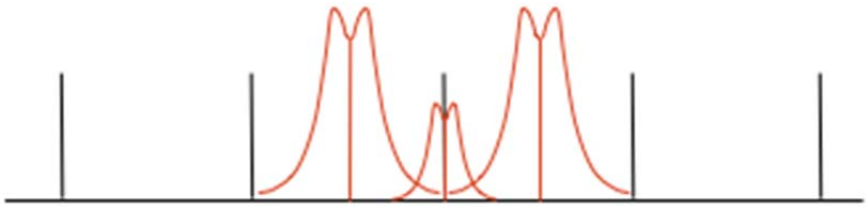


Figure 30: Worst case interferer between two channels

Properly channelized, the interfering signal still causes a failure with one of the signals, *but only one of the signals*, as illustrated below.

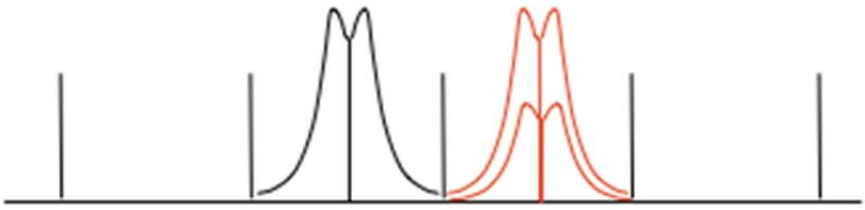


Figure 31: Best case interferer - single channel

Conceptually similar to *Slotted* Aloha in the time domain, channelization in the frequency domain causes interference to fall by half. We have however to accept the consequences of a reduction of technology neutrality may be beyond an acceptable level. Devices that do not need the bandwidth of the channel are either using too much frequency space or are prohibited to use the band to maintain a good GSE.

3.10 MIXED DEPLOYMENT SCALES

Cellular technology has driven much of the spectral efficiency gains in recent years. This is achieved by altering the size of cells in response to local traffic densities: macro cells with radii of many kilometres are used in rural areas to cost-effectively serve large areas; micro and pico cells with much smaller radii, operating at reduced power, provide underlying capacity enhancements, often increasing the capacity per unit area (and thereby spectrum efficiency) by orders of magnitude.

The central planning necessary to achieve this increase in capacity is inappropriate for SRDs, but the principle of achieving capacity increases by combining long- and short-range deployment scales is achievable, even on the same channel.

The success of this mechanism depends on the typical usage scenario's and user expectation for the applications vying for coexistence.

This is illustrated by the 'airport scenario' (which could be applied to many public spaces) below. A Building-area Wireless Data System' system such as 802.11 [13] WiFi system) 'floods' the terminal space with energy in the shared spectrum from multiple Access Points (APs). This raises the noise floor to (say) -80dBm throughout the terminal space. Data System users have a receiver BW matching the transmitted BW of 20 MHz and thus operate well within the area of coverage.

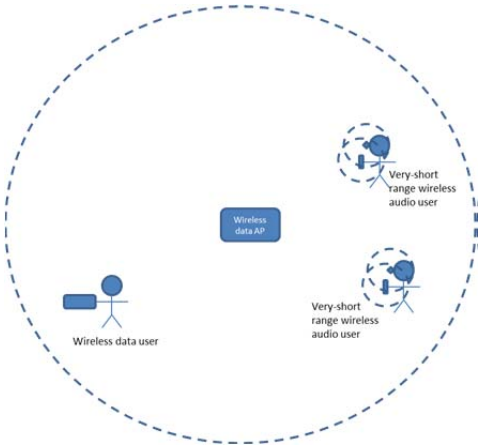


Figure 32: Airport" scenario of co-existence between different range systems

Transiting within the terminal are numerous wireless headset users using a Very-Short-range wireless audio system (such as Bluetooth). Each has a low-powered headset/handset combination that is streaming audio information. The interference radius of the wireless headset for 'headset use A' is small. It is a low powered device that is not particularly optimized for range; it only must reach the handset that is likely carried on or very near the user.

This well-known example shows how two systems with different operational range – Wireless Data and Very-Short Range wireless audio– are able to coexist successfully. The sheer proximity of the headset to the handset provides signals with sufficient SNR to overcome the signal from longer range Wireless Data system. The very short range of this requires very low transmit power, and as a result it does not interfere with other users.

4 ADVANCED TECHNIQUES – SCENARIOS & DISCUSSION OF POSSIBLE DEVELOPMENTS

This section is a discussion of new techniques that could be considered in order to improve further the prospects of co-existence of SRD technologies beyond what could be achievable by traditional spectrum access methods described in previous sections. For each new considered technique, it is aimed to have a brief description as well as analysis of what effect it would have on co-existence. Another interesting question to address it what happens if the considered new technique is used in conjunction with other techniques – the traditional ones or other new techniques. On the cautious side, each new technique would probably have some difficulties/costs associated, thus reducing incentives for industry to introduce it. Another interesting aspect is difficulty/ease for regulators to allow/mandate such new technique.

As a general objective it could be stated that overly restrictive and/or difficult to enforce technical requirements should not be introduced in order to ensure transparency and soundness of regulatory decisions and also not to distort the market development of SRD applications.

Thus a comprehensive cost-benefit analysis may be valuable to complement technical considerations whenever deciding on feasibility of introducing the respective new technique into practice.

4.1 SYNCHRONISATION

Many communication systems face synchronisation issues. For instance, frequency agile systems need to get the transmitter and receiver on same channel at the same time. Many single frequency systems also have time synchronisation requirements.

The problem is particularly acute where the communications traffic is low duty cycle. The system must either maintain sync during the quiet period, or acquire it again each time a packet is to be sent. Acquiring sync is not necessarily a trivial operation and can have considerable overhead penalties in terms of battery consumption or airtime.

4.1.1 Time synchronized systems

On a large number of applications, devices are battery operated, which makes the very economical consumption of energy a paramount requirement to their data transmission.

Synchronisation is particularly useful in those cases where a large number of devices are assigned to a central unit and both devices and central units are battery powered.

In the case of unidirectional communication the intervals of the transmissions can follow a deterministic scheme so that they can be predicted by the receiver. The timestamp of the next transmission from each transmitter in the system can be calculated using the intervals of the past transmissions. After a period of continuous receiving, during which the receiver “learns” the scheme of the transmissions of each device, the receiver needs to be active only for a very short period (typically a few milliseconds) to receive a single transmissions.

This kind of prediction is in particular implemented in metering devices and is also included in the CEN prEN13757-4 [14] standard for wireless meter reading (aka wireless m-bus).

When there is a bidirectional communication between the devices and the central units, each device is assigned to a time window and, if applicable, a frequency. Therefore transmissions from the central unit to the devices are possible without a significant impact on the battery life.

Particularly with regard to the bidirectional communication between the nodes of a network, e.g. data concentrators, a mutually synchronized scheme of time frames and time slots is widely used. The communication from a certain node to another and the direction of communication in respect to the network topology are each assigned to dedicated time slots. The transceiver of the network node needs to be powered only for a very short period (typically a few milliseconds). For the utmost share of its operating time the network node is without communication thus preserving the battery in a maximum efficient way. This is essential when the network node has to be operated from a single battery over many years (e.g. >10 years for metering applications).

The variation of the crystal frequency with ambient conditions can be taken into account for the prediction of the next instant of transmission provided that re-synchronisation is done continuously and that the synchronisation interval has been chosen appropriately.

A further benefit of synchronisation in time is the inherent prevention of intra-system-collisions of synchronized bi-directional systems.

In a quiet environment, this technique works very well and often no additional access mechanism is needed. If a re-synchronisation is considered necessary the reception window of the respective receiver can be widened to an adequate extent.

4.1.2 Acquiring Sync - Calling Channel

The time spent establishing the link is important not only for battery operated equipment but also for systems where the message latency is important.

There is also another reason for minimising the time spent setting up a link, and that is that it inevitably involves transmitter airtime. While airtime under LBT rules is possibly more “friendly” than random airtime, it is still airtime and adds to congestion.

Synchronisation between both ends of a frequency agile (FA or AFA) link is necessary. It is common to think in terms of the transmitter finding a clear time slot and frequency to send a message, but it is also necessary that the receiver is on the same frequency at the same time.

Note: Most links requiring synchronisation will be duplex, so the terms transmitter and receiver are sometimes used loosely. It may be better to think in terms of the originator and the recipient of the message.

There is a number of ways in which synchronisation in an AFA environment may be acquired and/or maintained:

- One method is to use regular transmissions (beacons) to synchronise the system. Another is to engage in a search pattern of calls (probes) and replies each time synchronisation is needed.
- For high data content systems these may be low overheads, but for systems with intermittent traffic they can multiply the airtime many times over. They also slow down the sending of a message. Therefore they result in both lower quality of service to the user, and higher apparent spectral occupation to everyone else.
- A third method of synchronisation is the use of a Calling Channel. A Calling Channel is a sub-band that the various equipments communicate on purely for the purpose of establishing contact; they then move to another frequency to pass the actual data.

The requirement of a Calling Channel is that it has very high availability at all times. I.e that the traffic on it is always of short duration and the aggregate duty cycle is low. The Calling Channel itself may not appear to be an efficient use of spectrum and time, but the trade-off is that the spectrum it serves is used more efficiently.

It is not appropriate or feasible that Calling Channels as such should be specified in the SRD bands. It would be too prescriptive to require their use for operating AFA, and it would be impossible to police whether they were being used only for that purpose.

If, however, it increased efficiency and/or improved quality of service for users, there is a case for a sub-band that could optionally be used as a Calling Channel.

This could be achieved by specifying application-neutral parameters, for instance timing parameters such as:

- Maximum Tx on time, e.g. 50 ms
- Maximum duty cycle over 1 hour e.g. 0.1%
- Maximum duty cycle over 1 sec e.g. 10%

The principle of application neutrality requires that any device that met the technical parameters would have access to that sub-band.

Typical parameters for either a short burst message (or a single hop in a hopping system) in 100 kHz channels might be:

- 5 to 20 ms per hop/burst
- data rate, 20 to 100 kbps,
- i.e. 100 to 2000 bits per hop/burst.

In many cases the payload (the actual information content) of a message is a few bytes or less. It is inevitable that some systems will use the “calling channel” to transfer the actual message. As long as the individual duty cycle limits are met and the aggregate duty cycle remains low, this should not be seen as a problem. In fact it could be seen as an advantage, since it minimises airtime for that traffic.

A good way to handle a very short message would be for the originator to send the entire message in the “calling channel”, together with information as to which other channel he would be listening on for the recipient to acknowledge and to send any reply traffic. Synchronisation is then achieved almost instantly and with very little overhead.

There is the question of whether the “calling channel” should require LBT in addition to the timing parameters or be open to any use that met the timing parameters. This requires further analysis. Allowing non LBT use increases the spectrum access for simplex systems, but at the expense of decreasing access for others. The balance is not clear cut; for instance, with very low duty cycles and very short messages the effect of LBT may not be large. Possibly the most important factor in this question is how it would affect user density.

Creating a very low duty cycle (VLDC) sub-band means a reduction in the rest of the band. But it does remove the need for beacon transmissions and will reduce the amount of probe transmissions.

4.2 VERY LOW DUTY CYCLE / LOW DUTY CYCLE

When employing very simple Duty Cycle and Aloha channel access techniques, transmissions are made at an uncontrolled time, without any checking or reference to other users and the sender hopes for the best.

The probabilities of success are analysed in section 3. It is shown that useful results are obtained if the overall occupancy of the channel is low, e.g. below 2.5% gives a success probability of 95%.

This of course is part of the rationale for duty cycle limits in certain sub bands. It is, however, recognised that the current duty cycle limits are relatively poorly defined (section 5.4 below). There is considerable interest in improving them and ETSI has formed STF411 with the task of developing a definition of Low Duty Cycle (LDC) as a passive mitigation technique.

The expectation is that there will be a set of LDC and/or VLDC rules that would provide certain minimum probabilities of transmission success.

In that case, it would be reasonable to describe LDC and VLDC as passive spectrum access techniques.

VLDC is an attempt to satisfy the reliability requirements of applications by reserving spectral capacity to ensure that at all times the traffic on any one channel is low. This is achieved by setting very low limits on transmissions that individual devices can make, and potentially restricting access to the band to special classes of devices. In terms of Single System Absolute Efficiency this would score badly, but it should be balanced against the utility provided, as measured by Group Spectrum Efficiency.

4.3 ULTRA LOW POWER

There may be an underlay regulation in the SRD bands possible, where in addition to the DC or/and LBT access rules a very low Tx power is specified. This very low power limit should be able to avoid (without any additional mitigation) interference to neighbouring devices using the other access techniques DC and LBT. But it has to be noted that this very low power application could not be protected by LBT (see 3.2.1 and 4.10), while the DC would retain its overall positive impact towards reducing interference potential.

4.4 LBT WITH ADAPTIVE THRESHOLD AND POWER

The effectiveness of LBT is analysed in section 3.3, with the assumption that the Transceiver decides on availability of a given channel based on fixed value of sensing threshold; i.e. if the threshold is exceeded, then the channel is assumed to be occupied, otherwise it is assumed to be available. The theoretically required threshold value for a well working LBT system is for a balanced power situation in the order of the noise floor of the LBT receiver. For higher thresholds the effectiveness of LBT decreases.

Now it is imaginable that the LBT device has several threshold values with corresponding Tx values. The following figure gives an example about the correlations between LBT threshold and Tx power of interferer (IT power) and victim (WT).

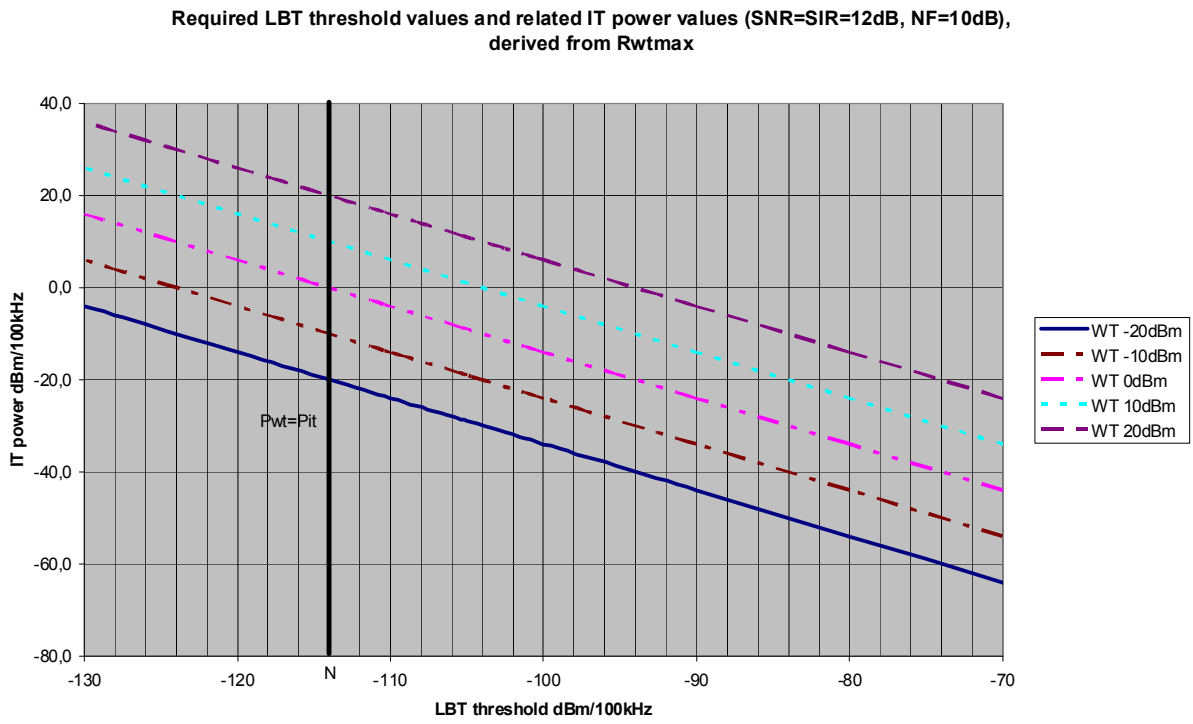


Figure 33: Adaptive LBT sensing threshold

The difficulty here is the same as for LBT in general that this may work in a balanced power situation, because the threshold value is normally designed for the system to be detected with the same Tx power. It therefore may be concluded that introducing an adaptive LBT-sensing threshold, while adding to the complexity of the system, does not offer complete solution to all the identified shortcomings of LBT.

4.5 MESH SYSTEMS

Mesh networks are networks in which the devices not only transmit and receive data for their own purpose but are also used to relay data from other devices in the same system. Mesh networks can therefore enhance their coverage area beyond the reach of an individual device and also provide more reliable communications.

The low setup cost and easy deployment are also important factors. A mesh system can also have reduced susceptibility to multipath effects and can employ self-healing techniques to keep the network intact.

Mesh networks are networks where any node can connect to any other node. This allows the situation where network nodes/devices not only transmit and receive data for their own purpose but are also used to relay data from other devices in the same system. Mesh networks can therefore enhance their coverage area beyond the reach of an individual device and also provide more reliable communications.

The pros of mesh network concept are potentially reduced setup costs due to absence of costly high-mounted access points (base stations) and easy deployment. A mesh system can also have reduced susceptibility to multipath effects and can employ self-healing techniques to keep the network intact. The cons of mesh network concept are the complexity of routing mechanisms and the inevitably incurred delays and data throughput losses as the number of relay nodes increases. These drawbacks had so far held back the wider deployment of mesh networks unless for systems where delays aren't critical. The goal of a mesh network may be the transmission of a single message to a central node but the network may also function without a central node to just enhance the functionality of one or more individual devices. This is for example the case with alarm sensors functioning as an individual self-contained unit but enhancing their coverage by transmitting an alarm or status message through a chain of other individual alarm sensors. The result is in this case a better chance that an alarm status is noted by the user.

Mesh networks may be constructed in different ways. The simplest form is when a device just rebroadcasts the data from an adjacent device or node. This technique is called flooding and involves a predetermined timing and protection mechanism to avoid self-congestion of the network. The message is not following a route but effectively tries all possible routes and has a high redundancy but relatively low spectrum efficiency.

Another technique is routing where data is transmitted from a start point to a defined endpoint.

Each node/device contains part of the information of the route within the network. A message can therefore propagate through the network in a more efficient way than with flooding.

In order to operate effectively, nodes on the network are required to communicate with one another periodically to ensure that neighbouring devices are still in operation. This typically accounts for less than 10% of the higher level application payload. Details of such a system are published in ETSI TR 102 886 [15].

Flooding mesh systems are always less efficient than routing systems in terms of medium utilisation, but may be effective in low traffic environments at getting a message through rapidly.

Mesh systems comprise an underlying set of transmitter/receiver pairs, and when considering the spectral efficiency of the 'system', the technology and techniques used should be the main consideration. In that they do not differ from any other applications in the band.

When considered as a system, however, their widespread roll out on other users in the band might be considered. Factors to note are:

- Mesh systems ubiquity will impact on most other systems in the band, adding traffic across a wide geographic area. Nevertheless, the overall traffic levels generated for typical applications such as some varieties of smart metering may be moderate.
- Mesh networks offer a versatile and effective way of connecting multiple points in an ad hoc way without the need to plan individual point-to-point links, and provide excellent reliability by offering alternative routes.
- Mesh networks allow communications over large distances without the need for high power radios coupled with directional antennas. In crowded areas, where nodes are closer together, the transmission power of the individual links can be reduced.
- Network management traffic will be generated over the air, which is in addition to higher level application data. Whether this should be considered as 'inefficient data' is a moot point, as it is a necessary overhead to deliver the advantages described above.
- The spectrum efficiency depends on the environment in which the network is rolled out. In a static environment where the number of devices for a particular geographical area is fixed (as in Smart Meter networks) and the location is static the network learns the most effective route for a message and may minimise the network management overhead mentioned in the paragraph above.
- This technology was found generally unsuitable for networks with mobile devices[5].

