UNIVERSITÀ DEGLI STUDI DI PALERMO

Information And Communication Technologies
ING-INF/03-Telecomunicazioni
Dipartimento di Ingegneria

Exploiting physical layer flexibility for high-capacity and ultra-dense wireless networks

Ph.D. Candidate

GIUSEPPE SANTAROMITA

Coordinator

ALESSANDRO BUSACCA

Tutors

ILENIA TINNIRELLO
GIOVANNI GARBO

Cotutor

STEFANO MANGIONE

XXXII CYCLE - ACADEMIC YEAR 2018-2019

Abstract

Recent advances on wireless technologies have been focused mostly on boosting network capacity by means of a more efficient, flexible and programmable physical layer. There are many different directions that have been explored for increasing the capacity of wireless links, such as the adoption of massive multiple antennas, aggregated carriers and bigger transmission bandwidths, beam-forming, as well as innovative solutions for spectrum access. An important element of these solutions is also the increasing demand for physical layer (PHY) flexibility, in terms of reconfiguration of transmission bandwidths, modulation and coding formats, signalling mechanisms, etc. for coping with very specific application requirements and network operating conditions.

In this thesis, we focus on the problem of exploiting PHY flexibility for optimizing wireless systems in different application scenarios: i) high-capacity applications, in which the emerging extensions of the IEEE 802.11 standards exhibit some inefficiencies, due to the usage of traditional contention and signalling mechanisms in wide-band channels; ii) massive and low-range IoT applications, in which access mechanisms have to be kept simple for avoiding unnecessary energy consumption and for covering long links; iii) localization applications, in which commodity technologies used for communications can be exploited for positioning a device. Device localization can also be relevant for high-capacity applications, for example for supporting beam-forming. In all these cases, PHY layer flexibility can enable several optimization strategies. However, from an implementation point of view, supporting PHY flexibility is not an easy task, because many PHY operations need to be deployed in hardware and current Software-Defined-Radio solutions are not currently available for commercial devices. We also consider these constraints, by proposing some specific forms of flexibility, which do not involve the complete re-definition of the modulation formats, but rather just work on innovative signalling mechanisms defined on top of the available multi-carrier OFDM PHYs.

More into details, the thesis include three main contributions. First, for supporting high-capacity applications in traditional Local Area Networks, based on WiFi, we propose a new contention mechanism devised to improve the efficiency of random access in OFDM systems. Our Medium Access Control (MAC) scheme, called ReCo (Repeated Contention) is based on the possibility of supporting a contention mechanism in the frequency domain, thanks to the transmission of random tones. The scheme has very interesting properties: it is simple to configure, robust and short-term fair. We demonstrate the feasibility of the scheme in a real implementation and we also compare the scheme performance with current IEEE 802.11g and 802.11ax standards.

Second, regarding PHY layers for massive long-range IoT applications, we study LoRa technology. We start from the characterization of the LoRa PHY, which is based on a chirp-based modulation and cyclic shifts for coding different symbols. Symbols times can be tuned in terms of different Spreading Factors (SFs), which result in quasi-orthogonal transmissions. We study the robustness of LoRa PHY to the interference generated by different SFs, in order to understand the real impact of interference on the cell capacity. For such a study, we developed a LoRa received on a SDR platform, able to collect low-level information about the received radio signal. Results about interference robustness have been exploited for deriving cell capacity models. We used these models and a public-available simulator for quantifying how different criteria of the configuration of LoRa networks (such SF allocations, power control and/or packet fragmentation, employ one or more gateways, etc.) may lead to significantly different capacity results.

Finally, we propose a novel methodology that can work with very low-cost SDR receivers to localize wireless transmitters. Transmitters localization can have a key role for enabling various technologies (such as beamforming) able to improve the performance of a wireless network. Our system is able to work with signals of whatever modulation, therefore to locate whatever transmitter and also to use external reference signals, as long as their position is known, such as signals transmitted by LTE Base Stations in proximity of our experimental testbed. This last contribution has been done in cooperation with the Spanish Research Center IMDEA.

Published Content

This thesis is based on the following published papers:

P1. D. Croce, M. Gucciardo, S. Mangione, **G. Santaromita**, I. Tinnirello, "*Impact of LoRa Imperfect Orthogonality: Analysis of Link-Level Performance*", in IEEE Communications Letters, vol. 22, no. 4, pp. 796-799, April 2018.

- The authors are arranged alphabetically;
- This work is partially included and its content is reported in Chapter 3;
- The author's role in this work is focused on the study concerning all the concepts proposed in the paper.

P2. A. Baiocchi, D. Garlisi, **G. Santaromita**, I. Tinnirello, "*Moving RTS/CTS to the Frequency Domain: an Efficient Contention Scheme for 802.11ax Networks*", in ITC31-Networked Systems and Services, Budapest, Hungary, August 2019.

- The authors are arranged alphabetically;
- This work is partially included and its content is reported in Chapter 2;
- The author's role in this work is focused on the feasibility analysis of the concepts proposed in the paper, as well as a co-leader role in the implementation of simulations.

P3. A. Baiocchi, D. Garlisi, A. Lo Valvo, **G. Santaromita**, I. Tinnirello, "*'Good to repeat': Making Random Access Near-Optimal with Repeated Contentions*", in IEEE Transactions on Wireless Communications, vol. 19, no. 1, pp. 712-726, January 2020.

- The authors are arranged alphabetically;
- This work is fully included and its content is reported in Chapter 2;

- The author's role in this work is focused on the preliminary practical feasibility analysis of the concepts proposed in the paper, as well as a co-leader role in all evaluations, both simulated and experimental. He is responsible of the application of these concepts to the emerging IEEE 802.11ax standard.

P4. D. Croce, M. Gucciardo, S. Mangione, **G. Santaromita**, I. Tinnirello, "*LoRa Technology Demystified: from Link Behavior to Cell-level Performance*", accepted for publication in IEEE Transactions on Wireless Communications, 2020.

- The authors are arranged alphabetically;
- This work is fully included and its content is reported in Chapter 3;
- The author's role in this work is focused on the study concerning the concepts proposed in the paper. In particular, he played a co-leader's role in the study of the LoRa physical layer (modulation, coding and co-channel rejection).

P5. D. Croce, M. Gucciardo, S. Mangione, **G. Santaromita**, I. Tinnirello, "*Performance of LoRa Technology: Link-level and Cell-level Performance*", accepted for publication in a chapter of the book LPWAN Technologies for IoT and M2M Applications, 2020.

- The authors are arranged alphabetically;
- This work is partially included and its content is reported in Chapter 3;
- The author's role in this work is focused on the study concerning the concepts proposed in the paper. In particular, he played a co-leader's role in the study of the LoRa physical layer (modulation, coding and co-channel rejection).

Other published paper not included in the thesis:

P6. A. Lo Valvo, I. Tinnirello, F. Giuliano, **G. Santaromita**, "*DEMO: Dynamic Adaptation of WiFi Channel Widths Without TX/RX Coordination*", in MobiCom 2017, pp. 507-509, Snowbird (UT), USA, Oct 16-20, 2017.

- The authors are arranged according to their contribution;
- The author's role in this work is focused on the analysis concerning the correlation mechanism in the receiver.

Acknowledgements

First, I want to thank my family, especially my sister Giulia and my parents, mom and dad, for the love, the support and for giving me the opportunity to attend university.

Second, I am happy to meet some people, who are practically my second family: Alberto, Alice, Princio and Ciccio. Without them, these years would not be the same.