There is a question whether a mesh system is a network /system or a population of independent devices with added functionality. This question cannot be answered in a simple way because it depends on the property of a single device to be able to perform its function alone even if no other devices are present. If this is the case we have single devices with added network functionality, if not and if a device needs to transmit its data to another point as its primary function we speak about a network. Several intermediate conditions exist.

The current view is that mesh devices are authorised as individual devices. Rules on duty cycle, etc, apply to the total transmissions from each device, rather than to the network as a whole or to paths within the network. In the case of devices that can only operate as part of a network and not independently, there is a potential difficulty in defining and/or testing their behaviour and thus applying a technical standard and placing the product on the market.

Typically 10% of devices in mesh networks fall under the category 'network/master devices' such as access points (AP) and relays. APs act in a role equivalent to base stations in 'star' networks, and are required for all systems where the radio component is required to link to a backhaul system (e.g. Alarm systems, Smart

[5] Ofcom UK report - http://stakeholders.ofcom.org.uk/market-data-research/technology-research/research/emerging-tech/dsa/

Meters, Automotive). Relays exist purely to enhanced the coverage and performance of the network and guarantee reliable ubiquitous communications. These nodes will carry the highest traffic and therefore it is imperative that they adhere to access rules in the same way as all other nodes.

It may be therefore concluded not to treat “mesh networking” as a subject of radio spectrum access regulations on a par with the spectrum access/mitigation techniques but to assume, mesh networking is an additional factor outside the scope of the regulations, just helping to improve operational resilience of the subject network of devices.

4.6 ADAPTIVE POWER CONTROL

Adaptive Power Control (APC) is a mechanism that may be employed in bi-directional radiocommunication systems that use the feedback from receiver (received power level indication, or link quality measures such as BER, etc.) to reduce the power of transmitter to the minimum sufficient for reliable operation of the subject link. Thus, by keeping the radio emissions level to the minimum, this method allows to both conserve energy (battery capacity) in the transmitting device as well as reducing unnecessary noise/interference generation within inter-system as well as intra-system interference scenarios. Given its universally applicable value of both saving the energy and reducing interference potential, today the APC is considered a useful design feature in many bi-directionally operated radiocommunication systems.

However, in SRD environments implementation of APC may become inhibited by considerations of additional design and RF implementation complexity in size/cost sensitive miniature devices. Therefore implementation of APC in SRD applications may be normally required and when co-existence studies show that without this feature the interference potential would be unacceptably high.

Also operation of devices in shared frequency environments (as opposed to dedicated centrally-managed single-system environments) could pose additional consideration on the implementation of APC, such as design of its power control algorithms. Example of such consideration was introduced in section 3.3.9 above that discussed the impact of Hidden Nodes in LBT-managed SRD spectrum access environments.

It was shown in section 3.3.9 that when an APC-enabled device is sharing the channel with an LBT-enabled device the variations of transmit power might distort the power balancing and subsequently affect the configuration of hidden/exposed zones around the victim/interfering system.

4.7 ACHIEVING HIGH RELIABILITY THROUGH MULTIPLE TRANSMISSIONS

Reliability of transmissions can be achieved for devices that have no feed-back channel by transmitting data packets/commands on two or three occasions or on multiple channels. If the probability of interfered transmission/lost packet is 5%, then repeating the transmission twice, with sufficient delay for each to be considered statistically independent (i.e. individual interfering episodes lasting less time than the re-transmit times), then the probability of blocking of all three transmissions is reduced to 0.125%, or of success to 99.9%.

Two major problems exist for this technique:

- The medium utilisation is a multiple of that for a single transmission, so the Absolute Spectrum Efficiency is reduced. In the case of short transmissions, however, the total medium utilisation may still be low.
- Success probabilities cannot be guaranteed because the reasons for failures can have longer time periods associated with them e.g. the receiver’s battery could be flat, a nearby transmitter could be ‘stuck on’ or the radio path to the receiving device could be blocked e.g. by a metal partition.

By the same arguments, however, any overhead or forward error correction could be criticised as inefficient medium utilisation. Techniques for achieving high reliability will inevitably lower the absolute spectrum efficiency, but their value should be judged in terms of the utility achieved, i.e. in terms of the Group Spectrum Efficiency.

On the other hand, high reliability services can be achieved at better absolute spectrum efficiency through the use of some form of acknowledgement protocol, especially in a shared band, but this means employing bi-directional transceivers rather than transmitters.

4.8 ADAPTIVE MODULATION

A system capable of transmitting and receiving variable data rates by utilising alternative modulation schemes is able, at a fixed transmit power, to maximise the rate at which data is conveyed between two end points. In this way, fixed packets of data are able to be transmitted at the shortest time commensurate with the range of the link. Therefore, when there is excess SNR on a link (e.g. the Eb/No is well in excess of that required for reception at a given speed) the data rate may be safely increased. Spectral efficiency benefits accrue in both directions: not only is the shorter transmission less of an interferer to others, it is less of a target itself.

4.9 FURTHER REGULATORY PROVISIONS DISCUSSION

The previous sections have provided a thorough overview of different principles and mechanisms for allowing shared use of SRD bands and thus improving the efficiency of spectrum usage. From the analysis provided, it may be concluded that there may be found no “silver bullet”-type mitigation technique that would, on its own, fit any conceivable operational scenario and still be application and technology neutral. As was shown, the neutrality principles sometimes conflict with spectrum use efficiency and therefore avoiding such conflicts requires careful consideration of all circumstances in order to find some balanced approach to regulating SRD access to a specific band.

In that sense, the provided discussions as well as some analytical/numerical analysis also show how it could be possible to draw suitable conclusions for some given set of operational requirements and band/application specific situations. For instance, based on consideration of channel access dynamics in Duty Cycle based access method (see section 3.1), the below table provides a rough example of how the overall channel occupancy may impact the efficiency of the prime method and how other access methods could be called for in case of obvious deficiency of primary method.

Table 5: Access techniques versus aggregate channel occupancy levels

Aggregate channel occupancy	Examples of possible/Suitable techniques	Comment
Below 2.5%	DC and LDC	Suitable for LDC use e.g.Alarms
2.5 to 8%	DC	
8 to 13%	Collision detection needed	
13 to 30%	Impossible without carrier sensing and other techniques	LBT not mandated, but manufacturers would find it necessary
30 to 50%	Advanced networking (access etiquette) techniques	Polling systems
Over 50%	Advanced networking (access etiquette) techniques, or centrally managed. Alternatively accept that a single user dominate the channel	AFA, DFS, DAA can achieve this if number of users <= number of channels.

Additional limitations and requirements for the use of the different techniques are further described in the previous sections and are not fully reflected in the table above (in particular, see section 3.3.9 for the hidden nodes and the exposed nodes for further details).

5 EXISTING SITUATION

5.1 DETERMINING LEVEL OF CONGESTION

In 2008 a monitoring campaign was performed within the framework of WGFM. The basic idea behind the campaign was to investigate the actual use of the frequency band 863-870 MHz but also to set a reference for future monitoring campaigns with the goal to detect a change in actual use over the years.

It was agreed that the overall analysis of the monitoring data from each administration would be performed by that individual administration and as much as possible normalized into a common occupancy figure.

Monitoring took place on the basis of monitoring plan that involved monitoring mainly at previously identified hotspots with actual expected traffic. These hotspots were selected by industry as well as the administrations involved in FM 22. The measurement plan was later published in [3].

Although a pure technical analysis is fairly straightforward a detailed analysis involves specific knowledge of the applications being present in the frequency band that was monitored. This was the main objection from some manufacturers / users against the hard occupancy figures presented in the final report of the monitoring campaign [16].

The occupancy was based on the frequency bins used by the analyser and not necessary on the channel occupancy of the devices. Also because of the timing of the SRDs in combination with the acquisition (sweep) time of the used analysers some transmissions could be lost giving an incorrect occupancy figure specially in those cases where devices with low duty cycle were measured.

The above referenced FM22 report covered this problem partly by presenting the occupancy figures on an hourly basis with a normalized acquisition time. Monitoring engineers argued that transmissions occurring for example only once an hour may still be “missed” in the registrations but that with such low occupancy figures the frequencies at least could be considered uncongested at that particular monitoring location.

Some conclusions from the campaign were that certain frequency segments had no or only very low occupation.

Manufacturers / users argued that this was not automatically a sign of under use a particular application could be designed in such a way that occupation only occurs in an emergency situation or that a low occupation is necessary for proper functioning of the application. Although true it requires a more detailed investigation to determine if sharing is possible anyway in these bands. Another conclusion was that certain frequency segments have strong location dependent occupancy figures that may indicate sharing possibilities.

Another important observation is that there is a distinct division in occupation in the segments 863-865 MHz, 865-868 MHz and 868-870 MHz this suggests that the allocation for non-specific SRDs through the entire range 863-870 MHz is not used as much as expected.

5.2 TYPICAL BANDWIDTHS

The most common bandwidths in use in the 863-868 MHz SRD band are in the range of 150 to 300 kHz. This fits the design of the readily available chipsets and also suits the crystal oscillators available at low cost.

There is a certain amount of equipment with bandwidths of approximately 16 kHz. This is either because they are required to use 25 kHz channel spacing or to gain the advantage of lower noise bandwidth and longer range. Operating at these bandwidths requires the use of more expensive temperature compensated oscillators (TCXOs) rather than plain crystal oscillators.

In the above cases, the modulation bandwidth of the transmitter and the reception bandwidth of the receiver are matched. There has historically been another type of system used in the SRD bands, using Surface Acoustic Wave (SAW) resonators instead of crystal oscillators as the frequency reference. SAW resonators can be used directly at UHF frequency range, whereas a crystal reference requires multiplication or synthesiser circuitry. The frequency error at UHF, however, can be of the order of several hundred kHz, and this results in systems with receiver bandwidths of 500 kHz or more in order to capture the signal, even

though the data bandwidth may be only a few kHz. However, there are also some modern chipsets which allows for data rates which actually use the whole bandwidth for data transmission.

The sub-bands 868.000 to 868.600 MHz and 868.700 to 869.200 MHz were originally created with such SAW-based systems in mind.

Since then, however, new generations of RF ICs targeting the SRD bands have allowed the use of crystal references with low cost and minimal circuitry, which generally eased the above described difficulties. Nevertheless, the use of SAW resonators may still be justified today in case of targeting very low cost/very low consumed power SRD device applications.

This is an example of conflict between technology neutrality and spectrum efficiency. Regulators will have to consider to what extend it is still necessary to accommodate poor frequency stability or excessively wide receiver bandwidths in the planning of SRD bands.

5.3 RECEIVER PERFORMANCE

Receiver performance plays a key role in frequency use, just as much as transmitter performance. It can be argued that narrower channel bandwidths (either Tx or Rx) will contribute to a lower probability of frequency collisions. Narrower bandwidths may, however, mean lower data rates which increase the probability of collisions in the time domain. Frequency use is usually considered from a transmitter point of view. In environments with a mix of narrowband and wideband signals the receiver bandwidth and other factors also play a role in the group spectrum efficiency figure. The following points are the most important factors that have an impact on the group spectrum efficiency

- IP3 and Blocking

When a receiver operates in an environment with many high level emissions in the proximity to the operating frequency, the receiver may become less sensitive due to its non-linearity. The level of the input signals causing this desensitisation may be expressed in terms of the third order intercept point (IP_3)

The point where an increase in input signal no longer results in an increase in output is the saturation level. At this level a strong signal not on the operating frequency prevents the receiver from responding to any other signals, an effect known as blocking. Blocking can be a problem when lower quality receivers are deployed in an environment with relatively strong signals. The result is that the most cost effective receivers (from the user's point of vie2w) may result in poor coexistence with other higher power systems and therefore have a negative impact on spectrum use and spectrum efficiency.

- Sensitivity and LBT threshold

Sensitivity as affected by strong input power is discussed in the previous paragraph, but the poor intrinsic signal sensitivity of the device is also a factor to take into account. For example, the sensitivity becomes a crucial co-existence factor in case of using LBT mechanism, which relies on receiver to sense the presence of other systems in the channel. LBT sensing threshold levels are usually determined, based on the geographical area the device needs to protect to have as little impact on other devices as possible. If a device has a significantly lower sensitivity than devices of the same kind LBT may not work effectively or give rise to hidden node problem (see section 3.3.9).

- Frequency selectivity

Usually a transmitter has a maximum allowed bandwidth based on a channelling scheme or the maximum amount of spectrum to be used at a particular moment. A receiver has a bandwidth based on the properties of the signal to be received and determined by the filtering employed. If the bandwidth of the receiver is too large it may suffer from emissions outside its receiving channel. In that case the receiver itself causes a spectral inefficiency, which is not always obvious since the problem is usually sought at the transmitter side.

To make the list complete some other parameters are mentioned in the following text. These may have an impact on the overall performance of an SRD but mainly in specific situations. They are also dependent on the actual design of the receiver and the frequency band they operate in.

- Image rejection

In a super-heterodyne receiver the wanted signal f_{in} is mixed with the Local Oscillator (LO) signal f_{LO} to produce a signal (f_{LO} - f_{in}) at f_{IF} the IF stage of the receiver. Another signal at the input at f_{LO} + f_{IF} also produces the same frequency when mixed with the LO.

These signals need to be rejected by the pre-selector or other filter at the front end. Inadequate filtering means the receiver is vulnerable to unwanted signals on the image frequency.

- IF rejection

A strong signal at the IF frequency can break through the early receiver stages directly into the IF stages if the front end filtering is inadequate.

- Reciprocal mixing

Occurs when high level unwanted signals mix with the noise sidebands of the LO, generating products at the wanted frequency.

- Frequency accuracy

Accuracy of the receiving frequency is necessary to make use of narrow filters optimised for the received signal.

We can distinguish between the stability as a function of time and temperature and the uncertainties caused by the production process. Poor frequency accuracy in either the transmitter or the receiver means that the receiver bandwidth has to be made larger than the signal bandwidth.

With receiver performance in mind SRDs can be divided into 3 different types.

- Unidirectional:

This type has a receiver on the listening side of the communications link

- Unidirectional with a listening mode such as LBT:

This type has a receiver on the listening side of the communications link but also a dedicated receiver at the transmitter side of the communications link. This receiver is used only for carrier sensing so does not need all the circuitry required for communication.

- Bidirectional

This type has at least one receiver on each side of the communications link.

Bidirectional systems open up a number of possibilities. For instance the communications may be managed by use of a separate control channel. At the very least acknowledgments are possible, so the sender knows whether a repeat is necessary. A bidirectional system may therefore be able to achieve the required overall performance with lower performance receivers than a unidirectional system.

5.4 DUTY CYCLE AND ACTIVITY FACTOR

To describe the time behaviour of SRDs usually duty cycle (DC) and activity factor (AF) are used as separate parameters in the statistical tools used for compatibility calculations. Activity factor is the proportion of time a transmitter is actually used/switched on, independent of whether the device is actually transmitting. DC is the ratio $TX_{on}/(Tx_{on}+TX_{off})$ during the period of activity. Activity factor and DC are not inter-changeable; the typical DC and activity patterns need to be multiplied within the time period for the calculation.

If DC is described over a relatively long period, as is presently done in most cases for SRD, activity factor and DC are exchangeable if the DC is 100%. A number of applications use this principle for a user defined DC, the DC pattern is defined by the manual activation of the device by the user in a specific time frame.

It turns out that the exact definitions of DC and AF are not so straightforward. The distinction between the two can vary according to different points of views. The following table provides an initial possible overview of the on-going discussion within ETSI STF 411 **Error! Reference source not found.** on this subject.

Table 6: Various distinctions between DC and AF

Table 6: Various distinctions between DC and AF

Device State	Off	Idle	Envelope On	Gaps due to Modulation
Example	Sleep	Rx, Ready state Warm up LBT	Transmitting	ASK, Pulse mod
Engineer's view	Activity Factor		Duty Cycle	
Regulator's view	Activity Factor	Duty Cycle		Ignore
EN Standard	Duty Cycle			Ignore
FCC Part 15.203(e)	Duty Cycle			
FCC Part 15.35			Duty Cycle	

The main difference in view lays in that the approach may be either a legal or a technological one. Those parameters that can be easily controlled in regulations are usually part of the DC figure in the eyes of a regulator. Those that are part of the technical performance and purely device controlled are part of the DC figure in the eyes of an engineer.

Moreover; it is believed that very few devices have any form of active control over their activity factor. For manually triggered devices, the manufacturer works on his own estimation of the likely rate of triggers. There is generally nothing built into such a device to stop a rapid sequence of operations. Automatically triggered devices may or may not have internal control over the activity factor and/or duty cycle.

An example of ambiguity with AF/DC definitions may be seen in the anecdotal reports that some manufacturers take the product of duty cycle and activity factor and quote that as the duty cycle. It is possible therefore that some devices with burst type operation are listed as very low duty cycle when they actually have rather high peak duty cycles. This is according to the regulation but creates an uncertainty.in terms of data loss or interference when mixing applications with different DC timing schemes.

In section 2.6 the suggestion is given to solve this in at a higher level, e.g., the access layers of the OSI model. For example, spread out the data to be transmitted over a longer time period, within the frame of acceptable time delay and use an error correction mechanism to reconstruct the data if part is lost as a result of a collision.

If all users in the group employ this example technique, the data throughput for each device is not optimal because of the data redundancy of the error correction mechanism but the group spectrum efficiency increases due to the lower impact of collisions.

5.5 SPECIAL TREATMENT OF SAFETY RELATED APPLICATIONS AND EXCLUSIVE FREQUENCY SPACE

There is a long standing discussion in the SRD world about whether certain applications merit special treatment, for instance access to exclusive or dedicated spectrum.

On one hand, when a manufacturer claims his application has critical implications, such as protecting property or lives, the response from regulators is often to question whether it should in that case be in the SRD bands at all.

On the other hand, if an SRD is capable of saving a life, why should it not be assisted rather than obstructed?

The philosophical side of this discussion is outside the scope of this report. It is noted, however, that keeping a channel empty for the benefit of one application is not necessarily spectrum efficient.

When a particular application has a requirement for high reliability, it is expected that the equipment should be designed so that it can achieve that reliability. It is then entirely appropriate that the spectrum access rules be framed so as to make that possible, within the overall goal of efficient use of the spectrum.

In other words, what is required is a partnership and a balance, between industry and regulators.

In this regard, it is noted that the guidance from the EC to CEPT on the 5th update of the EC Decision on SRDs [17] contains a strong presumption against application specific sub bands unless duly justified. On the other hand, certain sectors of industry dealing with inherently safety related equipment believe exclusive or dedicated bands are necessary for their applications.

Application specific bands can serve a number of purposes:

- They limit the numbers, or the density of devices using the band.
- They can segregate different types of signals. e.g., high duty cycle and low duty cycle signals are separated, which is necessary for spectrum sharing.
- They can segregate systems according to need. E.g. Systems requiring low latency do not have to compete for spectrum access with other systems.