Finally, I want to thank all my other colleagues and friends, from Palermo and Capo d'Orlando. A special mention to Ilenia Tinnirello, who is the best leader in the world.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In the last decade, we have assisted to an impressive proliferation of mobile-generated traffic, which is now the highest portion of the total internet traffic and will continue to grow with the emergence of Internet-of-Things (IoT) applications. As a result, we have assisted to an high proliferation of wireless technologies, although WiFi technology is still the dominant wireless access solution. According to Cisco traffic forecasts [1], WiFi traffic from both mobile and WiFi-only devices will account for almost half of total IP traffic by 2020. The increase of mobile devices and the advent of IoT applications have been characterized by an high-density deployment of base stations, based on heterogeneous technologies, such as 4G cellular base stations and WiFi Access Points, and by a diffusion of devices not limited to traditional user terminals. Indeed, many smart objects, such as domestic appliances, cameras, monitoring sensors, etc., are equipped with a wireless technology. Simultaneously, we attended a continuing advances in the development of flexible and programmable logic devices such as Field-Programmable Gate Arrays (FPGAs) that led Software Defined Radio (SDR) to be used to implement the signal processing functions of physical layer in wireless communication systems.

To support the ever growing volume of traffic, it becomes necessary to improve the WiFi technology and to use other wireless technologies, such as the emerging Long Range (LoRa) technology, which represents a critical example of wireless technology working in high-density networks. The aim of this thesis is to exploit physical layer flexibility to optimize wireless networks in various emerging scenarios with growing volume of traffic. Starting from the study and implementation of a generalized contention mechanism for WLAN technologies, we then focus on the analysis of LoRa technology, and finally we define a method for transmitter localization.

In the next Chapter, the focus will be on WiFi networks. Recently the original IEEE 802.11 standard has undergone an impressive improvement of the PHY layer capabilities. Breakthrough capacity improvements have been provided, by exploiting the latest PHY enhancements [2, 3], such as bandwidth aggregation, efficient modulation and coding schemes, advanced MIMO (up to 8 spatial streams in IEEE 802.11ac). As the air bit rate moves up to the Gbit/s range, the MAC and PHY protocol overhead, including the access contention procedure, do not scale accordingly. This is one of the leading motivations of the deep revision of the entire access paradigm that is currently maturing in the framework of the forthcoming IEEE 802.11ax standard, that promise of a throughput boost by a factor of four [3, 4, 5, 6]. Chapter 2 defines an innovative contention mechanism, called Repeated Contention (ReCo), that re-design completely the contention procedure of the framework of current IEEE 802.11 MAC protocol, in order to improve the performance of dense WiFi networks. Our work is supported by an experimental test-bed that realizes ReCo by means of simultaneous transmission and reception of short tones, which is feasible on top of programmable OFDM PHY layers. The throughput efficiency of ReCo is not sensitive to the number of contending stations. Efficiency and robustness is gained through the power of repeated contention rounds. We also apply the ReCo concept to the emerging IEEE 802.11ax standard, showing how it can boost performance of random access with respect to the current version of IEEE 802.11ax OFDMA Back-Off (OBO). Finally we propose a slightly modified version of ReCo, called ReCHO, which is designed to offer high throughput performance and robustness with respect to imperfect carrier sensing. The main idea is using narrow tones as signalling messages for performing channel access contentions and allowing the Access Point (AP) to echo these signals, in order to extend the sensing capabilities to all the stations associated to the AP so as to mitigate the hidden node problem.

In Chapter 3 the focus will shift to analyzes the LoRa technology, that has been conceived for Low Power Wide Area Networks (LPWANs), characterized by low data rate requirements per single device, large cells and heterogeneous application domains, which may lead to extremely high numbers of end devices (EDs) coexisting in the same cell. The goal of the work is to provide important guidelines for the design of LoRa networks. We analyze experimentally the link-level performance of LoRa and show that collisions between packets modulated with the same Spreading Factor (SF) usually lead to channel captures, while different spreading factors can indeed cause packet loss if the interference power is strong enough. Second, we model the effect of such findings to quantify the achievable capacity in a typical LoRa cell and that different criteria for deciding SF allocations within the cell may lead to significantly different results. We show that the use of power control and packet fragmentation can be detrimental more than beneficial in many deployment scenarios. Finally, we discuss the capacity improvements that can be achieved by increasing the density of LoRa gateways. This Chapter has been done in cooperation with the Ph.D. candidate Michele Gucciardo.

Finally, Chapter 4 defines a method for transmitter localization based on the well known Time Difference Of Arrival (TDOA, [7, 8]). In wireless communication systems the indoor position tracking is one of the key enabling technologies for various applications, such as to enable the beamforming technology in the IEEE 802.11ac and later standards networks, which could improve the performance in terms of energy saving and coverage range. The objective of this work is to provide a method for localization of the transmitter based on TDOA, drastically reducing errors by using appropriate signal processing techniques. We are working with very low-cost SDR receivers, such as RTL-SDR. Our method can work on a wide spectrum and with whatever type of radio signal, targets an accuracy of the order of the meter and it is based on a low-cost infrastructure, for a total cost less than a hundred dollars. We show how our methodology, which uses appropriate signal processing techniques for each step of the TDOA, provides excellent results on a simple test bed. We also show feasibility studies on possible future improvements, in order to integrate it later into the ElectroSense [9, 10] network, which is a crowd-sourcing initiative to collect and analyse spectrum data.

Chapter 2

Repeated Contention

2.1 Introduction

The MAC contention procedure in the standards IEEE 802.11 is based on random countdown of back-off time slots. While a back-off slot lasting 9 μs may have a negligible impact on the throughput of a WiFi interface operated at Mbit/s rates and carrying long frames (e.g., thousands of bytes), it rapidly becomes dominating as the bit rate scales up to Gbit/s and a lot of short messages are sent, e.g., generated by smartphone apps or sensors devices. The efficiency of the MAC protocol has been improved by allowing a station to transmit multiple data frames in a single channel access (aggregation of MAC Service or Protocol Data Units). Yet, the contention mechanism wastes a significant amount of capacity, due to countdown of idle back-off slots and to collisions. The consistent worsening of the MAC overhead impact has been highlighted for some time, e.g., see [11]. Throughput performance decay sensitively as the number of contending stations grows, thus making CSMA/CA unsuitable in crowded environments. Moreover, the Binary Exponential Back-off (BEB) affects adversely the jitter of service times [12]. This performance crippling carries over also to the currently proposed random access component of IEEE 802.11ax, namely the OFDMA Back-Off (OBO) procedure [5, 6].

In this Chapter, we define an innovative contention mechanism. We adopt the framework of current IEEE 802.11 MAC protocol, namely a contention phase aiming to elect a single winning station among the contending ones, with the winning station using the channel for transmitting its frames in the ensuing activity phase. We re-design completely the contention procedure of that framework and give an analytical model to predict its performance and provide guidelines for parameter dimensioning. We further extend our baseline design to address the emerging IEEE 802.11ax standard with multiple sub-channels. Finally, we add one more key idea, namely: echoing of the signalling tones by the AP. We re-define the contention rounds to allow the AP to echo the signalling tones sent by stations, in order to extend their sensing capabilities. We name this new enhanced version of ReCo as ReCHo. This new approach improves substantially access performance in imperfect channel sensing scenarios, where stations associated to a same AP can well be out of reciprocal visibility.

The driving principles of our project are: (i) simplicity of the design, that must be feasible based on current radio technology (e.g., without requiring full-duplex radio); (ii) high efficiency, through minimal contention overhead (time spent for contention and time wasted for collisions); and (iii) simple protocol configuration and robustness in the face of a varying number of contending stations. We propose a frequency domain repeated contention scheme that provides the following benefits:

- close to ideal saturation throughput;
- simple configuration of random access parameters, with no need of fine tuning or adaptive mechanisms. A fixed configuration is shown to provide excellent throughput performance over three orders of magnitude of the number of contending stations;
- full decoupling of the contention scheme from any detail of data transmission, with no need of transmitting information for coordinating the channel access within data frames;
- improvement of short-term fairness with respect to legacy IEEE 802.11, thanks to a complete regeneration of the contention mechanism after each channel access attempt;
- simple MAC protocol design;
- random access parameters do not need critical fine tuning or adaptive algorithms to be calibrated for best performance, e.g., as the number of contending stations varies.