On the other hand, there are arguments against application specific sub bands:

- They do not guarantee exclusivity. Only other types of SRDs which are under ERC/REC 70-03 [2] can be excluded.
- They may not be enforceable in practice. In the case of alarms, there is no proper definition of an alarm and many types of device could declare themselves as “alarms”. The picture is further muddied by devices that are sometimes alarms and sometimes not, such as image transmitters.
- Segregating signals by type is better done directly by means of signal parameters. High and low duty cycle signals can be separated without reference to application.
- In the case of some applications, the numbers and density argument is weakened if those applications are common.

5.6 EXAMPLES OF CURRENT SRD USE.

This section gives details of some of the applications and devices currently using the SRD bands, with discussions of expected developments and spectrum access requirements.

Further information on transmission timings can be found in Annex 7, which is a spread sheet giving an overview of many applications. The information in this was collated by ETSI groups STF 411 **Error! Reference source not found.** and TG28.

5.6.1 Automotive Industry

The applications for SRDs cover comfort and security as well as safety related features:

- **Remote Keyless Entry (RKE)** allows the driver to manually open the vehicle with a radio controlled key. Although representing a feature that mainly increases convenience, they are more and more considered as an essential requirement by customers. Even more convenient is the so-called:
  - Keyless or Passive Entry, which is based on bi-directional communication between the vehicle and the driver. Proximity sensors built into the door handles detect the approaching driver. A signal is sent from the vehicle to the driver's key, which responds with an identification signal. The same identification signal allows the driver to start the engine by simply pushing a button without putting the key into the ignition lock.
  - Passive Start. Starting the engine without this identification signal is impossible, which enhances the anti-theft protection.
- Additional functions can be integrated into cars as follows:
- **Remote Key**, for opening and closing of windows, sunroof, retractable hardtops etc., doors may automatically lock when the driver exceeds a certain distance from the vehicle. The so-called:
- **Personal Car Communication** allows the driver to receive information at a greater distance from the vehicle telling him e.g. that the lights or radio are left on or that the car is not locked. In addition it may be used to start remotely the Fuel Fired Heater or other car functions. An important aspect is the enhanced safety function of this application e.g. detecting persons left in a car.

- **TPMS (Tyre Pressure Monitoring System)** is considered an integral part of additional measures as required by the European Commission in order to achieve the EU CO2 policy target of 120 g/km.

  Background:

  The EU has set out a strategy to achieve a fleet average target for new vehicles sold in 2012 to 120 g CO_2/km. To achieve this target, the Commission is proposing to require mandatory CO_2 emission targets for manufacturers that equate to average CO_2 emissions of 130 g/km. To compensate for the gap of 10 g/km, additional measures have still to be defined. Current proposals include low rolling resistance tyres, low friction oils, more efficient air-conditioners, TPMS, etc.

  The role of TPMS is to eliminate the statistically proven incidents of low tyre pressure that lead to increased rolling resistance and consequent higher fuel consumption.

  The safety aspects of TPMS should also not be overlooked. An unreliable system may lead either to failure to warn of low tyre pressures or to false alarms.

- **Communication between truck and trailer** may be also provided by SRDs. The transmitted information enables the driver to monitor the status of important points on the trailer, such as brakes, lighting, individual pressure of all tyres etc.

- **Vehicle Alarm Systems**, like "Panic alarm" are systems which initiate alarm functions like the siren and flashing lights.

5.6.1.1 Pertinent Technical Details

Table 7: Envisaged requirements of automotive SRDs

Essential Parameters	
Cost limitations	Extreme cost pressure in particular for disposable device such as TPMS sensors
Technical restrictions	Few
Typical Rx specifications	Low specification
Spectrum access techniques	Duty Cycle
Expected DC	Low

5.6.1.2 Spectrum utilisation and Spectrum Efficiency discussion

Typical Automotive SRDs transmit bursts of 10 msec to ensure a fast response time.

One feature of many of these automotive applications is that they are very short range. Or rather that they only need to be reliable at very short range (some RKE fobs will operate over many metres, but users have become used to the fact that they do not work 100% of the time).

5.6.2 Alarms and Social Alarms

5.6.2.1 Application description

Alarms are devices designed to detect a condition and to communicate an appropriate warning or summons for help. Wireless alarms are found in the SRD bands for intruder detection, fire and smoke detection and other purposes. Social Alarms are a special category for providing assistance, support and monitoring to infirm or vulnerable people. Alarms are generally very low duty cycle.

A comprehensive description of wireless alarms can be found in ETSI TR 103 056 [18].

5.6.2.2 Pertinent Technical Details – Alarms

Table 8: Envisaged requirements of SRD Alarms

Essential Parameters	
Cost limitations	Medium
Technical restrictions	Narrowband
Typical Rx specifications	Usually high specification
Spectrum access techniques	Duty Cycle
Expected DC	Very Low
QOS requirements	Very high reliability

In the case of alarms, various user level organisations have set requirements that affect the technical parameters of the radio equipment.

5.6.2.3 Spectrum utilisation and Spectrum Efficiency discussion

Many alarm systems are uni-directional only. There is no spectrum access technique as such and they rely on the channel occupancy being low. A discussion of the increasing probability of success in these conditions can be found in section 3.1 and 4.7 above. At present several sub-bands in 868-870 MHz are restricted to alarm systems, with restrictions on duty cycle. It can be questioned whether this is a viable arrangement for the long term, because:

There is no guarantee that the aggregate duty cycle or occupancy of these channels will remain low (e.g below 3%). The regulations only limit the duty cycles of individual devices. As installations become larger and more numerous, the total occupancy will rise. This may cause a switch to bi-directional systems which may in turn increase the occupancy.

More than one administration has removed the protection of these sub-bands and allowed general purpose SRDs onto these frequencies.

A number of manufacturers are of the view that there is a need for battery powered devices and therefore they can only use duty cycle access. Others are of the view that battery powered devices can use other mechanisms.

The alarms industry is seeking further spectrum and means of ensuring high reliability communications, see ETSI TR 103 056 [18].

Because the exclusivity of spectrum for (social) alarms is uncertain, it is necessary to devise spectrum sharing and access mechanisms that meet the simultaneous application requirements of high reliability and low latency.

5.6.3 Building Management – Home Automation

5.6.3.1 Application description

This is a potentially wide field and overlaps with alarms, healthcare and metering. Traditionally, heating and lighting control has been seen as the core of home automation but there is now interest in integrated systems that combine a multitude of functions.

Entertainment and data networks may also rely on radio traffic in the home. Currently these are considered separate functions and generally operate on different frequencies (e.g. cordless audio in 863-865 MHz, WLANs in 2.4 GHz, video in 5.8 GHz). It is noted, however, that traffic from many functions such as home automation, entertainment, alarms may merge at the point of entry/exit of the building, e.g. telephone line or cable system.

5.6.3.2 Pertinent Technical Details

Table 9: Envisaged requirements of Home Automation SRD

Essential Parameters	
Cost limitations	Medium
Technical restrictions	Varies with function / Still evolving
Typical Rx specifications	Medium
Spectrum access techniques	Various (e.g. LBT+AFA, DC…)
Expected DC	Medium
QOS requirements	Varies with function. Lighting control requires low latency, heating control does not

5.6.3.3 Spectrum utilisation and Spectrum Efficiency discussion

Because home or building automation covers a range of functions, it is difficult to apply a single set of criteria. For instance, functions interfacing directly with the occupant, such as lighting control, need high reliability and very low latency. Functions concerned with security, intruder detection, fire detection, need very high reliability and medium latency. Other functions, such as heating can be treated more as background tasks.

5.6.4 Meter Reading

5.6.4.1 Application description

Meter reading, commonly referred to as Automatic Meter Reading (AMR), is an established user of radio links in the SRD bands.

The requirements for the metering market are particularly defined by EC directives, which focus on the efficient use of energy to mitigate anthropogenic climate change and to reduce the economic dependency of the EC on the import of primary energy resources. Metering and individual cost allocations are known to reduce significantly the consumption of primary energy resource by stimulating a change in consumers' behaviour. Thus AMR-systems have a high socio-economic benefit.

Descriptions of smart metering systems and the smart grid can be found in ETSI TR 102 886 [15], TR 103 055 [20] and CEN EN 13757-4 [14].

There are different applications to distinguish within the area of meter reading:

- **Metering**
  The measurement of heat, electricity, gas and water for resources provided by utilities. There is usually one electricity meter and one gas meter (if applicable) per flat and one water meter per building.
- **Sub-metering**
  The measurement of resource consumption for energy cost allocation and water cost allocation. With several heat cost allocators per flat – one per radiator – and typically two water meters per flat the number of sub-metering devices inside a building is significantly larger than the number of metering devices.
- **Energy data management**
  Metering data, sub-metering data and additional data are combined for value added functions like consumption analysis, device monitoring etc. or for the control of the central heating.

The majority of metering devices operate from a battery with anticipation of long service time, i.e. the guaranteed lifetime of more than 10 years must be achieved from a single battery. It is impossible to change the battery during the lifetime of the device.

This imposes extensive restrictions on the energy budget of the device in general and on the activity of the radio in particular.

5.6.4.2 Pertinent Technical Details

Table 10: Envisaged requirements of SRDs for Metering Applications

Essential Parameters	
Cost limitations	For plain AMR, particularly for sub metering, very low cost is needed to justify installation. For metering applications the use of smart meters may become compulsory.
Technical restrictions	guaranteed lifetime > 10 years => limited energy budget tamper proof, privacy concerns small outline (particularly sub metering devices)
Typical Rx specifications	Medium
Spectrum access techniques	Duty Cycle
Expected DC	Low to Medium
QOS requirements	Some systems have legal requirements for 15 minute latency

5.6.4.3 Spectrum utilisation and Spectrum Efficiency discussion

AMR is generally a very low traffic system. (The wireless m-bus standard EN 13757-4 [14] for instance refers to a 0.1% duty cycle sub-band (Annex 1 g2 in ERC/REC 70-03 [2], 868.700-869.200 MHz) with the recommendation not to exceed a duty cycle of 0.02%.) There are plans, however, towards Smart Metering and the Smart Grid in which extra functions of energy management are incorporated and the level of traffic in these systems will be higher.

5.7 CHANGES IN THE ENVIRONMENT: TRANSMITTERS IN ADJACENT BANDS

The interaction between SRDs and equipment in adjacent bands needs to be considered. One example is the changes occurring to 790-862 MHz (the “digital dividend”).

Whereas previously it has been used for TV broadcasting, in many European countries it is now being allocated for mobile communications (LTE).

TV broadcasting involves high power transmitters. However they are relatively few , and in known locations and the field strength at ground level within defined limits. They have not generally been known to cause problems to SRDs in the 863-870 MHz band.

This use will be replaced by large numbers of mobile cellular networks. The possibility exists of close physical proximity between SRDs and transmitters in the adjacent band.

This new compatibility scenario needs to be further assessed taking into account that license-exempt SRDs operate on a non-interference non-protected basis, and also taking into account the European commission decision 2006/771/EC and the R&TTE directive 1999/5/EC.

5.7.1 Use of LBT

The use of LBT requires setting a sensing threshold. Above the threshold the channel is considered occupied; below it is considered empty. If the background noise is above the threshold then clearly there is a problem.

5.7.2 Alarms and Low Duty Cycle equipment

Low Duty Cycle as an access and sharing mechanism relies on the fact that all the other users of the channel are also low duty cycle, and that the total occupancy of the channel is low. The analysis in section 3.1.1 applies.

This only works if all the interfering energy is grouped into short bursts of high amplitude with quiet periods between them. Therefore, some sub-bands are controlled by regulation to only allow signals meeting this requirement.

The effect of noise spill-over from adjacent bands is to introduce into these sub-bands interfering energy that is different in that it is omni-present and changes little with time and thus threatens to undermine the LDC-based sharing arrangements.

Alarm systems in particular tend to use very sensitive receivers and can be expected to be vulnerable to an increased noise floor. Indeed the receiver regulations for Category 1 performance in EN 300 220 force them down the route of using very sensitive receivers (receiver parameters are discussed in section 5.3).

It may be therefore concluded that whenever deciding on authorising any such sensitive applications, the care should be taken to safeguard them against the danger of interference spill-over from neighbouring bands.

5.7.3 Battery powered devices

It is common for battery powered devices to employ a wake/sleep cycle. The device wakes periodically, senses a channel for activity and then goes back to sleep if it is quiet. Either a general raising of the noise floor or unintended emissions into the band would cause severe difficulties for such devices.

6 DISCUSSION ON SPECTRUM ACCESS RULES

6.1 MINIMUM COMMON REGULATION

Previously it has been common to treat SRD devices as falling into different spectrum access categories, e.g.:

- Duty cycle limited devices
- LBT
- LBT+AFA
- Frequency Agile
- FHSS
- Etc

and to draft a set of rules for each category.

Now with the stated aims of technology neutrality, the natural question to ask is: rather than doing this is it possible to write one set of rules that applies to all devices?

One possibility is that we do not have rules about changing frequency, number of hops, use of LBT, etc. Devices would use these techniques not because the regulations required them, or rewarded them with extra airtime or spectrum, but because they gave a benefit to the device itself in a congested band. The idea would be to arrange things so that the spectrum access would be governed by a limited set of essential requirements and then it were the manufacturers to decide whether any additional techniques would give sufficient benefits both to the user of the technique and to other users in the band he might interfere with.

In the case of LBT the net potential benefit may occur naturally. Avoidance of a collision benefits both parties. In real life, however, LBT is not able to avoid every collision.

Similarly, in the case of changing frequency, the potential benefit occurs naturally with adaptive frequency agility, as avoidance or minimisation of co-channel operation benefits both parties. In the case of non-adaptive FHSS the situation might be however complicated if the hopping sequence overlapped with some higher occupancy channels, reducing mutual benefits with users of such channels.

Some possibilities of introducing technology neutral spectrum access regulations are presented below. It should be however noted that the examples presented here give just an idea of possible regulations, but they are not validated yet. They are just candidate possibilities, the effect of some of which are assessed in section 6.3.

6.1.1 Control by channel access time

This approach is based on controlling (limiting) the transmissions emitted within a given channel. If necessary to exceed the limit, the device can always get more airtime by using more channels. An attraction of this system is that it is fair also from the viewpoint of a potential victim.

For instance:

In any [200 kHz] bandwidth the transmission is limited to

Max Ton (duration) [100 ms]
Min Toff [100 ms]
Max duty cycle [10%]

Different parameters could be set in different sub bands within the band.

It may be necessary to define the measurement in more detail. E.g. by defining a measuring receiver and specifying an e.r.p. threshold.

6.1.2 Control by total airtime

This approach is based on the idea that the more airtime a device wants, the more mitigation techniques it has to deploy.

For instance:

Up to 0.1% duty cycle [or according to future LDC rules] - no restrictions other than power

Up to 1% duty cycle Max Ton [100 ms] Min Toff [100 ms]

Up to 10% duty cycle Max Ton [100 ms] Min Toff [100 ms]
Use of LBT

Over 10% duty cycle Max Ton [100 ms] Min Toff [10 ms]
Use of LBT
Use of AFA
Restriction to certain sub bands

6.2 PERFORMANCE ASSESSMENT OF SPECTRUM SCHEMES

As discussed in Section 2 above, a method of assessing various options is needed.

The question therefore arises: is it possible to write a set of criteria/constraints that can be used to judge and compare the overall effect of different sets of rules?

E.g., a list of key points or requirements could be stated:

- Preferably be compatible with existing regulations and their evolution.
- Users not constrained unnecessarily in uncongested case.
- Acknowledge that different applications have different criteria
- Stability of regulation – or at least careful planned change. Must also plan for likely technical improvements.
- Application neutrality pursued as s far as possible.

- Technology neutrality pursued as far as consistent with spectrum efficiency and sharing requirements
- Regulation as simple as possible
- Equitable and predictable sharing in congested case (spectrum, space, time, spectral density domains). Graceful degradation rather than catastrophic failure.

The following is a list of example criteria against which the operation of access rules could be judged. A good set of access rules would fulfill as many of these requirements as possible:

1. [75%] prob of getting a blind transmission of [50 ms] through in [1] attempts (maybe in some subbands only)
   e.g. low cost unidirectional system
2. [99.5%] prob of getting message of [100 ms] through within [3 sec] with more sophisticated system
   e.g. Alarm system needing high reliability
3. [98%] prob of getting message of [50 ms] through with latency less than [200 ms]
   e.g. Control system needing low latency
4. For single user, permit traffic of [200 kHz] BW at [80%] duty cycle?
5. Permit [5] co-located users [200 kHz] BW at [80%] duty cycle?
   e.g. spectrum may be used more fully when uncongested.

6.3 ASSESSMENTS OF PROBABILITY

In an attempt to enable the assessment of spectrum access techniques, two spread sheets were developed to calculate the probability of successful transmission in various scenarios. The spread sheets are described in Annexes 3 and 4, and they enable testing of success rates with various access rules and levels of congestion. They can be used with the example numbers above or other suitable ones.

The Annex 4 spread sheet in particular allows testing of criteria 1, 2 and 3. It shows the results of trying various strategies (e.g, multiple transmissions with and without LBT) in various environments subject to various sets of access rules. Band Segmentation

A band segmentation scheme may be implemented in order to implement the ideas presented in section 6.1.

This section provides an example of band segmentation which may need to be further considered and reviewed in order to possibly optimise each sub-band. The criteria listed above could be used as a means of evaluation.

The following supposes a piece of spectrum is divided into 3 sub-bands. Each sub-band is capable of supporting more than one channel of width 200 kHz, for example.

Figure 34: Example of piece of spectrum divided into 3 sub-bands

where:

(A) is intended for low cost unidirectional systems (e.g. some alarms). Criterion 1 discussed above would apply here.

(B) is intended for systems needing high reliability and/or low latency (e.g., more sophisticated alarms, lighting control). Criteria 2 and 3 would apply here.

(C) is intended for systems with high data throughput where traffic conditions allow. Criteria 4 and 5 would apply here.

The separation of data systems (C) from safety related (B) and simple applications (A) is required because the sharing mechanisms for (C) are not able to achieve the requirements of (A) and (B) (see section 3).

(A) requires very low duty cycle and ideally also a limit on transmit duration.

The means of achieving (C) involve networking protocols (e.g. ERC/REC 70-03, Annex 3, band a, "The equipment shall implement an adequate spectrum sharing mechanism in order to facilitate sharing between the various technologies and applications covered by this annex 3"). However, in the SRD bands, only peer to peer interactions are possible (i.e., no central controller). The key to enacting such peer to peer interaction may be some form of carrier sensing. In such case the access conditions for a channel may require LBT or equivalent technique since, these are the only practical methods of carrier sensing in this case. Due to the limitations of LBT, which are analysed in section 3.3, a frequency separation from applications with high reliability and known latency requirements is likely to be necessary.

Note the use of AFA in addition is implied if a device wishes to use more than one channel.

The means of achieving (B) needs further work consideration and work to define the appropriate access conditions.

It is suggested that, with careful choice of the parameters in section 6.2, the needs of all the different users can be balanced. This is discussed further in Annex 6.

The above discussion showed the possible thinking when deciding on the band segmentation approaches in technology neutral manner. It becomes clear that, again, the key to successful solution would involve careful balancing act.

7 CONCLUSIONS

It is important to distinguish between spectrum occupancy and spectrum efficiency. The value of using a particular part of spectrum comes from the utility it provides to users, which is not necessarily the same as the data traffic. A distinction should be made between the concepts of Single system Absolute spectrum Efficiency (SAE), which is based on the raw data transmitted, and Group Spectrum Efficiency (GSE), which is closer to the broader utility or service provided.

One conclusion is that some SRDs operating in "exclusive" bands might indeed benefit if those bands were to be low occupancy so that devices relying on access by duty cycle (DC) limits alone can operate effectively.

At the same time, it would be wasteful and inefficient to operate all the spectrum identified for SRDs in this way. In other sub-bands, whenever there is demand, occupancy and throughput levels will have to rise. Regulators and industry will have to devise means of achieving this. Since basic DC is only effective as a sharing mechanism up to relatively low levels of occupancy and throughput, this may require the introduction of more advanced sharing mechanisms.

A second conclusion is that different sub-bands should be optimised for different communication needs. Users of the SRD bands have a variety of needs and different criteria for a successful service, and this should be recognised in the management of the spectrum identified for SRDs.

Access mechanisms and spectrum management should be based on sound technical foundations – the equivalent of "evidence based" rule making. This report initiated some work relating to the derivation of the technical parameters and spectrum management for a given SRD sub-band. This work should be continued and extended.

The following conclusions on specific points are drawn:

Spectrum occupancy
Spectrum occupancy is the parameter most visible to observers and monitors of the spectrum. Section 3.11 shows the relationship between occupancy levels and access techniques. For monitoring purposes on an application level the distinction between spectrum occupancy and

channel occupancy needs to be made. In most general cases this is not necessary but when investigations are made in specific sub-bands, especially when considering application of new spectrum efficiency metrics proposed in this report and some advanced mitigation mechanism, this may be relevant.

High Reliability Use
There is a need to optimise some of the SRD spectrum to achieve high reliability use. The amount of spectrum required for such usage might be relatively small.

Application neutrality
The aims of this report are entirely consistent with the principle of application neutrality set out in CEPT Report 14 (section 2.7[19]. For instance, it may be better to designate a sub-band not for inherently safety related critical alarm systems, but instead as a sub-band where high reliability, low latency, low duty signalling is always possible. This is a clearer path for regulators to follow in order to provide a better service both to alarm manufacturers and to alarms users while remaining application neutral, thus not preventing further innovation in the given sub-band.

The principle of application neutrality means the end of segregation by application – whereby sub–bands were designated exclusively to a particular application, primarily within the European SRDs generic frequency ranges. In order to preserve technical efficiency, a suitable replacement could be partitioning of the bands based on technical objectives – e.g., sub-bands for high reliability, for low latency, for high throughput. However, this may lead to more detailed definition being needed in describing the technical requirements and this may lead to a reduction in technology neutrality if not performed properly.

At the same time it is worth noting that sometimes an SRD application may have very clear specific technical characteristics that may employ opportunistic sharing techniques to enable it to politely operate within spectrum allocated to radiocommunication services that otherwise would be interfered by generic SRDs. This may represent a higher spectrum use efficiency, beneficial to both uses.

Technology neutrality
The principle of technology neutrality is more difficult to realise and therefore may not always be realised by regulation without sacrificing spectrum use efficiency. It should be still possible to frame regulations so that, for instance, either analogue or digital modulation is allowed or a range of bandwidths is possible. In most cases, however, it is necessary to set specific technical conditions to allow successful sharing, so technology neutrality is at odds with spectrum efficiency. There may be a case for a “sandpit” area, akin to the concept of bands identified for ISM, where technology neutrality is applied as far as possible, to assist the emergence of new technologies.

Listen Before Talk

Listen Before Talk (LBT) is well known mitigation technique in the SRD field whereas the transceiver performs sensing of the channel before each packet transmission. This report carried out an extensive modeling with the aim of quantifying the precise benefits of LBT in various sharing scenarios. It was shown that the LBT is not a “silver-bullet” in that it has its limitations and shortcomings, most notably as described by the “hidden/exposed node” problems.
Moreover, the report considered the benefits of two related concepts, namely those of Carrier Sensing (CS) and Collision Detection (CD), known as part of so called Aloha channel access protocol. CD is the detection of a collision after the event. This happens in all systems that work at the higher levels of the OSI model, such as analysis of message success rates. CS operates before the transmission with the aim of preventing collisions. It thus closely resembles LBT and sensing elements of more advance frequency agility mechanisms such as DAA, DFS and AFA. The notable conclusion of this report is that LBT and CS/CD require further studies in anticipation that some kind of hybrid mechanisms, involving both CD and CS aspects, would be necessary if wanting to achieve high levels of throughput and spectrum use efficiency in high channel occupancy scenarios.

FHSS and Hybrid FHSS
The traditional generic FHSS may be only truly effective in scenarios with lower levels of *band* occupancy; basically it spreads the traffic over a wide spectrum to reduce the per-channel traffic to low levels. Hybrid or adaptive FHSS need further study to see how effectively it overcomes the

limitations of generic FHSS and what other types of spectrum access mechanisms it can most optimally share with.

Noting the nature of FHSS as band-level, not channel-level access mechanism, it may be suggested that regulations should not make special provisions for FHSS, but should instead apply per-channel access rules taking into account the correlation of channel transmissions in the spatial domain.

Advanced technologies (FDMA, CR…)
These spectrum access techniques may have received less attention and analysis to date in the SRD community than time domain techniques, mostly because of the higher involved complexities, including some of them needing central controlling entity with degree of “intelligence” etc., but they should be studied further as potential techniques for high occupancy, high traffic sub bands.

Mixed deployment scales

It may be possible to achieve spectrum use efficiency gains and overall spectrum capacity increase by combining longer- and shorter-range deployment scales (still in the overall context of limited SRD range). This would resemble the principle of combined deployment of umbrella macro-cells and pico-cells in the same area, even on the same channel. Systems operating with differing operating powers within sensible limits and ranges are able to effectively co-exist, thereby significantly increasing medium utilisation. The success of this mechanism depends on the typical usage scenarios and user expectation for the applications vying for co-existence. The achieved group spectrum efficiency depends on the choice of spectrum access parameters.

ANNEX 1: HIDDEN NODE ANALYSIS

A1.1 INTRODUCTION

In this Annex the compatibility of an LBT-like SRD (Interfering transmitter IT, transmitting to its wanted receiver WR) with a non-LBT SRD (Wanted transmitter WT transmitting to the victim receiver VR) is analysed. WR is able to monitor the WT, which is the basis for LBT.

A1.2 HIDDEN NODE THEORY

The following abbreviations and definitions are valid in this annex:

- Dimensions: r/m, P/dBm, S/dBm, SIR/dB, f/GHz, All antennas 0dBi
- VR Victim receiver
- WT Wanted transmitter (victim link)
- IT Interfering Transmitter
- WR Wanted receiver (Interfering Link) :
- N: Noise floor kTBF of VR,
- F: Noise figure of VR,
- S: Signal strength received at the VR from WT (Pwt)
- SNR: signal to noise ratio, or C/N at VR
- I: Interfering power at VR,
- SIRmin: Signal to interference ratio, or C/I at VR
- Pit Transmit power of IT
- Pwt Transmit power of WT
- Plbt: LBT power received at WR from WT (Pwt)
- Pthr: power threshold for the LBT mechanism at IT
- n: Path loss exponent n (n=2 free space loss)
- Rint: radius around VR; inside interference can occur (S-I<SIRmin)
- Rsig: radius around VR; inside the victim link works with S-N<SNR
- Rdet: radius around WT; inside the IT can detect the WT

The following figure explains the investigated scenario. Within a radius of Rint around the VR the IT can exceed the protection objective of the VR (e.g. C/I). Within a radius of Rdet around the WT the IT can detect the WT.

In the light blue area in the following figure 35 LBT is working effectively. The red area is the so called “hidden node”, where the IT is not able to detect the WT.

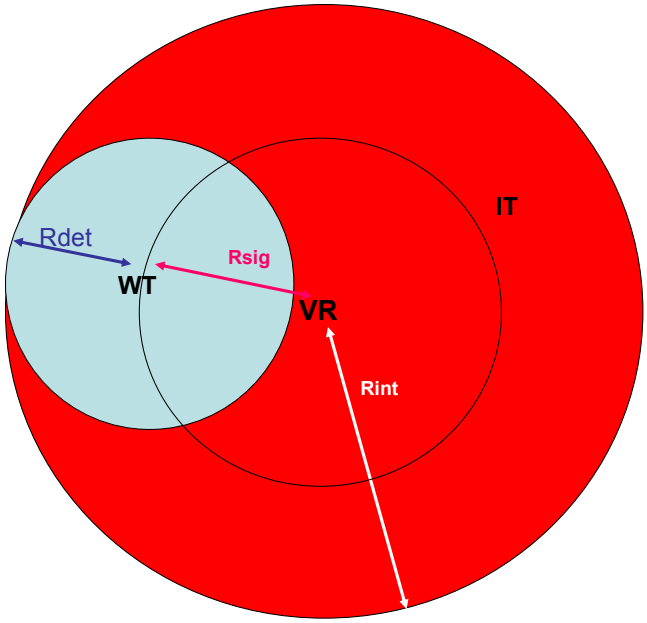


Figure 35: Illustration of the scenario

The formulas given hereafter are the basis for the further analysis.

S	=	N + SNR	=	Pwt	- PL(Rsig)	(1)
I	=	S – SIRmin	=	Pit	- PL(Rint)	(2)
Pthr	=			Pwt	- PL(Rdet)	(3)

Path loss:

PL = 32.5+10*n*log(R/m)+20*log(f/GHz) (4)

Figure 36 illustrates the assumed path loss model.

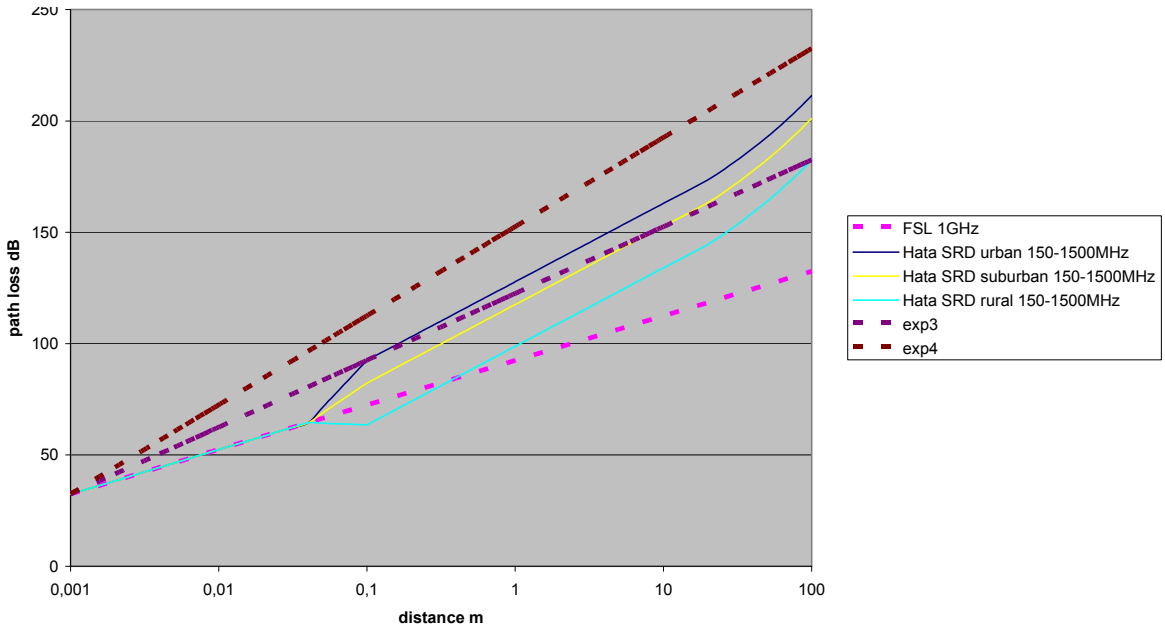


Figure 36: illustration of the used path loss model and comparison with extended hata SRD

The quantity of the circles from Figure 36 can be calculated as follows:

(1)+(4) -> 10n*log(Rsig) = Pwt-N-SNR-32.5-20logf (5)

(2)+(4) -> 10n*log(Rint) = Pit-N-SNR+SIRmin-32.5-20logf (6)

(3)+(4) -> 10n*log(Rdet) = Pwt-Pthr-32.5-20logf (7)

Relation Rint/Rsig: (6)-(5) -> 10n*log(Rint/Rsig)=Pit-Pwt+SIRmin (8)

Relation Rdet/Rint: (7)-(6) -> 10n*log(Rdet/Rint)=Pwt-Pit-Pthr+N+SNR-SIRmin (9)

Under the assumption Rsig+Rdet≤Rint the hidden node portion can be easily calculated as 1-(Rdet/Rint)^2. With (9) the Hidden node portion can be analytically describes as:

HN = 1- 10^[(Pwt-Pit-Pthr+N+SNR-SIRmin)/5n] (10)

A1.3 THE FOLLOWING SECTIONS A1.3 AND A1.4 SHOW THE HIDDEN NODE EFFECTS IN MORE DETAIL FOR A BALANCED AND AN UNBALANCED POWER SITUATION.EXAMPLE FOR A BALANCED SCENARIO

E.g. with Pit=Pwt, Pthr=-87dBm, N=-106 and SIRmin=12dB

HN=1-10^[(SNR-31)/5n] (11)

If the condition Rsig+Rdet ≤Rint is not fulfilled, the situation is more complex. But the following condition can be used to derive the SNR ratio where the hidden nodes are disappearing: Rsig+Rint≤Rdet.

The results for Rsig, Rint and Rdet for the path loss exponents 2 (=LOS), 3 and 4 are given in Figure 37 to 39 below. The other assumed parameters are: BW=200kHz, NF=15dB, Pthr=-87 dBm, SIRmin=12dB.

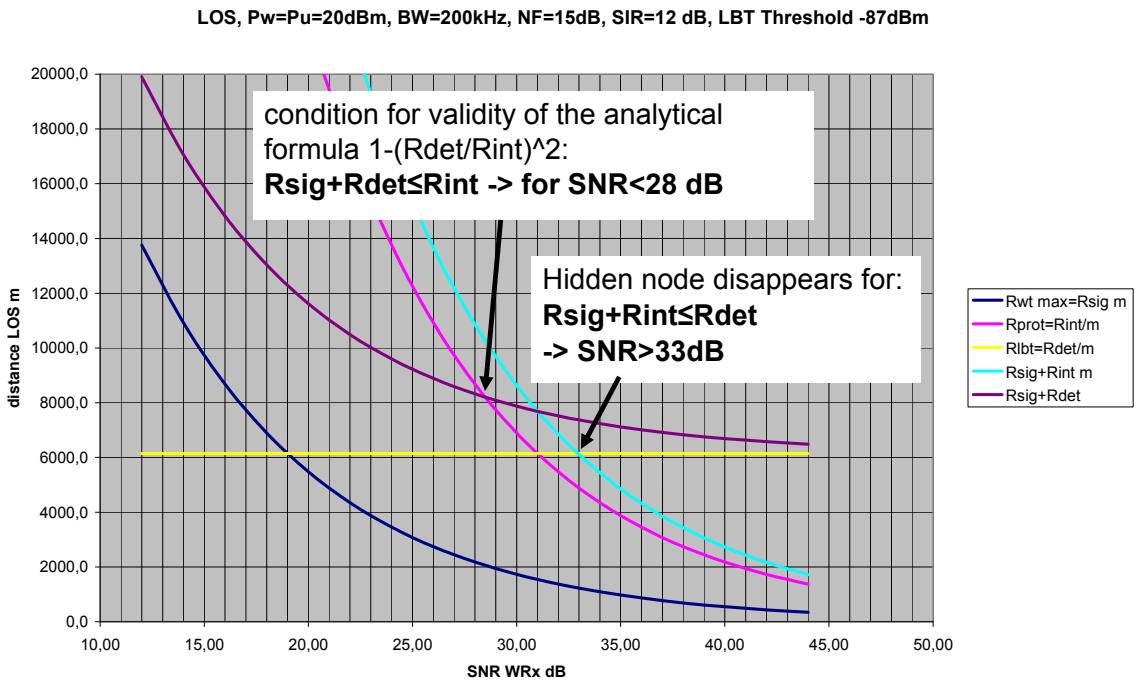


Figure 37: Distances for LOS conditions (exp.2)

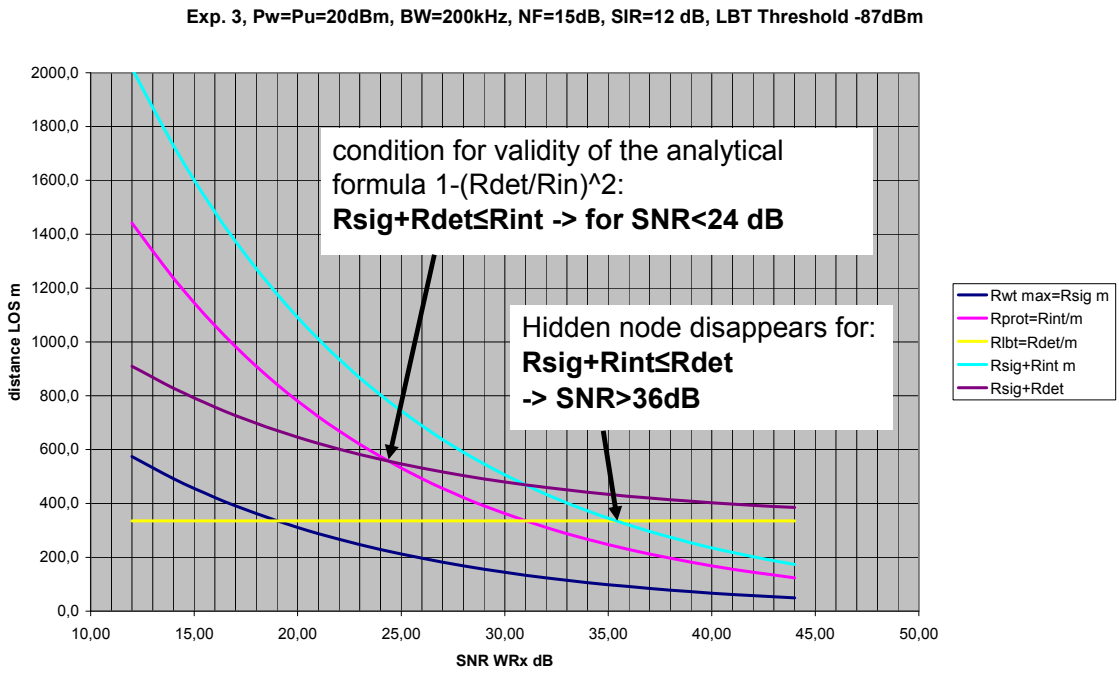


Figure 38: Distances for non LOS conditions (exp.3)

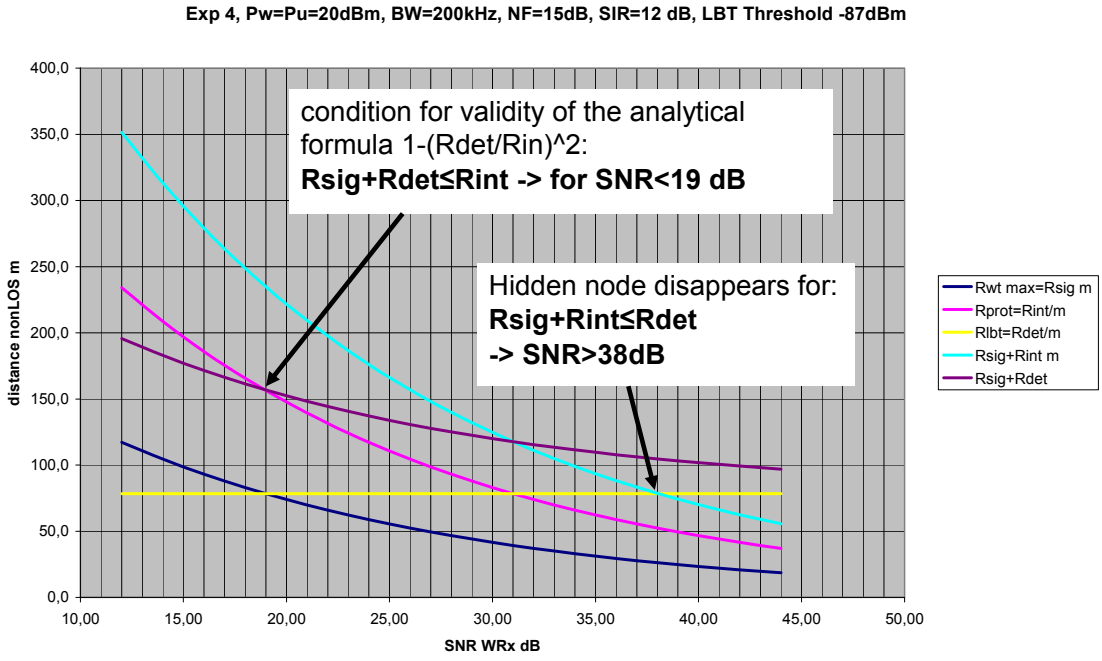


Figure 39: Distances for non LOS conditions (exp.4)

The following figure shows the results of Formula (11) as solid line and the linear intrapolation between Rsig+Rdet=Rint and Rsig+Rint=Rdet as dotted line.