The DCF contention time depends, mainly, on the number of contending stations and on the size of contention window (CW_{min} and CW_{max}). Although often, with $CW_{min} = 15$, the average contention time for DCF is comparable with ReCo (as we will see later), ReCo efficiency is higher because of the reduction of the collision rate. Collisions occur when the contention phase elects multiple winning station among the contending ones that, as we will see in § 2.3, will occur when multiple station choice (in the last round of the contention phase) the lowest level among the possible m.

As discussed in § 2.2, repeated contention round is not a new idea in itself, e.g., see [13, 14, 15, 16, 17] for some recent work. An early example of the idea of successive elimination rounds can be traced back to splitting algorithms, e.g., see [18, Ch. 4]. An innovative contribution of our work is setting up a mathematical framework that allows full understanding of the power of repeated contention and optimization of the contention parameters, including the probability distribution of back-off. Specifically, we prove an asymptotically tight bound of the collision probability, we give an algorithm to compute the back-off probability distribution that minimizes the number of surviving stations after one round and we exploit the algorithm to define a robust random access procedure (see Subsec. 2.4.2).

A second contribution is proving the practical feasibility of the frequency domain contention scheme in wireless nodes with limited complexity, which do not support full-duplex capabilities. In particular, we worked on WARP boards, for which a reference programmable implementation of IEEE 802.11g is available. We modified the PHY for enabling simultaneous transmission and reception of tones, and the MAC for supporting the ReCo logic. Our wireless node prototype allowed us to experimentally validate the efficiency and robustness of the proposed ReCo scheme.

A last contribution of the work is the definition of a new random access procedure, based on ReCo, for the emerging IEEE 802.11ax standard. We also highlight the impact of imperfect carrier sensing emerged in the real experiments is assessed also in a simulation environment in § 2.11. Then, we apply the ReCHo procedure to the emerging IEEE 802.11ax standard. We show through simulations the performance improvement brought by our two proposals with respect to the currently accepted OBO procedure, even in presence of imperfect channel sensing.

In the rest of this Chapter, after literature review in § 2.2 we define the proposed access procedure in § 2.3. We provide an analytical model and dimensioning criteria in § 2.4. Therefore, we illustrate a validation of ReCo in the following sections, in particular, in the order, the physical primitives for ReCo in § 2.5, Robustness and cover range of tone detection in § 2.6, an experimental validation in a real scenario in § 2.7 and an analysis of coexistence issues with legacy DCF in § 2.8. Application of the ReCo concept to IEEE 802.11ax is presented in § 2.10, after a brief overview of the emerging IEEE 802.11ax standard in § 2.9. Finally, after studying the impact of imperfect channel sensing in § 2.11, we describe the enhanced version ReCHo in § 2.12, showing its performance when it is applied to 802.11ax.

2.2 Related Work

Repeated contention round is not a new idea in itself. A general framework is set up by Zame et al. [19], aiming at defining a broad class of MAC protocols for distributed, sensing-based coordination protocols. The key idea is repeated *cycles* of contention that provide contending stations a history of channel observations, leading eventually to perfect coordination (hence no collisions) with high probability after a given number of contention cycles. Numerical examples provided in the paper point out that the number of contention slots required for a moderate number of stations (e.g., 32) can ramp up to several hundreds if not in the order of thousand. In terms of contention time, this corresponds to several tens of ms. As a result, the goodput is close to the theoretical maximum in settings where there is a limited number of stations and they have a large backlog to send. Moreover, it is not clear how adaptive the protocol could be as the rate of arrivals of new active stations or termination of previously active station grows up, since the time scale of convergence is much bigger than the time required to send a single frame.

Different mechanisms devised to provide constant contention times have been proposed [13, 14]. Here contending stations decide randomly to transmit a busy signal or not in a contention round. Stations that refrain from transmitting the busy signal, listen to the channel and drop out if they sense it busy. The Authors give a detailed approach to the optimization of the transmission probabilities in each round and several numerical examples. However, the optimization depends on the knowledge of the number of contending stations. Gowda et al. [14] illustrate the principle of repeated contention round to overcome the performance limitations of the traditional "linear" DCF contention, i.e., the one based on a single random extraction from a set of back-off values. Although recognizing the power of repeated contention, they end up defining a rather complicated access procedure. Moreover, they only state the repeated contention approach in time, as a generalization of standard DCF (akin ReCo in the time domain described in [20]). Another repeated round contention scheme is provided by Mao and Shen [15]. They target their contention scheme, named First Round-Bye (FRB), to handling different priority level flows. Repeated round supported by jamming is the basic means for priority management in a fully distributed way.

Frequency domain MAC procedure has been investigated in [16, 17], where up to two consecutive contentions are carried out by selecting random subcarriers rather than random back-off delays. The stations transmitting on the smallest frequency sub-carrier win the contention round. Frequency domain contention requires the capability of detecting other stations' signals while transmitting one's own sub-carrier. The feasibility of this operation is demonstrated experimentally in [17]. A Collision Detection (CD) scheme is defined in [21] for WiFi networks based on the full-duplex radio capability with standard CSMA access. The emphasis of the work is on optimization of the CD threshold. Full-duplex communication capabilities are exploited in a non-trivial way at MAC layer in [22]. The proposal is based on beacon (BCN) frames sent by the receiver during the data frame reception. The BCN frames act as acknowledgements that the reception is successful. BCN frames are also used to classify different network scenarios (collision region, transmitter-only region, receiver-only region; see Fig. 3 in [22]). The performance evaluation is focused on the resilience of the proposed MAC protocol to jamming attacks.

In [23] frequency domain contention is considered to define a random access reservation protocol. Reservation aims at electing a femtocell that transmits on a given channel, for mitigating the interference among nearby femtocells. A key point of that work is the assumption of a reliable feedback control channel from the receiver to the transmitter, which fits well the cellular paradigm. Moreover, the overhead of the reservation protocol is not a big issue in that context, given that the channel is required for intense, non-sporadic usage.

Frequency domain contention is exploited by Fayaz et al. [24] as well. They define a signaling protocol based on transmission and detection of tones, that aims at channelizing the available bandwidth so that non-interfering, concurrent links can operate simultaneously. The approach is suitable for infrastructured as well as ad-hoc networks. It requires the ability to transmit and detect sub-carrier tones at the same time. A synchronization scheme is discussed. A critical point is to fix the duration of the data phase so as to strike a good balance between potential low efficiency in case of long data phase duration and excessive weight of the signaling overhead for a short data phase.

Finally, ReCo allows any number of rounds to be configured, it also allows arbitrary probability of selection of signalling tones, parameters can be configured thanks to a mathematical description of the procedure that allows a deep insight. Moreover in ReCHo imperfect channel sensing can be handled even more successfully than with RTS/CTS and yet with substantially less overhead. Earlier examples of physical layer signalling, aiming at overhead reductions are [25, 26]. In [25], control messages like RTS, CTS and

ACK are encoded by using Correlatable Symbol Sequences (CSS). In [26] a PHY-based explicit signalling among the AP and the stations and frequency domain contention are proposed. The proposed scheme relies on additional control signals for solving the contention, whose duration is limited to one back-off slot. We believe ReCHo is actually simpler to implement than those proposals.

2.3 Repeated Contention Procedure

In this Section we introduce the baseline version of ReCo, that aims to define a single winning station at the end of the contention phase. Extensions to Multi-Winner (MW) ReCo, Multi-Contention (MC) ReCo and ReCHo, suitable for application to the emerging IEEE 802.11ax standard, are presented later, MW ReCo and MC ReCo in Sec. 2.10, ReCHo in Sec. 2.12.

As long as there are backlogged stations contending for the access to the channel, the channel time is divided into *access cycles*, made up of a *contention phase* and an *activity phase* (Fig. 2.1). During the contention phase, the time axis is divided into s consecutive *contention rounds*. All contending stations take part in the first round. Only stations winning round k qualify for the subsequent round $k+1$. Stations winning the last round are allowed to use the channel in the activity phase. The activity phase includes transmission of data frames (possibly multiple frames, thanks to aggregation), ACK and any other overhead. Whatever the outcome of the activity phase, namely either a successful transmission or a failure, all backlogged stations, take part in the next access cycle, by repeating exactly the same algorithm as performed in the previous cycles. No binary exponential back-off is used. Except of the count of re-transmission attempts, the entire access procedure is regenerated at each new access cycle.

Within each contention round, a backlogged station has to choose one 'level' among $m \geq 2$ possible choices. Regardless of the specific mechanism to implement the scheme, the key aspects for supporting repeated contention rounds are:

1. The m levels are strictly ordered. We label them with the integers of the set $\{1, \ldots, m\}$;

2. During one round every contending station can sense whether a level lower (or higher, based on the logic implemented) than its own has been chosen by any other station.

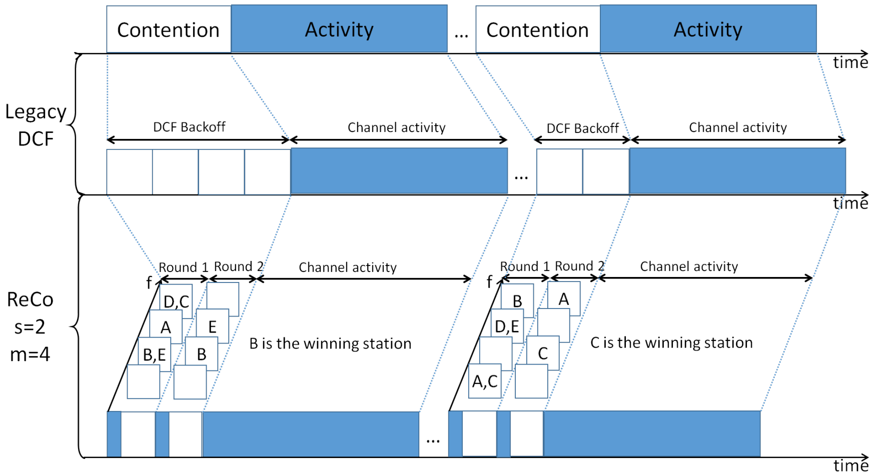


Figure 2.1. Channel access operations as a sequence of contention and activity phases. Comparison between Legacy IEEE 802.11 DCF (top) and ReCo (bottom).

We have decided to win the contention round for those who choose the lowest level so, the stations selecting the lowest level win the contention round and move forward to the next round. If we assume perfect channel sensing, there is *at least* one winner at the end of each round, and hence at the end of the contention phase. All losing stations drop out and wait for the next contention phase. The purpose is to elect a single winner.

Repeated contentions can be implemented very efficiently in the frequency domain. A set of m frequencies is defined, denoted with $\{f_1, \ldots, f_m\}$. As a matter of example, if the PHY layer is based on OFDM, the m frequencies can be identified with a subset of the sub-carriers. To improve the functioning of the channel sensing, it is necessary to select the tones properly, as explained in Sec. 2.5.

Based on extensive experiments with our ReCo implementation, we verified that tone detection works reliably by setting the contention round duration at least at two OFDM symbol times. To make a fair comparison with IEEE 802.11 DCF (Distributed Coordination Function), in the sequel we choose to set the contention round duration equal to the IEEE 802.11 DCF back-off slot, i.e., 9 μs, which is slightly more than two OFDM symbol times. Then, the time required to complete the contention procedure in one cycle is s back-off slots.

Let $q_i^{(r)}$, $i = 1, \ldots, m_r$ be the probability that frequency f_i is selected at round r, m_r being the number of frequencies used for contention in round r. The pseudo-code of the contention round algorithm for a station is given in

Algorithm 1 Pseudo-code of the contention phase algorithm.

```
$r = 0$;
$dropout = FALSE$
while $(r < s) \& (dropout == FALSE)$ do
    $r = r + 1$;
    $y =$ extract_random_number$(1, m_r, \bar{q}^{(r)})$;
    transmit_burst$(r, f_y)$;
    if $y > 1$ then
        $dropout =$ isbusy_channel$(r, [f_1, f_2, \ldots, f_{y-1}])$;
    end if
end while
```

Alg. 1. In the algorithm, we call the following functions:

- `extract_random_number`$(x, y, \bar{z})$: generates integer values between x and y according to the probabilities indicated in vector $\bar{z}$. Mathematically corresponds to $\min\{v \,|\, x \leq v \leq y \,,\, \sum_{j=x}^{v} z_j \geq \texttt{rand}\}$, with `rand` that generates samples uniformly distributed in [0,1].
- `transmit_burst`(r, f) transmits a busy tone on frequency f during contention round r.
- `isbusy_channel`$(r, [f_a, f_{a+1}, \ldots, f_b])$ checks if signal is detected on any one of the frequencies $f_a, f_{a+1}, \ldots, f_b$ during contention round r.

The algorithm states that a station contending in round r picks a frequency f_i at random, according to the probability distribution $q_i^{(r)}$ at round r, transmits that frequency and *at the same time* listens to check whether a frequency *lower* that f_i is being transmitted. Note that the contention round does not require the capability of decoding any frame. A station must simply check whether it receives a tone whose frequency is lower than its own choice.

Fig. 2.1 compares the legacy DCF contention process with ReCo in the frequency domain with $s = 2$ and $m = 4$, under the assumption that each contention round lasts exactly one DCF back-off slot. The boxes along the frequency dimension represent the tones used for the contention round. From the example, it is evident that the contention overhead in each contention round is variable for DCF and fixed for ReCo. Fig. 2.1 shows two consecutive contention phases, where five stations are contending, each marked by a letter. After the first contention round, two stations survive and are admitted to the second round, after which only one station survives (station B in the first contention phase, and station C in the second contention phase).

2.4 Analytical model

Let n be the number of contending stations with non-empty transmission queues at the beginning of the contention phase, also called backlogged stations. Let also s be the number of rounds and m be the number of levels (e.g., frequencies).

Let us focus on a single round. Let q_i denote the probability that a station picks level i, $i = 1, \dots, m$. Let also $G_i = \sum_{j=i}^{m} q_j$ be the corresponding Complementary Cumulative Distribution Function (CCDF). The probability $P_{k,h}$ that h stations survive after a single contention round, given that k stations are contending at the beginning of that round, is

$$P_{k,h} = \begin{cases} \sum_{i=1}^{m-1} \binom{k}{h} q_i^h G_{i+1}^{k-h} & h = 1, \dots, k-1 \\ \sum_{i=1}^{m} q_i^k & h = k. \end{cases} \tag{2.1}$$

We can form the $n \times n$ matrix $\mathbf{P}$ whose k-th row entries are $P_{k,h}$, for $h = 1, \dots, k$, and 0 for $h = k+1, \dots, n$ ($k = 1, \dots, n$). $\mathbf{P}$ is the one-step transition probability matrix of a Markov chain $\mathcal{X}$ on the state space $\{1,2, \dots, n\}$ with an absorbing state at 1. The state probability vector at time t is denoted with $\mathbf{x}(t)$, $t \geq 0$, where $x_i(t) = \mathcal{P}(\mathcal{X}(t) = i), i = 1, \dots, n$. It is $\mathbf{x}(0) = [0 \dots 0\ 1]$, i.e., at the initial time $t = 0$ the Markov chain is in state $\mathcal{X} = n$ with probability 1. The number of stations that survive s contention rounds (winning stations) is $\mathcal{W} = \mathcal{X}(s)$. We have a success after the completion of s rounds with probability $\mathcal{P}(\mathcal{W} = 1)$. The probability distribution of $\mathcal{W}$ can be calculated as $\mathcal{P}(\mathcal{W} = h) = x_h(s)$, $h = 1, \dots, n$, with $\mathbf{x}(s) = \mathbf{x}(0) \prod_{r=1}^{s} \mathbf{P}_r$, where $\mathbf{P}_r$ is made up by using the probability distribution $q_i^{(r)}$, $i = 1, \dots, m_r$, for round r ($r = 1, \dots, s$).