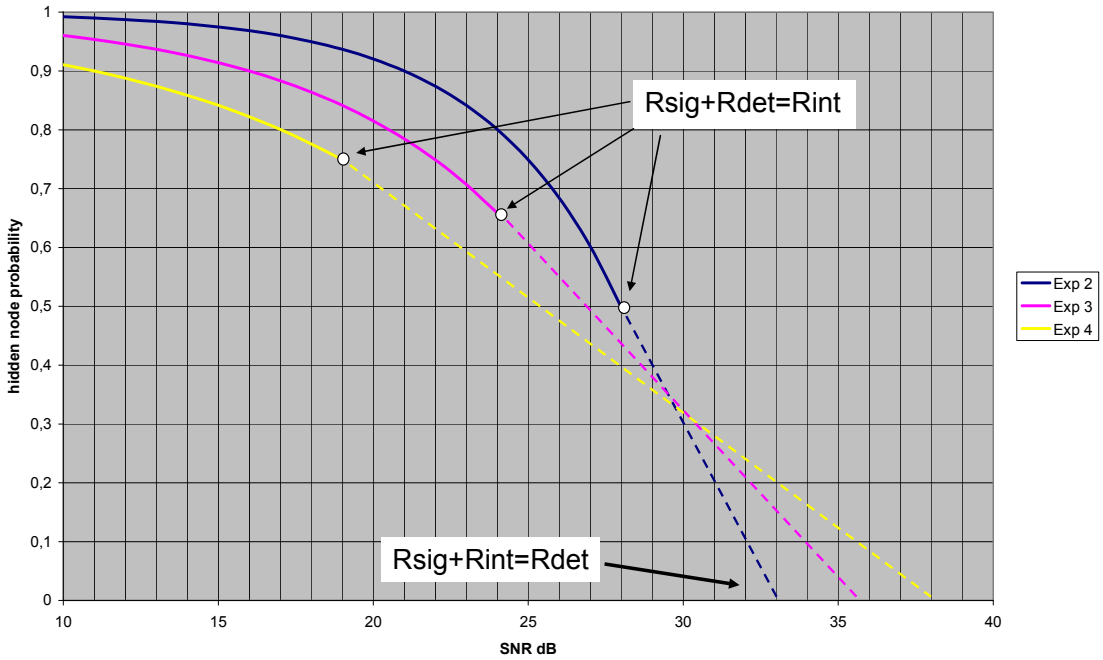


Figure 40: Hidden node probability for the balanced case as function of SNR at the Victim Receiver and the propagation condition

A1.4 THE UNBALANCED SCENARIO

The hidden node probability is even more critical if the victim link transmits with less power.

With formula (10) and Pit=20dBm, Pwt=0dBm, Pthr=-87dBm, N=-106 and SIRmin=12dB the hidden node probability is:

HN = 1- 10^[(SNR-51)/5n]

The following figure shows these results in comparison to the balanced case.

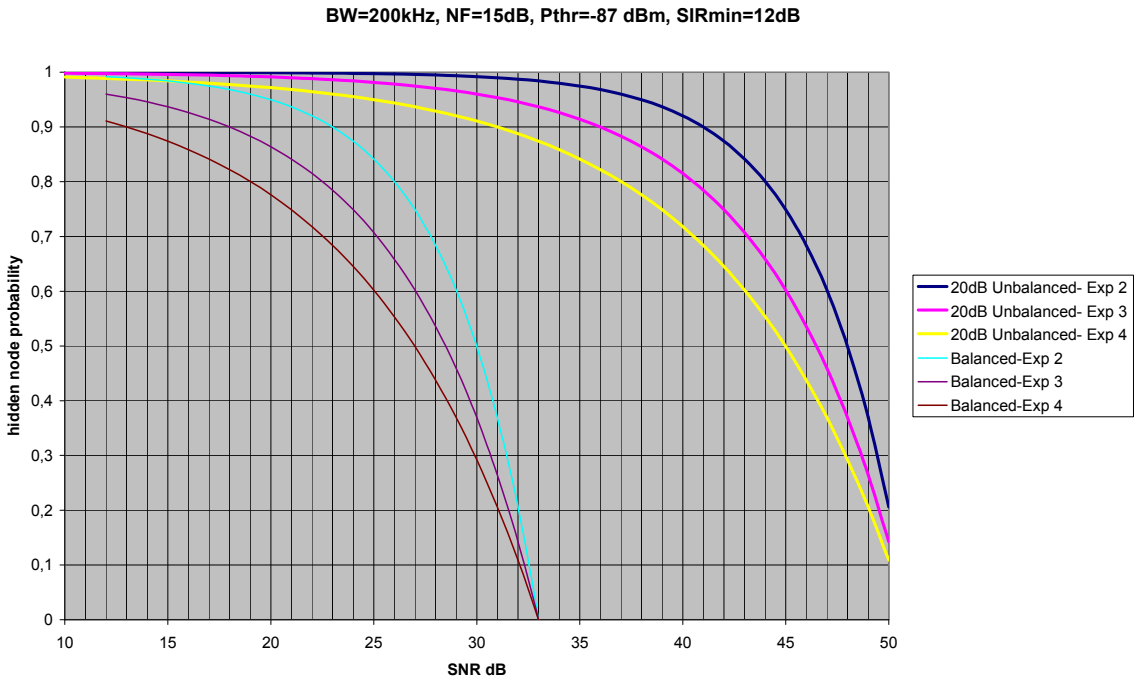


Figure 41: hidden node probability for the unbalanced case as function of SNR at the Victim Receiver and the propagation condition

ANNEX 2: LBT SEAMCAT ANALYSIS

A2.1 INTRODUCTION

The SEAMCAT simulations provided in this Annex aims to verify the available pure theoretical studies on the hidden node effects provided in Annex 1. Figure 42 shows how SEAMCAT works.

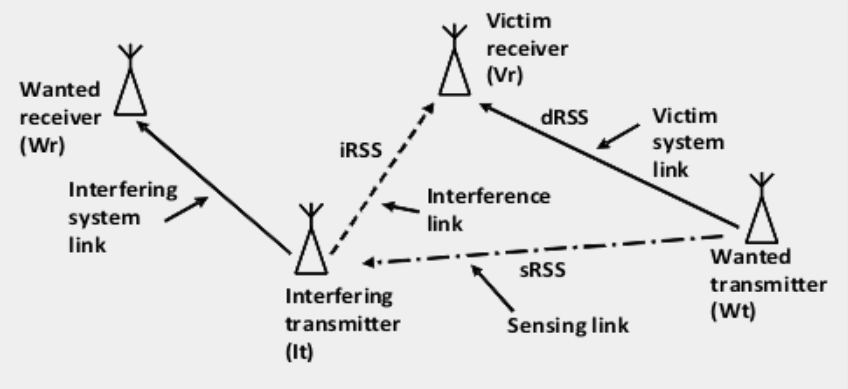


Figure 42: SEAMCAT simulations

Section A2.1 provides results for balanced Tx power values for victim and interferer, while Section A2.2 considers an unbalanced situation. SRD with LBT vs other SRDs with same power

The following table provided the assumed parameters for the victim link.

Table 11: Wanted transmitter to Victim receiver path = Victim System Link

Parameters	
F	868 MHz
Bandwidth BW	200 kHz
Noise floor	-106 dBm/BW (F=15dB)
C/I	10 dB
Height over ground	1.5m
Antenna gain Tx and Rx	0 dBi
Tx power	20 dBm/BW , see Figure 43
CR feature is selected	
coverage radius	User defined 0.1km
Propagation model	• Low margin case: Extended Hata SRD, suburban, indoor-indoor, above roof • Medium margin case: Extended Hata SRD, suburban, outdoor-indoor, above roof • High margin case: Extended Hata SRD, suburban, outdoor-outdoor, above roof

The Table below gives an overview about the assumed dRSS and SNR distributions (details are given in Appendixes 1, 2 and 3).

Table 12: dRSS distribution for the 3 scenarios

Parameters	Low margin	Medium margin	High margin
dRSS is higher as the following figure for 90% of cases (SNR in brackets)	-100 dBm (6 dB)	-87 dBm (19 dB)	-75 dBm (31 dB)
dRSS is higher as the following figure for 50% of cases (SNR in brackets)	-80 dBm (26 dB)	-67 dBm (39 dB)	-57 dBm (49 dB)

Figure 43 shows the transmitter mask of all involved transmitters and Figure 44 details of the chosen propagation model.

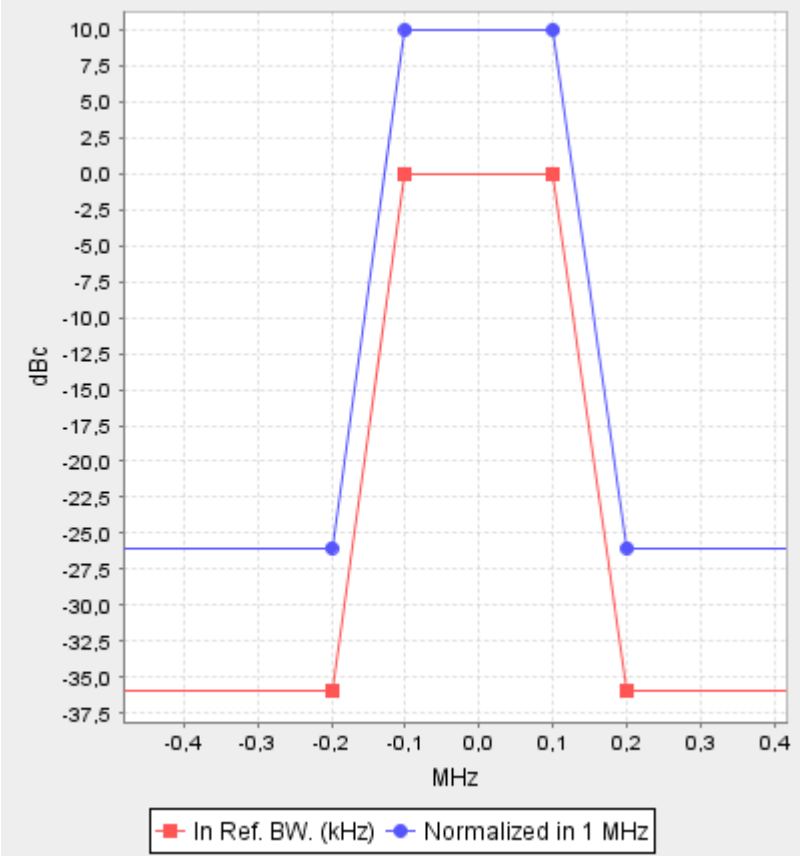


Figure 43 : Transmitter Tx mask

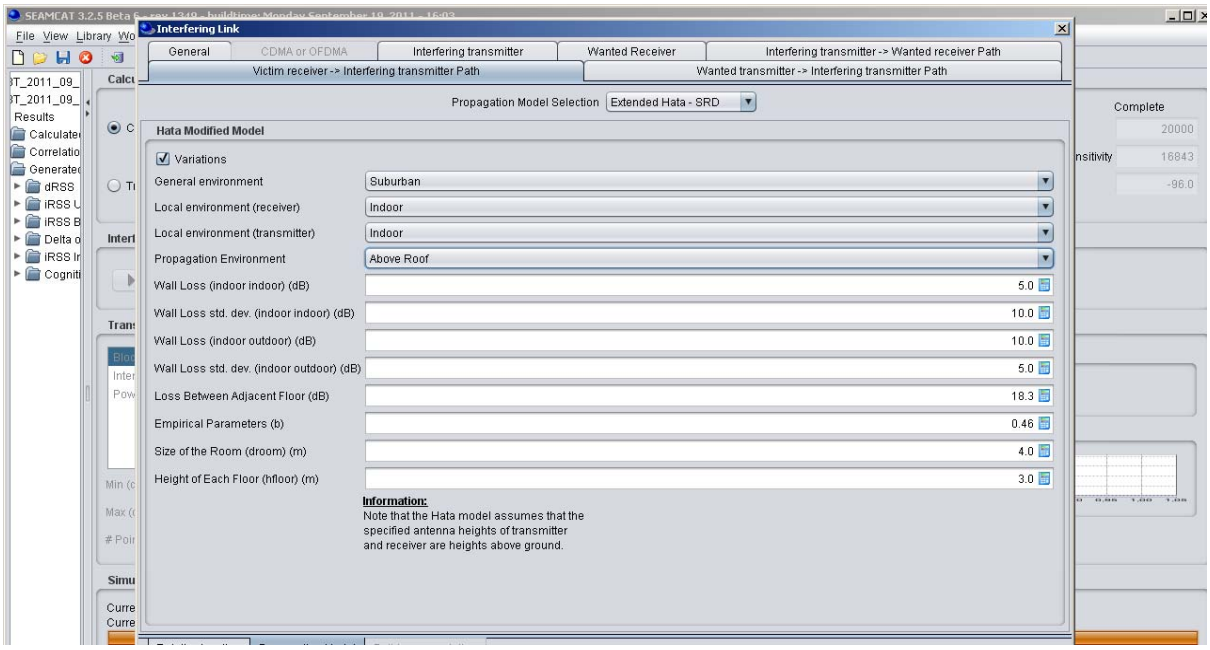


Figure 44 : Propagation model of all links for the low margin case

The following tables provide the assumed parameters for the interfering link.

Table 13: Interfering transmitter IT to Wanted receiver WR path = Interfering system link

Parameters	Values
Frequency	868 MHz
Bandwidth BW	200 kHz
Sensitivity	-96 dBm/BW
Height over ground	1.5m
Antenna gain Tx and Rx	0 dBi
Tx power and mask	20 dBm/BW, see Figure 3
CR feature is selected	
Wanted transmitter to victim receiver path, coverage radius	User defined 0.1km
Propagation model IT to WR	• Low margin case: Extended Hata SRD, suburban, indoor-indoor, above roof • Medium margin case: Extended Hata SRD, suburban, outdoor-indoor, above roof • High margin case: Extended Hata SRD, suburban, outdoor-outdoor, above roof

Table 14: Interfering transmitter IT to victim receiver VR path = interference link

Parameters	Values
Interferer distribution	Uniform density
Interferers density	20/km^2
Activity factor	1
Probability of transmission	1
Protection distance	0km
Number of active transmitters n	1
Simulation radius (=Rsimu^2=n/(pi*density)	0.126 km
Propagation model (IT to VR)	• Low margin case: Extended Hata SRD, suburban, indoor-indoor, above roof • Medium margin case: Extended Hata SRD, suburban, outdoor-indoor, above roof • High margin case: Extended Hata SRD, suburban, outdoor-outdoor, above roof

The following table provides the assumed parameters for the sensing link.

Table 15: Wanted transmitter to interfering transmitter path = sensing link

Parameters	Values
Detection threshold	-87 dBm and others
Probability of failure	Not selected
Sensing reception bandwidth	200 kHz
e.i.r.p. max inblock limit	20dBm / 200kHz from -10MHz to +10MHz
Propagation model (WT to IT)	• Low margin case: Extended Hata SRD, suburban, indoor-indoor, above roof • Medium margin case: Extended Hata SRD, suburban, outdoor-indoor, above roof • High margin case: Extended Hata SRD, suburban, outdoor-outdoor, above roof

Figure 45 shows the distribution of the devices in space.

Figure 45: Distribution of the devices in space

Table 16 gives the results for the low margin case,

Table 16: SEAMCAT results for the low margin case

	Case 0a (LBT switched off)	Case 1a (LBT -87)	Case 2a (LBT -87)	Case 3a (LBT -87)
Interfering Transmitter Uniform density /km2	20	20	20	20
Activity factor	1	1	1	1
Wanted transmitter Tx power	20dBm	20dBm	20dBm	20dBm
LBT Threshold if sRSS > Thr ▪ then Tx =-1000dBm ▪ else Tx specification	0 dBm	-87 dBm	-100 dBm	-120 dBm
sRSS mean (Std dev)	-80 dBm (17 dB)			
dRSS mean (Std dev)	-80 dBm (17dB)			
iRSS mean (StdDev)	-81 dBm (17dB)			
Probability of interference (C/I<10dB) for the low margin case	58%	19%	6%	0.28%
	Improvement by LBT	67 %	90 %	99.5 %
	Hidden node portion	33 %	10 %	0.5 %

Table 17 for the medium margin case.

Table 17: SEAMCAT results for the medium margin case

	Case 0a (LBT switched off)	Case 1a (LBT -87)	Case 2a (LBT -87)	Case 3a (LBT -87)
Interfering Transmitter Uniform density /km2	20	20	20	20
Activity factor	1	1	1	1
Wanted transmitter Tx power	20dBm	20dBm	20dBm	20dBm
LBT Threshold if sRSS > Thr ▪ then Tx =-1000dBm ▪ else Tx specification	0 dBm	-87 dBm	-100 dBm	-120 dBm
sRSS mean (Std dev)	-68 dBm (14 dB)			
dRSS mean (Std dev)	-67 dBm (14dB)			
iRSS mean (StdDev)	-69 dBm (14dB)			
Probability of interference (C/I<10dB)	64%	6%	1%	0%
	Improvement by LBT	90.6 %	98.4 %	100 %
	Hidden node portion	9.4 %	1.6 %	0 %

Table 18 for the high margin case.

Table 18: SEAMCAT results for the high margin case

	Case 0a (LBT switched off)	Case 1a (LBT -87)	Case 2a (LBT -87)	Case 3a (LBT -87)
Interfering Transmitter Uniform density /km2	20	20	20	20
Activity factor	1	1	1	1
Wanted transmitter Tx power	20dBm	20dBm	20dBm	20dBm
LBT Threshold if sRSS > Thr ▪ then Tx =-1000dBm ▪ else Tx specification	0 dBm	-87 dBm	-100 dBm	-120 dBm
sRSS mean (Std dev)	-57 dBm (13 dB)			
dRSS mean (Std dev)	-57 dBm (13dB)			
iRSS mean (StdDev)	-59 dBm (13dB)			
Probability of interference (C/I<10dB)	66%	1%	0.1%	0%
	Improvement by LBT	98.5%	99.8%	100%
	Hidden node portion	1.5%	0.2%	0%

Table 19 gives a summary of the simulations.