In the following we address first the case where the same probability distribution $\{q_j\}_{1 \leq j \leq m}$ is used in all s rounds (homogeneous case). Then we generalize the analysis to the heterogeneous case.

2.4.1 Homogenous repeated contention

In the homogeneous case we have $\mathbf{x}(s) = \mathbf{x}(0)\mathbf{P}^s$. Let $\mathbf{Q}$ denote the square matrix obtained by taking the last $n-1$ rows and columns of $\mathbf{P}$. $\mathbf{Q}$ is the one-step transition probability matrix of the transient states of $\mathcal{X}(t)$.

The collision probability p_c can be expressed as $p_c(s) = \mathcal{P}(\mathcal{W} > 1) = \mathbf{e}_1 \mathbf{Q}^s \mathbf{e}$, for $s \geq 1$; $\mathbf{e}$ is a column vector of ones of size $n-1$, $\mathbf{e}_1$ is a row vector of size $n-1$ whose entries are $e_1(j) = 0$ for $j \neq n-1$ and $e_1(n-1) = 1$.

The matrix $\mathbf{Q}$ is lower triangular, with diagonal elements given by P_{kk} in eq. (2.1) for $k = 2, \ldots, n$. Hence, its dominant eigenvalue is $\eta \equiv Q_{11} = \sum_{i=1}^{m} q_i^2$. Since $\mathbf{Q}$ is also a non-negative matrix, the left and right eigenvectors $\mathbf{v}$ and $\mathbf{u}$ associated to η are positive. Then, the asymptotic behavior of the collision probability as $s \to \infty$ can be written as $p_c(s) \sim \kappa\eta^s$, where $\kappa = \mathbf{e}_1\mathbf{uve}$.

We can state this result as follows: the collision probability decays geometrically as the number of rounds s grows, with a decay rate $\eta = \sum_{i=1}^{m} q_i^2$. Note that η is *minimized* for $q_i = 1/m$, $i = 1, \ldots, m$, i.e., when the level selection probability distribution is *uniform*. In that case it is possible to find closed forms for the dominant eigenvalue and associated eigenvectors of $\mathbf{Q}$. It is $\eta = 1/m$, $\mathbf{v} = [1\ 0\ \ldots\ 0]$ and $\mathbf{u} = [2\ 3\ \ldots\ n]^T/2$. Hence the asymptotic expansion of the collision probability is $p_c(s) \sim n/(2m^s)$ as $s \to \infty$. We summarize this in the following.

Property *The collision probability $p_c(s)$ behaves as $p_c(s) \sim \kappa\eta^s$ for $s \to \infty$, with $\eta \geq 1/m$. The equality holds if and only if $q_i = 1/m, i = 1, \ldots, m$. Then, it is, it is $p_c(s) \sim \frac{n}{2m^s}$ as $s \to \infty$.*

Since it is $\mathbf{e} = [1\ 1 \ldots 1] \leq [2\ 3 \ldots n]/2 = \mathbf{u}$, we have also:

$$p_c(s) = \mathbf{e}_1\mathbf{Q}^s\mathbf{e} \leq \mathbf{e}_1\mathbf{Q}^s\mathbf{u} = \frac{1}{m^s}\mathbf{e}_1\mathbf{u} = \frac{n}{2m^s} \tag{2.2}$$

since all involved vectors and matrices are made up of non-negative entries. The inequalities are entry-wise. In view of the Property stated above and of eq. (2.2), we can provide an asymptotically tight upper bound for the collision probability in case of uniform probability distribution $\{q_i\}_{1\text{ł}i\leq m}$:

$$\hat{p}_c(s) = \min\left\{1, \frac{n}{2m^s}\right\}, \qquad s \geq 1 \tag{2.3}$$

The accuracy of the upper bound is discussed in [20].

2.4.2 Non-homogeneous repeated contention

Different probability distributions can be used at each round of ReCo. We have seen that a uniform probability distribution results in a collision probability that decays geometrically with the number of rounds and is proportional to the number n of stations. We could then design the probability distribution of the first round so as to reduce the number of stations surviving the first round as low as possible.

Let $\overline{S}(n)$ be the mean number of stations surviving one round when n station take part in the round. We have:

$$\overline{S}(n) = n \sum_{i=1}^{m} q_i G_i^{n-1} \tag{2.4}$$

We will see that *the probability distribution* $\{q_k^*\}_{k=1,\ldots,m}$ *that minimizes* $\overline{S}(n)$ *makes it very close to 1.* However, the optimal probability distribution is a function of n. We will see also that it is possible to construct a heuristic probability distributions $\{\tilde{q}_k\}_{k=1,\ldots,m}$, *independent of* n, yielding values of $\overline{S}(n)$ close to the minimum $\overline{S}^*(n)$. After a first round with such a probability distribution, few stations survive. At that point, few more rounds with a plain uniform probability distribution will attain a low collision probability. Let us determine an algorithm to calculate $\{q_k^*\}_{k=1,\ldots,m}$. It is the probability distribution that minimizes (2.4) for any given $n \geq 2$, under the constraint $\sum_{i=1}^{m} q_i = 1$. Using Lagrange multipliers, we have to minimize the function

$$f(\mathbf{q}) = n \sum_{i=1}^{m} q_i G_i^{n-1} - \Lambda \sum_{i=1}^{m} q_i \tag{2.5}$$

Imposing that the gradient be null, we obtain the m equations $nG_j^{n-1} + n(n-1)\sum_{i=1}^{j} q_i G_i^{n-2} - \Lambda = 0$ in the $m+1$ unknowns Λ and q_i, $i = 1, \ldots, m$: Taking differences, we find

$$G_j^{n-1} - G_{j-1}^{n-1} + (n-1) q_j G_j^{n-2} = 0 \tag{2.6}$$

for $j = 2, \ldots, m$. This can be re-arranged as follows:

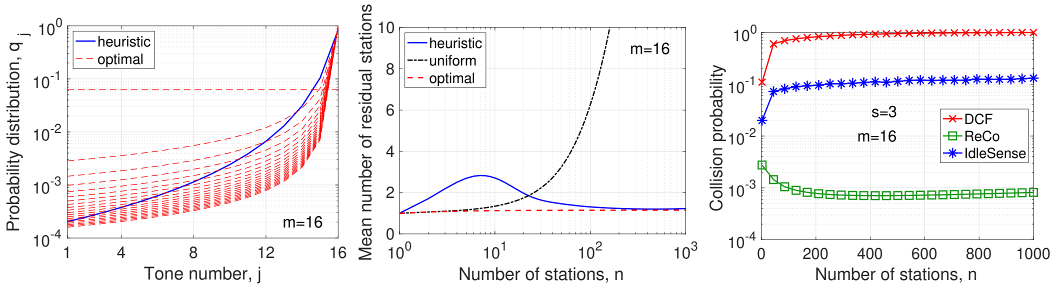
$$1 + (n-1)\frac{q_j}{G_j} = \left(\frac{G_{j-1}}{G_j}\right)^{n-1} = \frac{1}{(1 - q_{j-1}/G_{j-1})^{n-1}} \tag{2.7}$$

holding for $j = 2, \ldots, m$. Letting $z_j \equiv 1 - q_j/G_j$, we find

$$z_{j-1} = \frac{1}{[n - (n-1)z_j]^{\frac{1}{n-1}}} \tag{2.8}$$

for $j = m, m-1, \ldots, 2$, starting with $z_m = 1 - q_m/G_m = 0$. Once the z_j's have been calculated, the optimal probability distribution is found by

$$q_j^* = \begin{cases} 1 - z_1 & j = 1 \\ (1 - z_j)\left(1 - \sum_{i=1}^{j-1} q_i^*\right) & j = 2, \ldots, m \end{cases} \tag{2.9}$$



(a) (b) (c)

Figure 2.2. Subfigure (a) shows the probability distribution q_k: optimal distributions for $n = 1, \ldots, 1000$ (red dashed lines) and heuristic distribution (solid blue line), for $m = 16$. Subfigure (b) shows the mean number of stations surviving one round, for $m = 16$. Subfigure (c) shows the collision probability of DCF, Idle Sense and ReCo as a function of the number of stations n, when $m = 16$

Substituting back the optimal solution into the expression of the mean number of surviving stations, we find $\overline{S}^*(n) = 1 + (n-1)q_1^*$. It can be shown that

$$\overline{S}^*(n) \approx 1 + \overbrace{\log(1+\log(1+\cdots+\log(1+}^{m-2 \text{ times}}\log n))) \tag{2.10}$$

as $n \to \infty$. This explains why $\overline{S}^*(n)$ is so close to 1, even for large values of n.

A sample of optimal probability distributions $\{q_k^*\}_{1\le k\le m}$ for some values of n between 2 and 1000 is plotted in Fig. 2.2(a) for $m = 16$ (dashed red lines). We have found that a power-law heuristic probability distribution $\{\tilde{q}_k\}_{1\le k\le m}$, "interpolates" the family of optimal probability distributions; so we let

$$\tilde{q}_k = \frac{1/(m+1-k)^\alpha}{1+1/2^\alpha+\cdots+1/m^\alpha}, \qquad k = 1, \ldots, m. \tag{2.11}$$

The solid line curve of Fig. 2.2(a) is obtained by setting $\alpha = 3$ in (2.11).

The function $\overline{S}(n)$ is plotted in Fig. 2.2(b) for $m = 16$. We compare the values of $\overline{S}(n)$ obtained with: (i) the optimal probability distributions (one for each value of n; red dashed line); (ii) the heuristic probability distribution $\{\tilde{q}_k\}_{1\le k\le m}$ (blue solid line); and (iii) the uniform probability distribution (black dash-dot line).

The optimal probability distributions make $\overline{S}^*(n)$ very small, close to 1, for all n from 1 up to 1000 (high values could be interesting for IoT scenarios). The heuristic probability distribution is quite successful in cutting down to very low levels the number of stations surviving after one round. For low values of n (below 10), the uniform probability distribution is almost optimal, unlike the heuristic one. On the contrary, for large n, the uniform probability distribution fails to reduce the number of surviving stations substantially. Therefore, we expect good results by using the heuristic pdf in the first round (so that few stations survive, less than 3 on average over the range $1 \leq n \leq 1000$), then letting a uniform pdf deal with such a small number of surviving stations in the ensuing rounds.

The collision probability resulting from three rounds, the first of which uses the heuristic pdf in eq. (2.11) and the two subsequent ones use the uniform pdf, is shown in Fig. 2.2(c) for $m = 16$. We compare DCF (cross markers), ReCo (square markers) and Idle Sense (asterisks), an adaptive algorithm to adjust the transmission probability of DCF in an optimal way [27, 28]. Idle Sense has been selected since it achieves the best performance among the variants of IEEE 802.11 DCF. It has been simulated, by implementing the refined algorithm described in [28].

While Idle Sense is quite effective in maintaining the collision probability at a constant level as n varies, ReCo outperforms both DCF and Idle Sense by at least one order of magnitude. Results improve fast as m grows. Note that Idle Sense adaptation algorithm requires parameter values depending on the specific MAC PDU format, air bit rate, inter-frame space sizes. On the contrary, the optimization of the probability distributions used in ReCo contention are completely independent of any detail of data transmission.

2.4.3 Saturation throughput and short-term fairness

Fig. 2.1 shows the channel time evolution as a sequence of contention and activity phases. The duration of activity phases is the sum of two contributions: (i) overhead time T_{oh} accounting for PHY/MAC overhead and inter-frame spacings; (ii) payload transmission time.

Let U_i be a random variable representing the time the frame payload takes to be transmitted by the i-th station. If stations are statistically equivalent, $U_i \sim U$ for all i. It is $U = L/R$, R being the air bit rate of the MAC interface and L the MAC PDU payload length. Both quantities take a discrete spectrum of values, so that we model U as a discrete random variable. Let $U \in \{a_1, a_2, \ldots, a_\ell\}$ with $a_1 \leq a_2 \leq \cdots \leq a_\ell$, and $Q_j = \mathcal{P}(U \leq a_j)$ for $j = 1, \ldots, \ell$. For notation convenience we set also $Q_0 = 0$.

The activity time A in case r stations transmitting results to be $A = T_{oh} + \max\{U_1, U_2, \ldots, U_r\}$ for $r \geq 1$. By the independence assumption, the payload times U_i are mutually independent, so it is $\mathcal{P}(\max\{U_1, \ldots, U_r\} \leq a_j) = Q_j^r$ ($j = 1, \ldots, \ell$), and $\mathrm{E}[\max\{U_1, \ldots, U_r\}] = \sum_{j=1}^{\ell} a_j (Q_j^r - Q_{j-1}^r)$, for $r \geq 1$. For

$r = 1$ we find the mean $\mathrm{E}[U] = \sum_{j=1}^{\ell} a_j(Q_j - Q_{j-1})$.

We evaluate the saturation throughput ρ for n stations continuously backlogged. By considering that the end of each access cycle is a regeneration instant for ReCo, we can express the normalized saturation throughput as the ratio of the mean time spent in successful payload transmission and the average duration of the access cycle:

$$\rho_{ReCo} = \frac{v_1 \mathrm{E}[U]}{s\delta + \sum_{h=1}^{n} v_h \sum_{j=1}^{\ell} (T_{oh} + a_j)(Q_j^h - Q_{j-1}^h)} \tag{2.12}$$

where δ is the back-off slot time, $v_h \equiv v_h(s,m,n) = \mathcal{P}(\mathcal{W} = h)$, for $h = 1,\ldots,n$.

As reference comparison terms, we consider the throughput achievable under legacy DCF and under perfect scheduling. For the legacy DCF, the normalized throughput ρ_{DCF} can be found as a simple generalization of the model proposed in [29][30]:

$$\rho_{DCF} = \frac{n\tau(1-\tau)^{n-1}\mathrm{E}[U]}{\delta(1-\tau)^n + \sum_{j=1}^{\ell} (T_{oh} + a_j)(Y_j - Y_{j-1})} \tag{2.13}$$

where τ is the transmission probability in a back-off slot, and $Y_j = (1 - \tau + \tau Q_j)^n$, $j = 0,1,\ldots,\ell$.

The probability τ can be computed, given the number n of stations, the DCF maximum retry parameter, M, and the contention window sizes, W_i, $i = 0,1,\ldots,M$, by solving a non-linear equation system (see [29][30]), namely

$$\tau = \frac{1 + p + p^2 + \cdots + p^M}{\beta_0 + \beta_1 p + \beta_2 p^2 + \cdots + \beta_M p^M} \qquad p = 1 - (1-\tau)^{n-1} \tag{2.14}$$

with $\beta_i = (W_i + 1)/2$ for $i = 0,1,\ldots,M$.

In case of ideal scheduling, no contention and back-off are required. Each access cycle consists of a successful transmission. Then

$$\rho_{ideal} = \frac{\mathrm{E}[U]}{T_{oh} + \mathrm{E}[U]} \tag{2.15}$$

As for the fairness, all considered random access protocols, DCF, ReCo and Idle Sense, are long-term fair, in the sense that they guarantee the same average number of access opportunities to each station, over a long time.