Table 19: Summary SEAMCAT simulations

Victim Link SNR margin	Probability of interference		
	Without LBT	LBT with -87dBm threshold	Hidden Node portion
Low margin SNR 90% >10dB	58%	19%	33 %
Medium margin SNR 90% >20dB	64%	6%	10 %
High margin SNR 90% >30dB	66%	1%	2 %

A2.2 SRD WITH LBT VS OTHER SRDs WITH DIFFERENT POWER

In this section the unbalanced scenario is analysed, that means the victim link works with 20dB less Tx power. As a consequence of this the coverage radius of the victim link was changed accordingly. In absence of this two changes in the victim link (Tx power, Coverage radius) all other parameters are the same as in section A2.1.

Table 20: Wanted transmitter to Victim receiver path = Victim System Link

Parameters	Values
F	868 MHz
Bandwidth BW	200 kHz
Noise floor	-106 dBm/BW (F=15dB)
C/I	10 dB
Height over ground	1.5m
Antenna gain Tx and Rx	0 dBi
Tx power	0 dBm/BW
CR feature is selected	
coverage radius	User defined 0.033km
Propagation model	Low margin case: Extended Hata SRD, suburban, indoor-indoor, above roof Medium margin case: Extended Hata SRD, suburban, outdoor-indoor, above roof High margin case: Extended Hata SRD, suburban, outdoor-outdoor, above roof

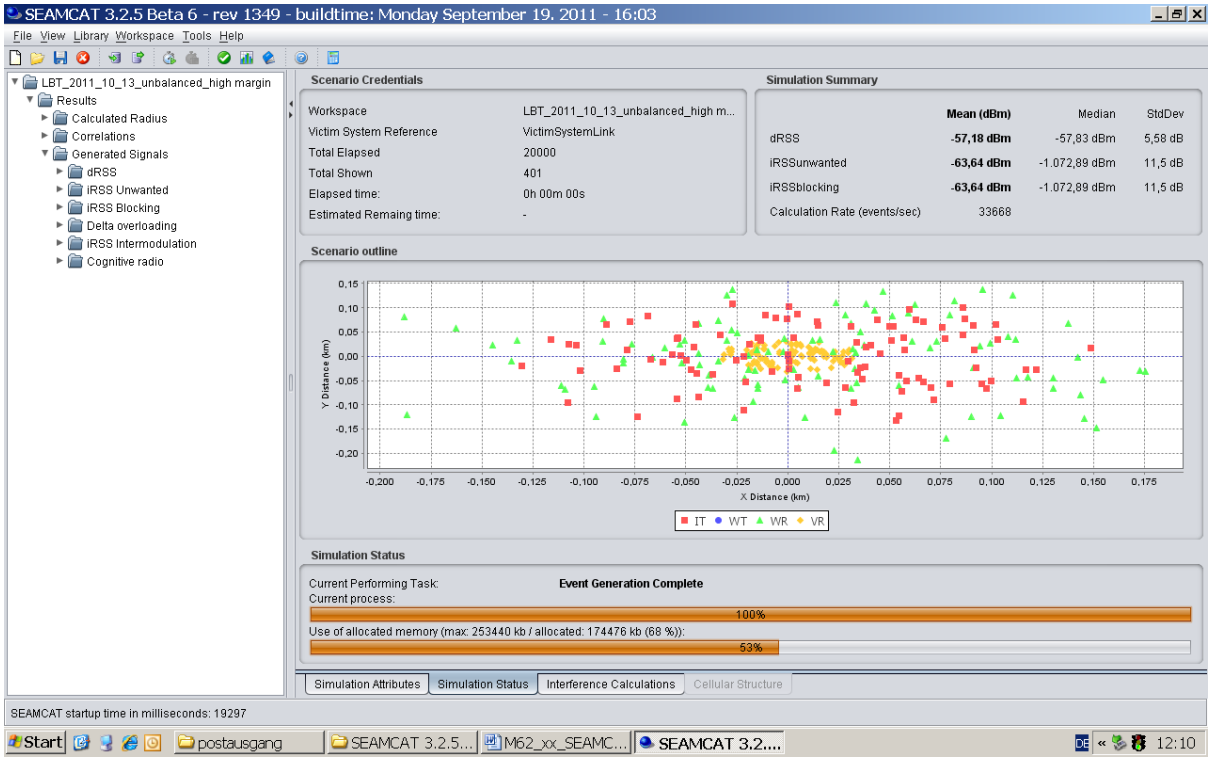


Figure 46: SEAMCAT results

Table 21 gives a summary of the simulations.

Table 21: Summary of the unbalanced SEAMCAT simulations

Victim Link SNR margin	Probability of interference		
	Without LBT	LBT with -87dBm threshold	Hidden Node portion
Low margin SNR 90% >10dB	59%	44%	75 %
Medium margin SNR 90% >20dB	68%	30%	44 %
High margin SNR 90% >30dB	70%	14%	20 %

APPENDIX 1: SIGNAL DISTRIBUTIONS FOR THE LOW MARGIN CASE

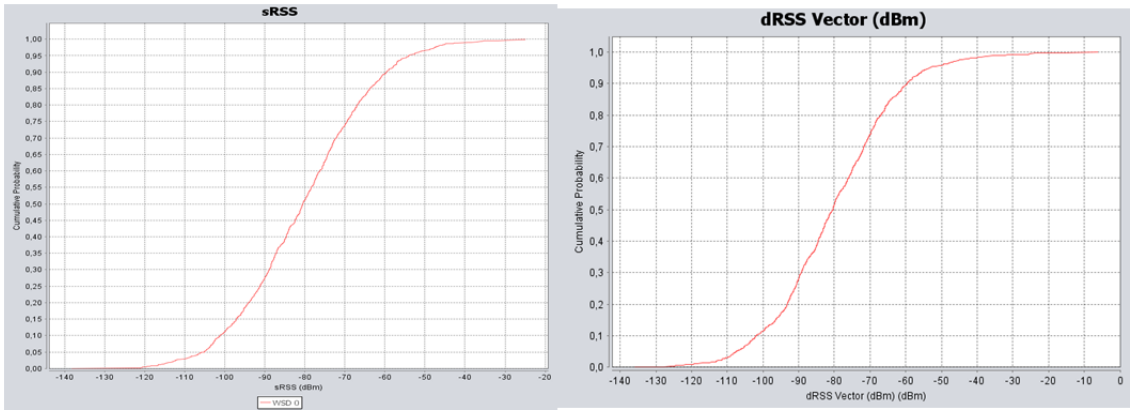
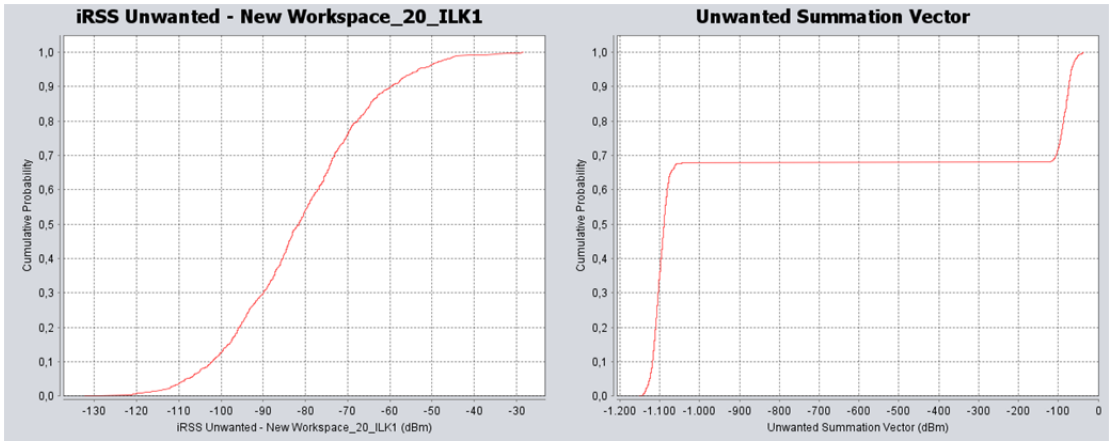


Figure 47: sRSS and dRSS for the low margin case



iRSS without LBT, iRSS with LBT (-87 dBm)

Figure 48: iRSS for the low margin case

APPENDIX 2: SIGNAL DISTRIBUTIONS FOR THE HIGH MARGIN CASE

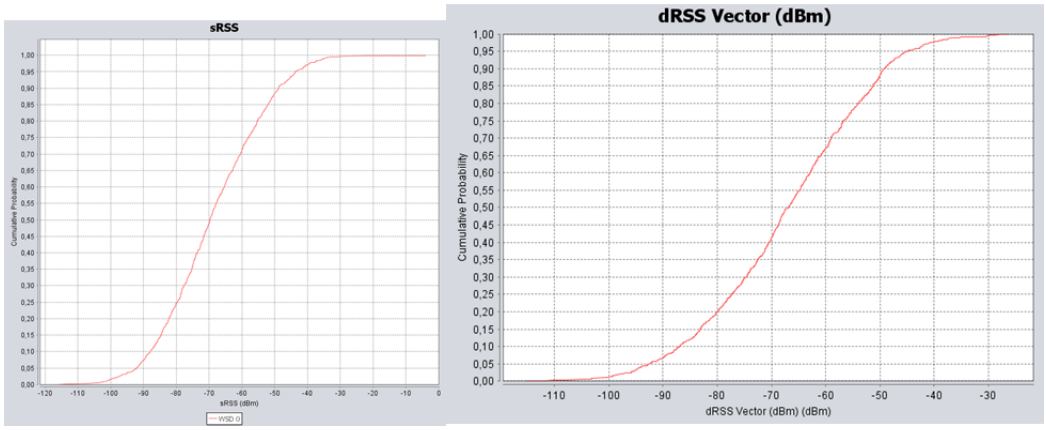
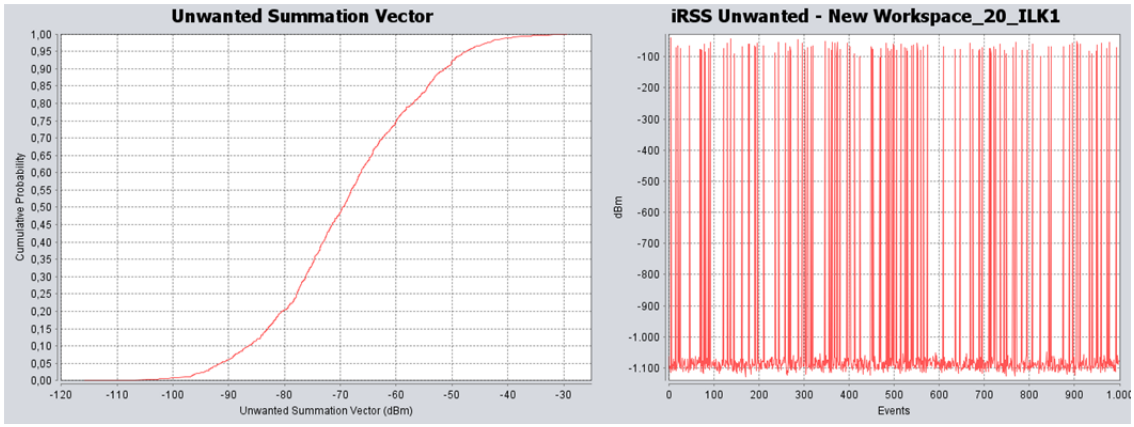


Figure 49: sRSS and dRSS for the high margin case



iRSS without LBT, iRSS with LBT (-87 dBm)

Figure 50: iRSS for the high margin case

APPENDIX 3: SIGNAL DISTRIBUTIONS FOR THE VERY-HIGH MARGIN CASE

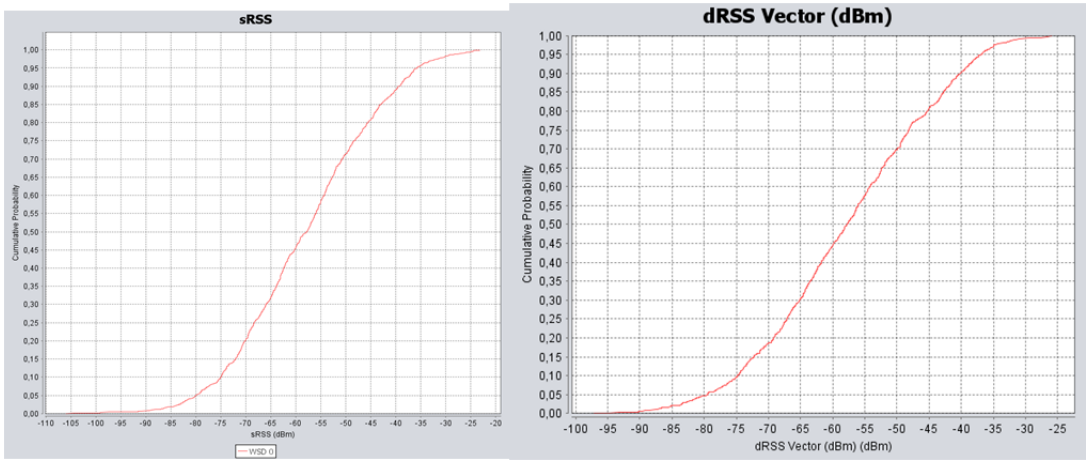
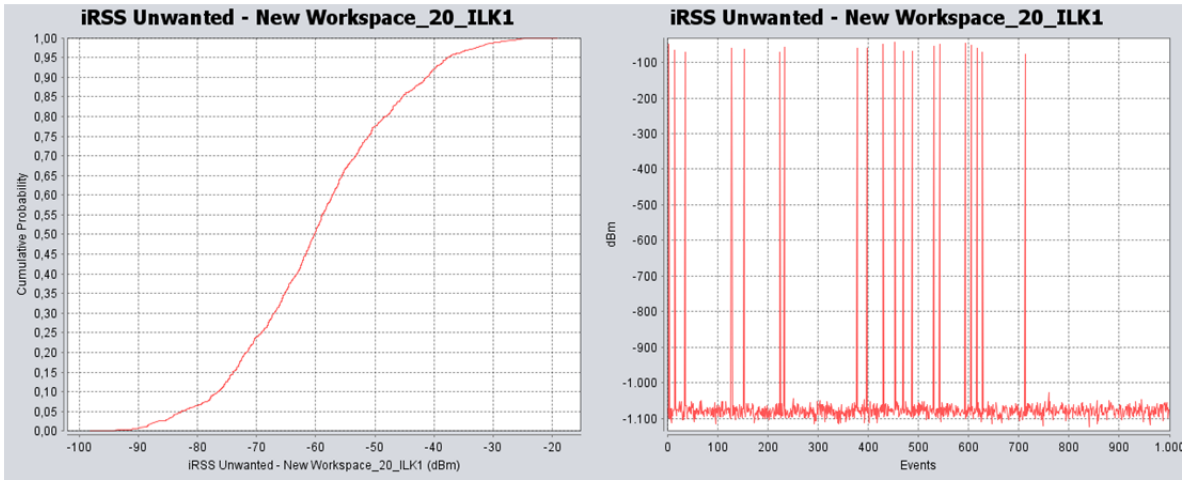


Figure 51: sRSS and dRSS for the very-high margin case



iRSS without LBT, iRSS with LBT (-87 dBm)

Figure 52: iRSS for the very high margin case

ANNEX 3: DC AND LBT SPREADSHEET SIMULATION

A3.1 THE TIME DOMAIN MONTE CARLO SIMULATION

The performance of LBT in the time domain is analysed using methods from the Signal Detection Theory.

The simulation performed in the spread sheet “RoCxxx” is a pure Monte Carlo analysis to investigate the collision conditions in the time domain. This Annex contains a short introduction into the methodology and a detailed description of the spreadsheet.

A3.2 GENERAL REMARKS ON TEST METHODOLOGY

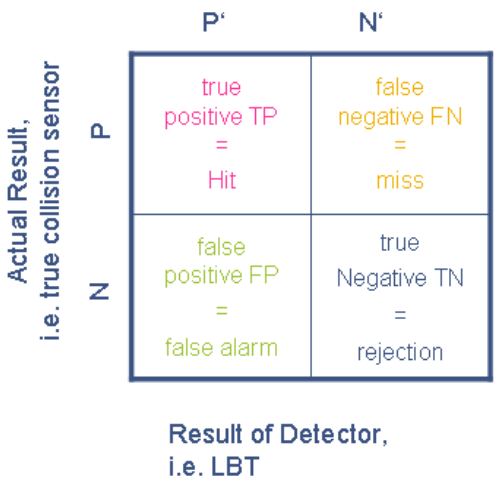
In general every kind of test can be assessed by its ability to diagnose the outcome correctly. LBT is a test on the particular criterion that an intended transmission of a radio transceiver will cause a collision with the transmission of another radio device
The test results in a binary outcome.
Positive: collision detected
Negative: no collision detected

The purpose of the simulation is to evaluate the contingency table of the test and some other statistical figures, which in the end lead to the receiver operating characteristic.

The figures of the contingency table are:

1. True Negatives (no collision detected, channel free)
   In these cases both devices transmit successfully without any conflict.
2. False Negatives (no collision detected, channel occupied)
   In these cases the LBT device cannot detect the signal of the other device, and its transmission causes a collision. The transmissions of both devices are destroyed.
3. False Positives (collision detected, channel free)
   In these cases the LBT device detects the transmission of another device and hence prevents its own transmission. But no collision would have occurred, because the detected transmission would have been terminated before the LBT-device would have started its transmission.
4. True Positives (collision detected, channel occupied)
   In these cases the LBT device detects a signal and hence prevents its own transmission.
   A collision would have occurred if the LBT device would not have prevented its transmission.

Contingency Table



Related figures from the statistical terminology

- Sensitivity (True Positive Rate TPR)
  TPR = TP / (TP + FN)
- Specificity (SPC)
  SPC = TN / (FP + TN)
- False Positive Rate
  FPR = FP / (FP + TN)
- False Discovery Rate (FDR)
  FDR = FP / (FP + TP)
- Accuracy (ACC)
  ACC = (TP + TN) / (TP + FP + FN + TN)
- F1 score (also an accuracy value)
  F1 = 2 TP^2 / (2 TP + FN + FP)

Figure 53: Contingency table

The sensitivity of a test is the proportion of cases for which the outcome is positive that are correctly identified by the test. The specificity is the proportion of cases for which the outcome is negative that are correctly identified by the test.

Generally, both the sensitivity and specificity of a test need to be known in order to assess its usefulness. A discriminating test would have sensitivity and specificity close to 100%. However, a test with high sensitivity may have low specificity and vice versa. The decision to make use of a certain test will also depend on whether the reaction on the positive result, here to retain the transmission, has a detrimental effect in cases in which the result is a false positive.

Variations of the conditions and parameters of a test result in a change of the proportions of the cases in the contingency table. A very illustrative depiction of the applicability of a test is the receiver operating characteristic (ROC) curve. The ROC is the graph of the true positive rate, (= sensitivity) against the false positive rate (= 1 – specificity).

A perfect test would have sensitivity and specificity both equal to 1. Then the ROC curve would start at the origin (0, 0), go vertically up the y-axis to (0, 1) and then horizontally across to (1, 1). A good test would be somewhere close to this ideal, like the green curve in the diagram below.

If a test has no diagnostic capability, then it is a guess and will equally produce false positive or true positive results. This equality is represented by a diagonal line from (0, 0) to (1, 1) on the graph of the ROC curve, like the red line in the diagram below.

The blue curve represents a test that is somewhere between a very good test and a complete guess.