In the short-term, i.e., over time scales comparable with the packet transmission time, things are quite different. Let us consider the access time Θ, defined as the time elapsing since when a packet becomes the head-of-line in a station until when it is transmitted with success or it exhausts all retransmission attempts. A measure of short-term fairness is offered by the Coefficient Of Variation (COV) of the access time C_Θ, defined as the ratio between the standard deviation of Θ and the mean of Θ.

We define also the access gap J. Let $P, P_1, P_2, \ldots, P_k, P'$ a sequence of successfully transmitted packets, with P and P' belonging to a same station A and all other packets belonging to stations other than A. Then we say that the access gap is $J = k+1$. With n identical stations, in case of deterministic round-robin, it would be $J = n$. Since the considered protocols are long-term fair, they all give $\mathrm{E}[J] = n$. The COV of J, C_J, gives a measure of short-term unfairness.

Since access cycles starting times are renewal points for a given station, it can be verified that $\mathcal{P}(J = k) = (1 - 1/n)^{k-1}(1/n)$, $k \geq 1$. Then, it is $C_J = \sqrt{1 - 1/n}$. A simple upper bound of C_Θ can be found by bounding the variance of the intervals where collisions occur with the variance of intervals with successful transmissions. This is an upper bound since payload times U vary within a finite interval. Then, it can be found that $C_\Theta = \sqrt{1 - p_s/n + p_s/nC_U^2}$, p_s being the probability of success. This expression is exact if U is fixed, hence $C_U = 0$.

As for the DCF, the derivation of the access time variance is quite involved. It can be found in [31]. The variance of J for DCF can be found following a similar reasoning. Numerical evaluation of this expression highlights that the largest contribution by far is due to the binary exponential back-off algorithm. The variability of the payload times has a minor impact. Finally, C_J and C_Θ for Idle Sense are obtained via simulations.

We give numerical examples. We evaluate the saturation throughput for the IEEE 802.11ac PHY parameters. The back-off slot duration is $\delta = 9$ μs. We consider a 40 MHz channel, the air bit rate set to 200 $Mbps$ (1 spatial stream, 256-QAM with code rate 5/6), and $T_{oh} = 162.9$ μs. Payload lengths are uniformly distributed over the set {80,1500,9000,11454} bytes, taking also into account the aggregation of 4 MPDU. For the standard IEEE 802.11 DCF the contention window sizes are $W_i = \min(16 \cdot 2^i, 1024)$ for $i = 0, \ldots, 7$. As for ReCo, we set $s = 3$ and assume that we can use 8 tones in 20 MHz bandwidth. Consistently, we set $m = 16$ for IEEE 802.11ac. The duration of each contention round is identified with the back-off slot duration. Tab. 2.6 summarizes the analysis parameters and gives their numerical values.

Fig. 2.3(a) shows the normalized throughput ρ as a function of the number of contending stations n for the ideal MAC (triangle markers), ReCo (square

Table 2.1. Numerical values of parameters used in the analysis.

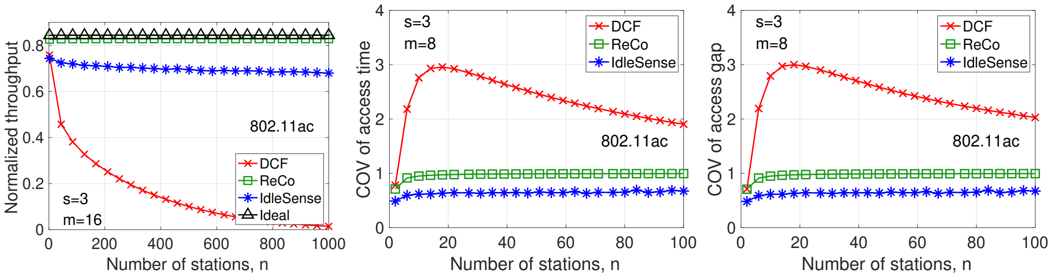
δ	Back-off slot	9 μs
W_{ch}	Radio channel bandwidth	40 MHz
R	Bit rate	200 $Mbps$
T_{oh}	PHY/MAC overhead time plus inter-frame spacings	162.9 μs
L	MAC PDU payload	80,1500,9000 or 11454 bytes
CW_{min}	Minimum contention window	15
CW_{max}	Maximum contention window	1023
s	Number of rounds for ReCo	3
m	Number of tones for ReCo	8 or 16
	ReCo contention phase time	$s \cdot \delta$

markers), IEEE 802.11 DCF ('x' markers) and Idle Sense (asterisk markers).

The most relevant outcome is that ReCo exhibits close-to-ideal performance results, and that the achieved throughput is almost insensitive to the number of contending stations in the range between 1 and 1000. While Idle Sense exhibits excellent performance, except at small n levels, ReCo is definitely superior. We observe that ReCo throughput performance are achieved with a *fixed* parameter configuration and a relatively small value of m. There is no need of implementing an estimator of the number of contending stations as in Idle Sense. This is a critical point whenever the offered traffic is volatile and intermittent, so that the number of contending stations varies quickly over time, possibly by large amounts. ReCo does not suffer the offered traffic variability, given that a *static* parameter setting is essentially optimal for n ranging over three orders of magnitude.

As for fairness, Fig. 2.3(b) and Fig. 2.3(c) illustrates the short term fairness of an IEEE 802.11ac network with n stations in saturations. Payload times are fixed. We have tested numerically also variable payload times, with various distribution for packet sizes and air bit rates, checking the the COV of access time and access gap exhibit little modifications with respect to the values plotted in figures.

Fig. 2.3(b) compares the COV of Θ for DCF, ReCo and Idle Sense. It is apparent that DCF introduces a sensitive variability of the access time, with a COV as high as 3. ReCo offers access times that are geometric, hence with COV close to 1. Idle Sense gives an even smaller COV, i.e., a more regular service. The little increase of variability induced by ReCo with respect to



(a) (b) (c)

Figure 2.3. Normalized throughput vs. stations number: comparison among ideal (no collisions), ReCo, IdleSense and IEEE 802.11 DCF, in case of IEEE 802.11ac with $m = 16$ (a). Short-term fairness for a IEEE 802.11ac network with n stations in saturation. Parameters for ReCo are $m = 8$ and $s = 3$. (b) Coefficient Of Variation (COV) of access time as a function of n. (c) COV of the access gap J as a function of n.

Idle Sense is the price to pay for the simplicity of ReCo. ReCo algorithm is essentially state-less. The state of the station is renewed at each contention phase. As opposed, Idle Sense maintains a number of variables to track the optimal value of the contention window. By introducing a minimum added complexity to ReCo, in the order of that of Idle Sense, it could be possible to reduce the variability of the access time by temporarily reducing the priority of a station that has just transmitted with success. Same remarks apply to the COV of the access gap J (Fig. 2.3(c)). The statistics of J do not depend in any way on the statistics of the payload times and on the values of the overhead and other peculiarities of the different IEEE 802.11 amendments (except of the parameters of the adaptive algorithm of Idle Sense).