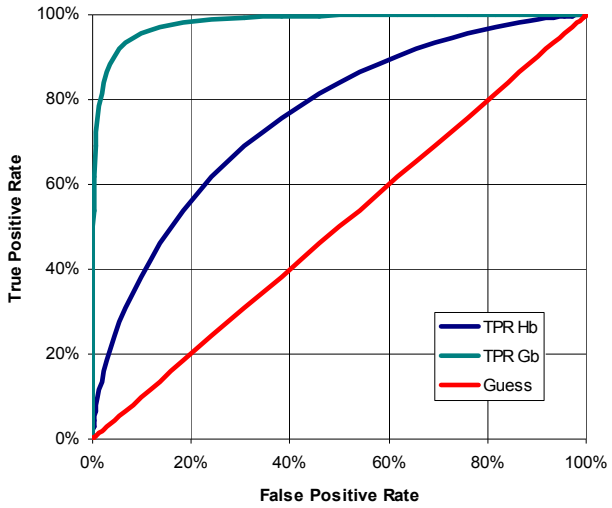


Figure 54: Receiver operating characteristics

A3.3 WORKSHEET "RESULTS"

The following parameters are defined in the sheet "Results":

Listen time	The time while a LBT-device tries to receive a foreign signal
Dead Time	The time between the end of the listening period and the begin of the transmission
Transmit Time	The duration of the transmission
DuteCycle	The proportion of the occupation of the channel in time, i.e. one transmission per interval
Sample Time	The minimum duration of a foreign signal to be detected by the LBT-device

The numbers in the yellow cells can be edited to vary the parameters.

For a certain set of parameters the calculation is performed in multiple runs for 100 to 1000 devices, where the increment is 100 devices. The results of each run are linked to row 11. After each run they are copied into the table below.

The button “Update Results” triggers a complete calculation for 100 to 1000 devices and refreshes the table with the updated figures. The button “Shuffle Random Numbers & Update Results” has the same effect but creates a new set of random numbers once before the calculation.

The most significant figures are given in the table in rows 13 to 23:

Table 22: Explanations relating to the worksheet

Column	Meaning	Description
B/2	Number of Devices	
3/C	LBT Temporal Spectrum Use Efficiency	Relation between the throughput and the offered load $LBT\ Temporal\ Spectrum\ Use\ Efficiency = \frac{LBT\ Throughput}{LBT\ Throughput + LBT\ Packet\ Loss\ Rate}$
4/D	LBT Throughput	Rate of successful transmissions, i.e. the rate of those transmissions, which do not suffer a collision and which are not retained in order to prevent a collision
5/E	LBT P 99.9 % back off delay	The time needed to achieve a probability of 99.9% to get all transmissions successfully sent. Here it is assumed that due to energy limitations every device will only transmit once per transmission interval. This delay results from the multiple repetitions of lost and retained transmissions.
6/F	DC Temporal Spectrum Use Efficiency	Relation between the throughput and the offered load $LBT\ Temporal\ Spectrum\ Use\ Efficiency = \frac{DC\ Throughput}{DC\ Throughput + DC\ Packet\ Loss\ Rate}$
7/G	DC Throughput	Rate of successful transmissions, i.e. the rate of those transmissions, which do not suffer a collision
8/H	DC P 99.9 % back off delay	The time needed to achieve a probability of 99.9% to get all transmissions successfully sent. Here it is assumed that due to energy limitations every device will only transmit once per transmission interval. This delay results from the multiple repetitions of lost transmissions.
9/I	Sensitivity TPR = Victim Protection Rate	Proportion of cases for which the outcome is positive that are correctly identified by the test (also known an true positive rate TPR)
10/J	Specificity	Proportion of cases for which the outcome is negative that are correctly identified by the test
11/K	False Positive Rate	Proportion of cases for which the outcome is positive on the whole actual negative cases (also known as type I errors or α errors)
12/L	Accuracy	Proportion of true outcomes (both true positives and true negatives) on the total cases
13/M	Positive Rate	Proportion of positive cases on the total cases
14/N	Negative Rate	Proportion of negative cases on the total cases
15/O	Test Positive Rate	Proportion of positive outcomes on the total cases
16/P	Test Negative Rate	Proportion of negative outcomes on the total cases
17/Q	False Discovery Rate	Proportion of the detected false positive results on the whole detected positive results
18/R	F1 score	Another measure of the accuracy of a test

The diagrams with the statistics and the receiver operating characteristic are then derived from the corresponding figures in the table.

It should be noted, that the curves are not smooth as one would expect. The reason therefore is that for one run of simulation only one set of random numbers is used. If multiple runs would be performed and the results would be averaged the shape of the curves would smoothen.

A3.4 WORKSHEET "CALCULATION"

The calculation itself is done in the sheet "calculation".

The calculation starts from a column of random numbers, which represent the timestamp of the transmission of each device.

All relevant figures for each device are provided in the following table.

Table 23: Relevant figures for each device

Column	Meaning	Description
1/A	Device	Index of the device
2/B	Start Listen	Begin of the listening period associated to the begin of transmission in column 4/D in the same row
3/C	Start Dead Time	End of the listening period associated to the begin of transmission in column 4/D in the same row
4/D	Start Tx	Timestamp which in the case, that the device actually transmits, is the start of the transmission To speed up the performance these numbers are copied from the sheet "Random numbers", which contains all precalculated random numbers for the relevant cases (100 to 1000 devices).
5/E	End Tx	Associated end of the transmission to the timestamp in column 4/D in the same row
6/F	LBT Causing Collision	Indicator that the transmission of an ideal LBT device would cause a collision with one or more preceding transmissions If the end of any transmission in the preceding rows is later than the begin of the transmission in column 7/G, i.e. two transmissions overlap. "1" indicates that the transmission in this row would cause a collision.
7/G	True LBT Tx Start	Start of the transmission of an ideal LBT device. A Transmission would be possible without causing collision. The cell remains empty if a collision would occur, i.e. if the ideal LBT device would prevent its transmission. This is an auxiliary column which is needed for the calculation of column 6/F.
8/H	True LBT Tx End	End of the transmission of an ideal LBT device. A Transmission would be possible without causing collision. This is an auxiliary column which is needed for the calculation of column 6/F.
9/I	LBT sensing collision	Indicator that an LBT device detects a collision with one or more preceding transmissions If the end of any transmission in the preceding rows is later than the begin of the listening period in column 7/G, i.e. two transmissions overlap, a "1" indicates that the LBT device in this row would detect a collision.
10/J	LBT Tx Start	Start of the transmission of an real LBT device. The cell remains empty if the real LBT device prevents its transmission.
11/K	LBT Tx End	End of the transmission of an real LBT device. The cell remains empty if the real LBT device prevents its transmission.
12/L	LBT causing Collision	Indicator that the transmission of an real LBT device would cause a collision with one or more preceding transmissions If the end of any transmission in the preceding rows is later than the begin of the transmission in column 10/J, i.e. two transmissions overlap, a "1" indicates that the transmission in this row would cause a collision.
13/M	LBT suffering collision	Indicator that an LBT device suffers a collision from one or more succeeding transmissions If the start of any transmission in the succeeding rows is earlier than the end

Column	Meaning	Description
		of the transmission, i.e. two transmissions overlap, a “1” indicates that the transmission of the LBT device is destroyed by a collision with another transmission. This is needed to get the total collisions of LBT devices.
14/N	LBT collisions	Indicator for the collision of an LBT transmission, either if it causes a collision or if it suffers a collision, i. e. “LBT suffering collision” and “False Negative” are ORed. The entries in this column are summed up to get the total LBT collisions and not to count the same entry multiple times.
15/O	True Positive	Indicator for TP-condition, i.e. collision detected, channel occupied
16/P	False Positive	Indicator for FP-condition, i.e. collision detected, channel free
17/Q	True Negative	Indicator for TN-condition, i.e. no collision detected, channel free
18/R	False Negative	Indicator for FN-condition, i.e. no collision detected, channel occupied
19/S	DC causing Collision	Indicator that the transmission of a DC device would cause a collision with one or more preceding transmissions If the end of any transmission in column 5E in the preceding rows is later than the begin of the transmission in column 4/D, i.e. two transmissions overlap. “1” indicates that the transmission in this row would cause a collision.
20/T	DC suffering Collision	Indicator that an DC device suffers a collision from one or more succeding transmissions If the start of any transmission in the succeding rows is earlier than the end of the transmission, i.e. two transmissions overlap, a “1” indicates that the transmission of the LBT device is destroyed by a collision with another transmission. This is needed to get the total collisions of DC devices.
21/U	DC Collisions	Indicator for the collision of an DC transmission, either if it causes a collision or if it suffers a collision, i. e. “DC causing collision” and “DC suffering collision” are ORed. The entries in this column are summed up to get the total DC collisions and not to count the same entry multiple times.

Relevant figures for the entirety of devices are provided in the following table.

Table 24: Relevant figures for the entirety of devices

Cell	Meaning	Description
Z3S5: Z6S8 / E3:H6	Contingency Table for LBT devices	Table of the four conditions as described above: F4/Z4S6 : true positive outcomes (TP) F5/Z5S6 : false positive outcomes (FP) G4/Z4S7 : false negative outcomes (FN) G5/Z5S7 : true negative outcomes (TN) F6/Z6S6 : positive outcomes (= TP + FP) G6/Z6S7 : negative outcomes (= FN + TN) H4/Z4S8 : actual positive cases H5/Z5S8 : actual negative cases
Z2S10: Z8S12 / J2:L8	Transmission Totals LBT	Figures that characterise LBT spectrum access
Z3S10/J3	LBT Temporal Spectrum Use	Proportion of the transmission interval which is occupied by the entirety of the LBT devices
Z4S10/J4	LBT Temporal Spectrum Use Efficiency	see explanation in clause ‘Worksheet “Results”’
Z5S10/J5	LBT Packet Loss Rate	Rate of those transmissions which are lost due to a collision with another transmission

Cell	Meaning	Description
Z6S10/J6	LBT Packet Delay Rate	Rate of those transmissions which have to be repeated because they either are lost due to a collision or are retained because the LBT mechanism detects a potential collision
Z7S10/J7	LBT Throughput	see explanation in clause 'Worksheet "Results"'
Z8S10/J8	LBT P 99.9 % back off delay	see explanation in clause 'Worksheet "Results"'
Z3S14: Z7S17 / N3:Q6	Contingency Table for DC devices	Table of the four conditions as described above (but there are no positive outcomes, because DC devices have no detection capability; the 'outcome' is always assumed as negative): O4/Z4S15 : n.a. O5/Z5S15 : n.a. P4/Z4S16 : false negative 'outcomes' P5/Z5S16 : true negative 'outcomes' O6/Z6S15 : n.a. P6/Z6S16 : negatives 'outcomes' Q4/Z4S17 : actual positive cases Q5/Z5S17 : actual negative cases
Z2S19: Z8S21 / S2:U8	Transmission Totals DC	Figures that characterise DC spectrum access
Z3S19/S3	DC Temporal Spectrum Use	Proportion of the transmission interval which is occupied by the entirety of the DC devices
Z4S19/S4	DC Temporal Spectrum Use Efficiency	see explanation in clause 'Worksheet "Results"'
Z5S19/S5	DC Packet Loss Rate	Rate of those transmissions which are lost due to a collision with another transmission
Z6S19/S6	DC Packet Delay Rate	Rate of those transmissions which have to be repeated because they are lost due to a collision
Z7S19/S7	DC Throughput	see explanation in clause 'Worksheet "Results"'
Z8S19/S8	DC P 99.9 % back off delay	see explanation in clause 'Worksheet "Results"'
Z2S23: Z8S25 / W2:Z8	LBT Statistics	Statistical figures to assess the applicability of LBT as a test method
Z3S23/ W3	True Positive Rate (Sensitivity)	see explanation in clause 'Worksheet "Results"'
Z4S23/ W4	Specificity	see explanation in clause 'Worksheet "Results"'
Z5S23/ W5	False Positive Rate	see explanation in clause 'Worksheet "Results"'
Z6S23/ W6	False Discovery Rate	see explanation in clause 'Worksheet "Results"'
Z7S23/ W7	Accuracy	see explanation in clause 'Worksheet "Results"'
Z8S23/ W8	F1 score	see explanation in clause 'Worksheet "Results"'

A3.5 WORKSHEET "RANDOM NUMBERS"

The sheet "Random numbers" contains all pre-calculated sets of random numbers for the relevant cases (100 to 1000 devices). This ensures that the same set of random numbers is available for several users compared to the alternative approach to generate the numbers
Just before running the simulation. Additionally the use of pre-calculated numbers speeds up the performance of the simulation

ANNEX 4: SIMULATION SPREADSHEET

A4.1 INTRODUCTION

In order to further the Performance Assessment objective in section 6, a spreadsheet was developed to test the effect of different strategies adopted by an SRD device and the effect of different regulatory parameters.

The spreadsheet calculates the probability of successful transmission of a message in various scenarios. In this instance, a scenario is a combination of a particular level of traffic and a particular set of regulatory parameters.

A4.2 DC AND LBT STRATEGIES

For the given scenario, the spreadsheet calculates the probability of success for various strategies. These strategies are the four combinations of wanted transmitter using DC or LBT and the interfering traffic using DC or LBT. It is to be stressed that this is not mandatory LBT (i.e, imposed by regulation as means of sharing) but voluntary LBT (ie, the device takes steps to improve its performance). In the spreadsheet, devices using LBT are also required to meet the limits on duty cycle, transmission time, etc.

The rationale is simply that a device placed on the market should be expected to do what is reasonably possible to maximise its performance, even if the details of that are not forced upon it by the regulations.

A4.3 LATENCY/RELIABILITY TESTS

The spreadsheet also attempts to model latency and reliability outcomes. It does this by setting a time limit for successful transmission and a target success probability. Within the time limit the device makes a number of transmissions according to the regulatory limits imposed and its own strategy. The result is given as the probability of success within the time limit for each strategy, and an indication of whether this exceeded the threshold.

A4.4 OUTPUT DISPLAY

The output display is presented as the results of three tests.

Test A is the success probability for a single transmission with each of the four strategies.

Test B is a latency/reliability test that represents low latency systems

Test C is a latency/reliability test that represents high reliability systems. In this an additional limit on the number of attempts may be imposed.Calculations

The calculations of collision probability are as set out in section 3.3.1 and Annex 5 of this report. In addition, an estimate of the predicted wait time if LBT is used has been made using Queuing Theory. For the purposes of this analysis, it is believed an estimate is sufficient and the underlying calculation is described in section 2.11.3 and in the spreadsheet.

A4.5 PRELIMINARY RESULTS

Shown below: a typical result display from the spreadsheet

Regulatory Limits		
Max Ton	sec	0.5
Min Toff	sec	0.05
Max DC long term		0.01
Max DC short term		1

LBT Parameters		
Response Time	sec	0.002
Dead time	sec	0.001
Risk window	sec	0.003

Occupier Behaviour		
Number of users		25
Assumptions:		
Use of Toff		
Individual Tx rate	by T limits	1.818182
	by DC limi	0.02
Overall	per sec	0.02
Total Tx rate		0.5
Total normalised traffic		**0.25**

Test	**A**	**Single transmission**	
This block shows success rate of an individual packet			
Message Duration	sec	0.1	
Target success rate		95.0%	

Strategy		Prob
Random Transmit		**73.95%**
Wanted Tx uses LBT		**94.93%**
Interferers use LBT		**77.66%**
All use LBT		**99.70%**

Test	**B**	**Multiple transmissions**	
This block represents Low Latency Systems			
Message Duration	sec	0.05	
Max Latency	sec	0.2	
Target success rate		95.0%	

Strategy	Attempts	Prob
Random Transmit	2	**94.16%**
Wanted Tx uses LBT	0.9	**96.37%**
Interferers use LBT	2	**95.01%**
All use LBT	0.9	**99.50%**

Test	**C**	**Multiple transmissions**	
This block represents High Reliability Systems, eg Alarms			
Message Duration	sec	0.1	
Max Latency	sec	3	
Max attempts		5	
Target success rate		99.5%	

Strategy	Attempts	Prob
Random Transmit	5	**99.88%**
Wanted Tx uses LBT	5.0	**100.00%**
Interferers use LBT	5	**99.94%**
All use LBT	5.0	**100.00%**

Figure 55: Typical results

A4.6 VOLUNTARY USE OF LBT

For a single transmission, the use of LBT by the sender always reduces the collision probability. The penalty is the need for a receiver and the waiting time. In most cases, however, apart from very high congestion, the expected wait time is low. A typical figure with messages of 100 ms in a moderately congested environment (10% normalised traffic) was 11 ms.

LBT does not eliminate collisions entirely, although if all parties use it, the collision probability was reduced by a factor of about 8 in typical scenarios (note: in its current version, the spreadsheet does not account for

the hidden node problem). If only one party uses it (either the message sender or the potential interferers) then the collision probability is approximately halved.

A4.7 LATENCY/RELIABILITY

Two versions of the latency/reliability test are shown. One may be set in the region of 100 to 250 ms and 95%, considered appropriate for remote controls such as lighting; the other in the region of 3 to 5 seconds and 99.9%, considered appropriate for alarm systems.

For the very low latency case, the use of LBT gave significant benefits. This was certainly so where an LBT system would send 3 or 4 messages compared to a non LBT one sending 4 or 5. The improved performance of each message more than offset the reduced rate of message sending.

It was, however, possible to construct scenarios with the opposite effect, where the use of LBT gave a worse result. This occurred in high congestion scenarios with long message times, when the expected wait caused by LBT approached the latency time limit. In those cases the LBT system was able to send zero or one messages while the non LBT could send 2 or 3. It should be noted, though, that these are unusable scenarios and neither strategy gives an acceptable result. The traffic on the channel is just too difficult to deal with and the only good strategy would be to find another channel.

For the 3 to 5 second latency, almost all combinations of scenario and strategies gave a very high probability of success. Before further analysis of this case, it will be necessary to decide what rules for short term duty cycle might be applied. The scenarios chosen all had no short term duty cycle limit, in line with current regulations which average duty cycle over an hour. Please see discussion on duty cycle and activity factor in section 5.4.

A4.8 MAXIMUM ON TIME

Another feature of the spreadsheet tests the effect of imposing a maximum on time (duration) and minimum off time between transmissions.

Imposing such limits is attractive as they fit well with the idea of short term duty cycle limits as well as long term ones. For instance it solves the perceived problem of 1% duty cycle allowing transmissions of 36 seconds duration.

Initially it was thought that driving durations lower would always yield benefits to latency, but an interesting effect was observed. Below a certain level, taking the duration limit lower results in a rapid rise in collisions, because the traffic is broken up into very many small packets. On the other hand as the duration limit was taken higher, LBT wait times go proportionally higher.

The trade off is illustrated in the figure below. This is one scenario, with normalised traffic of 0.2 (moderate congestion). The multiple random and multiple LBT curves show the probability of success of sending a 50 ms transmission within a latency limit of 1 sec, and how this varies according to the duration limit (Ton).

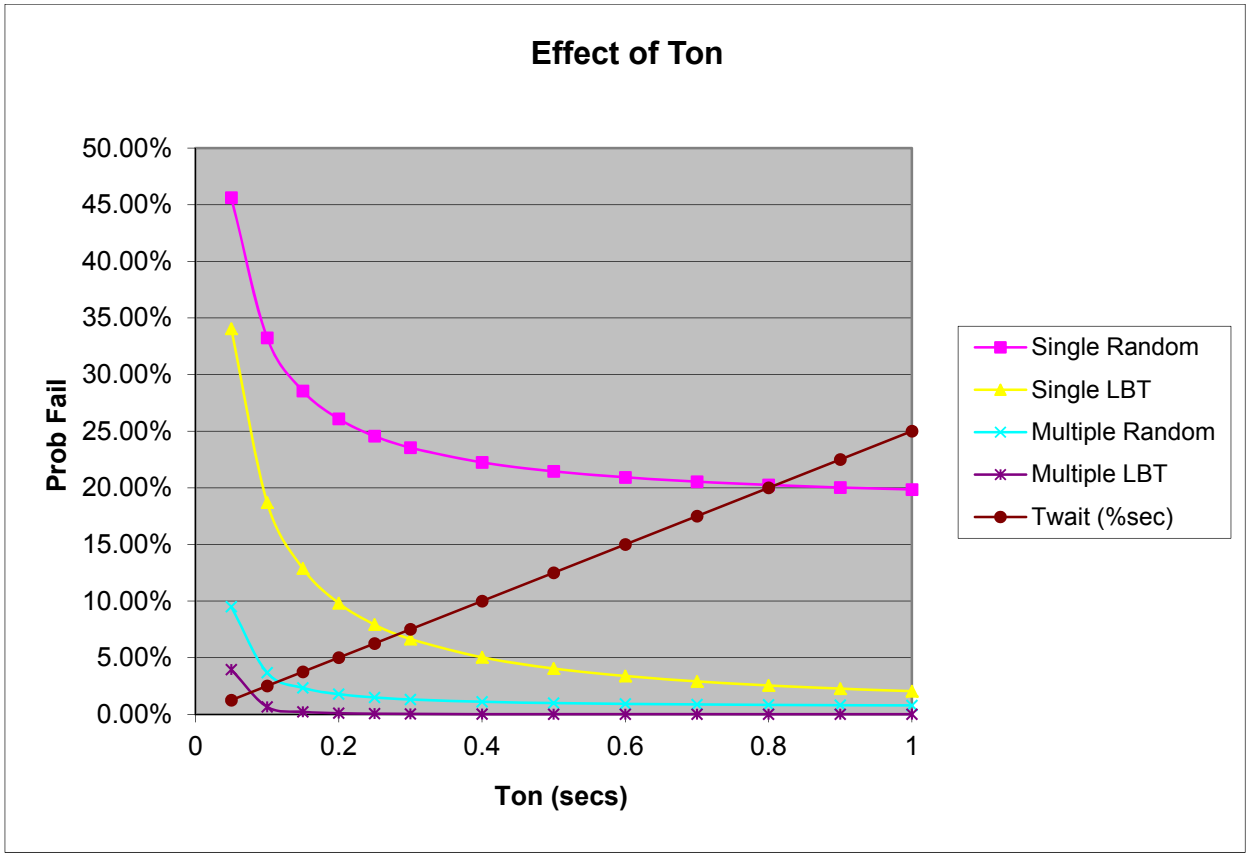


Figure 56: Effect on Ton

In this particular scenario, the optimum value for a Ton limit would be of the order of 200 ms, and similar values of around two to four times the typical transmission length are obtained in other scenarios.

ANNEX 5: COLLISION PROBABILITIES WITH DC AND LBT

This Annex develops further the analysis of collision probabilities begun in Section 3.3.1. In particular the analysis is extended to the case of multiple devices on the channel.

The equations presented here are the ones used in the spreadsheet described in Annex 4.

A5.1 LISTEN BEFORE TALK

When an LBT device attempts to transmit a message, there are three possibilities. The transmission is stopped, or it suffers a collision or it gets through.

$$P_{STOP} + P_{COLL} + P_{THRU} = 1$$

The proportion of actual transmissions that are successful is

$$P_{SUCCESS} = \frac{P_{THRU}}{P_{COLL} + P_{THRU}} = \frac{P_{THRU}}{1 - P_{STOP}}$$

A5.2 EFFECT OF RETRIES

Generally the stopped transmissions will be retried. If the retries are at random times the system is said to be operating non-persistent CSMA. The effect of retries can be modelled by applying a factor of $1/(1-P_{STOP})$ to the average repetition rate where appropriate.

In the analysis here, F_{LBT} and F_{DC} represent the traffic on the channel. F_{LBT} thus differs from the rate of attempted transmissions by the factor of $(1-P_{STOP})$. The rationale for this is that we are interested in the success probabilities of various strategies on an occupied channel, so it is the actual traffic on the channel that is important. We are also interested in the effects of various regulatory parameters and these can only be applied to the actual rather than the intended traffic. Furthermore, if a correction factor were to be applied to F_{LBT} then it could be argued that a larger factor should be applied to F_{DC} to allow for its lower success rate.

A5.3 LBT DEVICE AND SINGLE DC DEVICE

Consider the case of an LBT system occupying the same channel as a system transmitting blind but with a duty cycle limit.

The parameters of the duty cycle (DC) limited system is:
Transmit duration T_{DC}
Average repetition rate F_{DC}
Suppose an individual LBT transmission and an individual DC transmission are related as shown below.

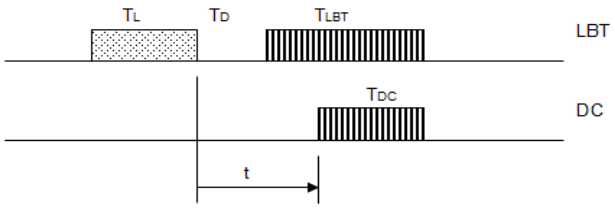


Figure 57: LBT and DC devices coinciding

The LBT system detects the DC transmission if it falls in a certain window

$$-\left(T_{DC}+T_L-T_R\right)<t<-T_R$$

I.e., there is a detection window of

$$T_{DC}+T_L-2T_R$$

If we set T_L to the minimum value of T_R (to optimise the LBT system), we can immediately write

$$P_{STOP\_LBT}=\left(T_{DC}-T_R\right)F_{DC}$$

There is a collision that is not prevented by the LBT process if the following conditions are met

$$-T_R<t<T_D+T_{LBT}$$

This is equivalent to there being a danger window in the relative timing of size

$$\left(T_D+T_R+T_{LBT}\right)$$

Therefore

$$P_{COLL\_LBT}=\left(T_D+T_R+T_{LBT}\right)F_{DC}$$

We can also note that a through transmission requires a certain window and the probability of this being clear is:

$$P_{THRU\_LBT}=1-\left(T_D+T_{LBT}+T_{DC}\right)F_{DC}$$

As a check

$$P_{THRU\_LBT}+P_{STOP\_LBT}+P_{COLL\_LBT}=1$$

and therefore

$$P_{SUCCESS\_LBT}=\frac{P_{THRU\_LBT}}{1-P_{STOP\_LBT}}=\frac{1-\left(T_D+T_{LBT}+T_{DC}\right)F_{DC}}{1-\left(T_{DC}-T_R\right)F_{DC}}$$

The probability of a collision is given by the size of the danger window and the relevant rate of transmissions. The situation is the same whichever party is considered the victim or interferer, since it is assumed that the collision destroys both messages. For the DC device, the probabilities are:

$$P_{STOP\_DC}=0$$

$$P_{COLL\_DC}=\left(T_D+T_R+T_{LBT}\right)F_{LBT}$$

$$P_{SUCCESS\_DC}=P_{THRU\_DC}=1-\left(T_D+T_R+T_{LBT}\right)F_{LBT}$$

A5.4 LBT DEVICE AND LBT DEVICE

Consider the case of two systems, each operating LBT, that attempt to make transmissions with the relative timing shown below.

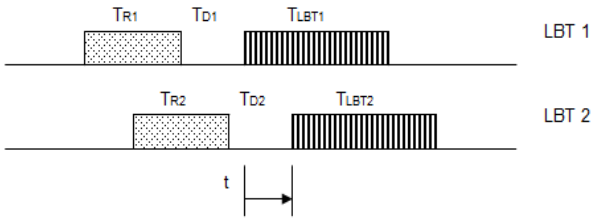


Figure 58: Two LBT transmissions coinciding

A collision occurs and is not prevented by either LBT process if the following conditions are met

$-T_{LBT2} < t < T_{LBT1}$ condition that the transmissions would collide

$-T_{D1} - T_{R1} < t < T_{D2} + T_{R2}$ condition that they would collide and do not detect each other

Assuming Tlbt1 > Td2+Tr2 then only the second condition is important, and the size of the danger window is

$$T_{D1} + T_{R1} + T_{D2} + T_{R2}$$

Therefore:

Probability of LBT 1 suffering a collision $P_{COLL\_LBT1} = (T_{D1} + T_{D2} + T_{R1} + T_{R2})F_{LBT2}$

Probability of LBT 2 suffering a collision $P_{COLL\_LBT2} = (T_{D1} + T_{D2} + T_{R1} + T_{R2})F_{LBT1}$

The detection windows and the stop probabilities are found in the same way as for the single LBT and single DC device:

$$P_{STOP\_LBT1} = (T_{LBT2} - T_{R1})F_{LBT2}$$
$$P_{STOP\_LBT2} = (T_{LBT1} - T_{R2})F_{LBT1}$$

If the two devices have the same parameters, then for each device:

$$P_{STOP} = (T_{LBT} - T_R)F_{LBT}$$

$$P_{COLL} = 2(T_D + T_R)F_{LBT}$$

$$P_{SUCCESS} = \frac{1 - P_{STOP} - P_{COLL}}{1 - P_{STOP}} = \frac{1 - (2T_D + T_R + T_{LBT})F_{LBT}}{1 - (T_{LBT} - T_R)F_{LBT}}$$

A5.5 EXTENSION TO MULTIPLE DEVICES

Consider now the various cases of a device sending a single transmission on a channel shared with multiple other devices. If all the interfering devices adopt the same strategy, then for all the combinations, we can say that:

P_{THRUN} is the probability that ALL of the other signals allow the transmission through, therefore:

$$P_{THRUN} = \left(P_{THRU}\right)^N$$

and, when the device under consideration uses LBT, then P_{STOPN} is the negative of the probability that ALL of the other signals do not cause it to stop.

$$P_{STOPN} = 1 - \left(1 - P_{STOP}\right)^N$$

Therefore:

$$P_{SUCCESSN} = \frac{P_{THRUN}}{1 - P_{STOPN}} = \frac{\left(P_{THRU}\right)^N}{\left(1 - P_{STOP}\right)^N} = \left(P_{SUCCESS}\right)^N$$

For reference the detailed equations for several cases are set out below.LBT device and N DC devices

$$P_{STOPN} = 1 - \left(1 - P_{STOP}\right)^N = 1 - \left(1 - \left(T_{DC} - T_R\right)F_{DC}\right)^N$$

$$P_{THRUN} = \left(P_{THRU}\right)^N = \left[1 - \left(T_D + T_{LBT} + T_{DC}\right)F_{DC}\right]^N$$

Therefore:

$$P_{SUCCESSN} = \frac{P_{THRUN}}{1 - P_{STOPN}} = \frac{\left[1 - \left(T_D + T_{LBT} + T_{DC}\right)F_{DC}\right]^N}{\left(1 - \left(T_{DC} - T_R\right)F_{DC}\right)^N}$$

A5.6 DC DEVICE AND N SIMILAR LBT DEVICES

$$P_{STOPN} = 0$$

$$P_{THRUN} = \left(P_{THRU}\right)^N = \left[1 - \left(T_D + T_R + T_{:BT}\right)F_{LBT}\right]^N$$

Therefore:

$$P_{SUCCESSN} = \frac{P_{THRUN}}{1 - P_{STOPN}} = \left[1 - \left(T_D + T_R + T_{LBT}\right)F_{LBT}\right]^N$$

A5.7 LBT DEVICE AND N LBT DEVICES

$$P_{STOPN} = 1 - \left(1 - P_{STOP}\right)^N = 1 - \left(1 - \left(T_{LBT} - T_R\right)F_{LBT}\right)^N$$

$$P_{THRUN} = \left(P_{THRU}\right)^N = \left[1 - \left(2T_D + T_R + T_{LBT}\right)F_{LBT}\right]^N$$

Therefore:

$$P_{SUCCESSN} = \frac{P_{THRUN}}{1 - P_{STOPN}} = \frac{\left[1 - \left(2T_D + T_R + T_{LBT}\right)F_{LBT}\right]^N}{\left(1 - \left(T_{LBT} - T_R\right)F_{LBT}\right)^N}$$

A5.8 DC DEVICE AND N SIMILAR DC DEVICES

$$P_{STOPN} = 0$$

$$P_{THRUN} = (P_{THRU})^N = [1 - 2.T_{DC}.F_{LBT}]^N$$

Therefore:

$$P_{SUCCESSN} = \frac{P_{THRUN}}{1 - P_{STOPN}} = [1 - 2.T_{DC}.F_{LBT}]^N$$

ANNEX 6: EXAMPLE BAND SEGMENTATION SCHEME

This Annex shows an example of a band segmentation scheme as discussed in section 6.4.

The assumption is that a piece of spectrum might be divided into 3 sub-bands. Each sub-band is capable of supporting more than one channel of width 200 kHz, for example.

Figure 59: Example of piece of spectrum divided into 3 sub-bands

where:

(A) is intended for low cost unidirectional systems (eg some alarms).

(B) is intended for systems needing high reliability and/or low latency (eg, more sophisticated alarms, lighting control).

(C) is intended for systems with high data throughput where traffic conditions allow.

By choosing the relative sizes of A, B and C, the overall spectrum efficiency can be optimised. The occupancy is kept low in A as this is necessary for this type of technology. The occupancy in C is encouraged to rise to the limits that network technology can allow. B is an intermediate case, where results such as low latency and high reliability (eg, for important alarm systems) are emphasised.

For instance, an example scheme might look like this:

Table 25: Example of band segmentation

Sub-band	A	B	C
Max T_{ON}	50 ms	200 ms	400 ms
Min T_{OFF}	100 ms	200 ms	50 ms
Max Duty Cycle per channel	0.1%	2%	-
Other			LBT
Expected user	Accessible with low cost simplex system	With suitable protocol, possibly including LBT, user can get low latency and/or high reliability	With suitable protocol, user can get high throughput in the uncongested case

Note that T_{ON} T_{OFF} and Duty Cycle all apply per channel. If a device changes frequency the timings are reset. Thus, Adaptive Frequency Agility is rewarded, but not required, and there is no need to write specific rules for it.

ANNEX 7: OVERVIEW OF DEVICE DUTY CYCLES

ETSI STF411 and TG28 have collected some information about the duty cycles currently in use by various SRD devices.

This has been collated and presented in a spreadsheet “Duty Cycle Overview table_ Ed3.xls”. A sheet from this is shown below.

This is part of the work on Low Duty Cycle (section 4.2 of this report) and is a work in progress.

Table 26: Example of band segmentation

Application	Latency in sec	Max Cumulated TxON time over 1 second [in seconds]	TxON time /s %	Average activity per day	TxOFF in s	Cumulated TxON per day in seconds	Max equivalent DC with current definition	Comment	Number of devices Estimated in Europe
Automotive		100ms to 30s							
remote keyless entry	200ms	0.15	15.00%	10	0.1	1.5	0.002%		Millions
convertible roof	200ms	1	100.00%	120		120	0.139%	4 actions of 30s transmission	Hundred thousands
TPMS	500ms	0.03	3.00%	20		0.6	0.001%	10 times an hour when moving / 2 hours per day	Millions
ITS CAM	100ms	0.015	1.50%	7200		108	0.125%	Average use of 2 hours per day	Launching
Home & building control		10ms to 36s							
Mains powered devices	100ms	0.025	2.50%	96	0.1	2.4	0.003%		Millions
		0.2	20.00%	20	0.1	4	0.005%		Millions
Battery powered devices	500ms	0.6	60.00%	10	0.1	6	0.007%		Millions
		1	100.00%	10	0.1	10	0.012%		Millions
Repeaters	100ms	0.025	2.50%	20	0.01	0.5	0.001%		
Smoke detectors	30s	1	100.00%	1.20	0.1	1.2	0.00139%	Minimum operation 1 test transmission of 36s per month	Hundred thousands
Low cost point to point devices	1s	1	100.00%	10	0.01	10	0.012%	Remote controlled mains adaptors (DIY)	Millions
Telemetry, telecomand		350ms to 1s							
	100ms	1	100.00%	1400	0.1	1400	1.620%		
Metering		25ms to 1.2s							
without in home display	8s	0.025	2.50%	1	0.1	0.025	0.000029%		Millions
with in home display	1s	0.025	2.50%	96	0.1	2.4	0.003%		Hundred thousands
Repeaters	100ms	0.025	2.50%	50	0.01	1.25	0.001%		Millions
EN13753 Mode R2	15min	1	100.00%	20	0.1	20	0.023%		
Alarms		25ms to 1s							
Intrusion alarm	3s	0.025	2.50%	24	0.1	0.6	0.001%		Millions
Social alarm	2s	0.15	15.00%	4	0.1	0.6	0.001%		Millions
Battery power devices	1 min	1	100.00%	24	0.1	24	0.028%		Millions
Imaging	3s	1	100.00%	6	0.1	6	0.007%	3 mins of transmission once a month	New request from customers
Referee voice system		0.1	10.00%	3600	0.9	360	0.417%		

ANNEX 8: LIST OF REFERENCES

[1] ECC Report 37: Compatibility of planned SRD applications in 863-870 MHz
[2] ERC/REC 70-03: Short Range Devices (SRD)
[3] ITU-R Report SM.2154: Short-range radiocommunication devices spectrum occupancy measurement techniques
[4] ECC Report182 on Survey about the use of the frequency band 863-870 MHz
[5] European Commission Directive 2009/140/EC on the authorisation of electronic communications networks and services
[6] R&TTE Directive: Radio and telecommunications terminal equipment
[7] Recommendation ITU-R SM.1046-2: Definition of spectrum use and efficiency of a radio system
[8] Taking Sides on Technology Neutrality - Chris Reed Professor of Electronic Commerce Law, Centre for Commercial Law Studies, Queen Mary University of London.2007
[9] B.A. Witvliet, "PhD Research Plan - Spectrum Access Mechanism for Licence Exempt radio systems", Radiocommunications Agency Netherlands, 2010.
[10] ERC Report 101 on Comparison of the minimum coupling loss method, enhanced minimum coupling loss method, and the Monte-Carlo simulation
[11] EN 300 220, ElectroMagnetic Compatibility and Radio Spectrum Matters (ERM);Short Range Devices (SRD);Radio equipment to be used in the 25 MHz to 1 000 MHz frequency range with power levels ranging up to 500 mW; Part 1: Technical characteristics and test methods
[12] Tanenbaum, “Computer Networks” 4th Ed, 2003, Pearson Education Inc.
[13] IEEE 802.11: Wireless local area network (LANs)
[14] prEN13757-4: Communication systems for meters and remote reading of meters - Part 4: Wireless meter readout (Radio Meter reading for operation in the 868-870 MHz SRD band)
[15] ETSI TR 102 886: Technical characteristics of Smart Metering (SM) Short Range Devices (SRD) in the UHF Band;System Reference Document, SRDs, Spectrum Requirements for Smart Metering European access profile Protocol (PR-SMEP)
[16] FM22(09)077 Report on monitoring 863-870 MHz
[17] EC Decision on on harmonisation of the radio spectrum for use by short-range devices
[18] ETSI TR 103 056: Short Range Devices (SRD); Technical characteristics for SRD equipment for social alarm and alarm applications
[19] CEPT Report 14 - Develop a strategy to improve the effectiveness and flexibility of spectrum availability for Short Range Devices (SRDs)
[20] Survey about the use of the frequency band 863-870 MHz