2.5 Physical primitives for ReCo implementation

In order to use the ReCo mechanism in a real wireless network, we implemented a prototype of wireless node able to execute the ReCo contention mechanism in the frequency domain. Prototypes of wireless nodes exploiting tone-based contention mechanisms have been mainly built on top of software defined radio (SDR) platforms, such as the GNU Radio/USRP platform [17, 32]. In our case, we decided to work on the WARP SDR platform, which includes an FPGA, in order to exploit both PHY layer programmability (to

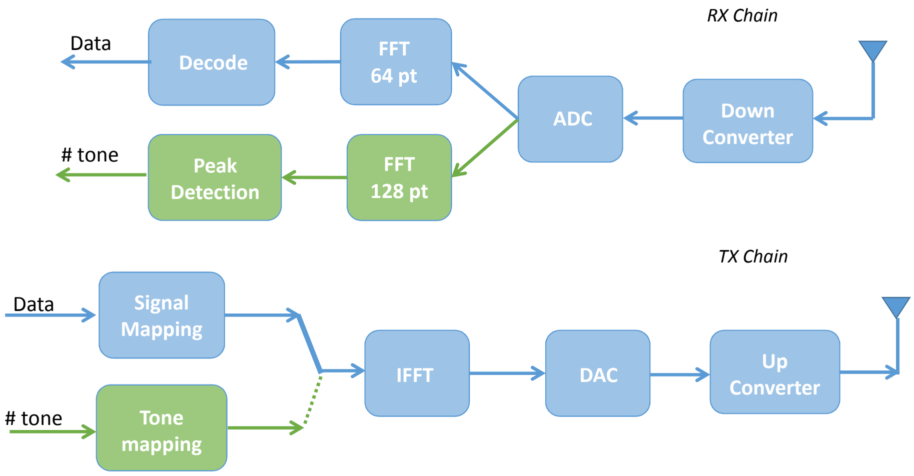


Figure 2.4. Modifications to the WARP reference design for IEEE 802.11g PHY enabling tone transmissions and receptions.

implement tone transmission and reception mechanisms) and MAC layer programmability (to implement the ReCo logic) inside the platform. Moreover, a reference implementation of a complete IEEE 802.11g MAC/PHY stack is available for the WARP platform.

In ReCo, contentions are resolved by identifying the stations transmitting the tone at the lowest frequency. This implies the capability of transmitting and receiving tones simultaneously. While full duplex radio obviously enables such a capability, by means of advanced digital and analog cancellation mechanisms [33], the same requirement can be satisfied by a much simpler PHY layer based on the integration of two independent transceivers in one node. Indeed, as discussed in what follows, recognizing tones is much easier than recognizing modulated signals, especially if tones are opportunistically spaced.

Fig. 2.4 shows the high-level architecture of the original IEEE 802.11g PHY reference design (blue blocks) and the modifications implemented for supporting tone contentions (green blocks). In principle, multi-carrier modulation can be easily adapted for transmitting and detecting tones. On the transmitter side, it is enough to null all the available sub-carriers apart one selected in a list of possible tones. On the reception side, it is enough to sample the reception power at each sub-carrier and compare such a power with a threshold, as we will see shortly.

We considered the possibility to rely on IFFT/FFT blocks working on a number of sub-carriers higher than 64. By increasing the tone duration (or

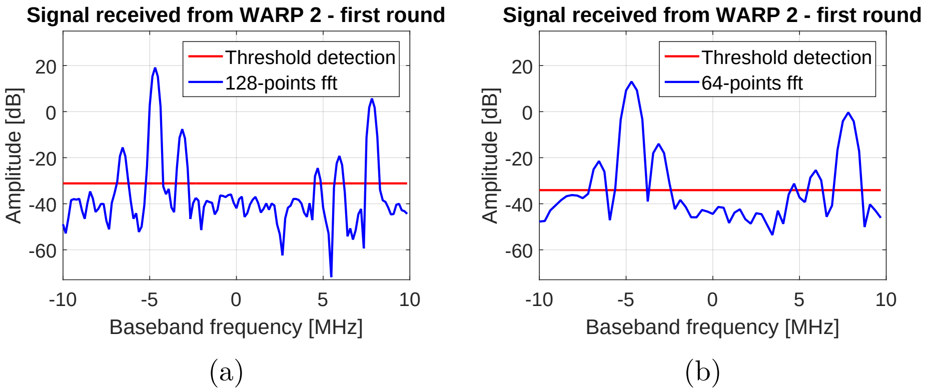
Table 2.2. Output of algorithm that provides the best sub-set of all available sub-carriers, with $m = 11$ and 20 MHz bandwidth

sub-carrier index	38	44	50	55	61	3	9	15	20	26	32
frequency [MHz]	-8.44	-6.56	-4.69	-3.12	-1.25	0.62	2.50	4.37	5.94	7.81	9.68

equivalently the number of IFFT/FFT samples generated for a given sampling interval), detection of tone transmissions is more robust to channel attenuation and interfering signals. The contention overhead grows proportionally to the time required for each tone transmission.

Apart from the capability of transmitting and receiving tones, a crucial aspect is performing these operations in parallel. The tone used for transmission will obviously create a self-interference with the receiver chain, which will result in high power values at the same frequency used for transmission. Moreover, because of the non-linearity of the transceiver, the transmitted tone will create some (attenuated) spurious signals at other frequencies (namely, at the mirror frequency and at twice the frequency of the tone). Self-interference signals can prevent the detection of tones transmitted by other nodes at the same frequency, but this capability is not required by the ReCo scheme. On the other hand, spurious signals could affect the ReCo contention mechanism because they can be erroneously considered as valid tones transmitted by contending nodes. The probability of this event can be reduced to zero (or minimized) if the frequencies at which the spurious signals appear do not overlap (or have a small overlapping probability) with the ones chosen as contention tones. To this purpose, we have implemented a simple algorithm that provides the best sub-set of all available sub-carriers as contention tones, in order to maximize the distance between spurious signals and contention tones. Table 2.2 shows a list of 11 sub-carriers that the algorithm has calculated among a set of 64 available ones in a 20 MHz bandwidth. The ReCo scheme has been configured with $m = 11$ for comparing the contention results with the time version ReCo scheme presented in [20].

Fig. 2.5 shows the FFT samples received by a wireless node, while transmitting a tone at -4.69 MHz with respect to the central frequency, lasting 128 (Fig. 2.5(a)) or 64 (Fig. 2.5(b)) samples. From the figure it is evident that it is possible to perform parallel transmissions and receptions of tones, despite of the self-interference generated by the transmitted tone. This interference results in the power peaks at -4.69 MHz, corresponding to the transmitted tone, and at 4.69 MHz and -9.38 MHz, corresponding to the spurious tones generated at the mirror frequency and twice the frequency of



(a) (b)

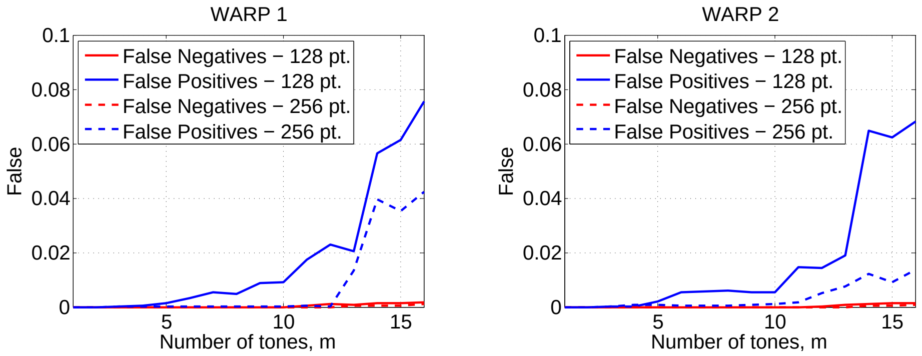
Figure 2.5. FFT samples received by a wireless node in a first contention round with 5 contending stations, with tones lasting 128 samples (a) or 64 samples (b). Shorter tones result in an increased width of the power peak lobs.

the tone. The second spurious tone is not relevant, being its power comparable with the noise. The narrow bandwidth of spurious tones and the opportunistic selection of contention tones allow the correct identification of other four tones transmitted at $-6.56\ MHz$, $-3.12\ MHz$, $5.94\ MHz$, and $7.81\ MHz$. Working on 128 (or 256) samples allows to improve the frequency resolution of peak detection, thus leading to a more robust identification of tones.

Tone detection is based on a simple comparison between the value of the FFT samples at the frequencies of the potential tones and a threshold value, which is tuned as a function of the background noise. The power of background noise is estimated by averaging the FFT samples at frequencies different from the ones in which tone transmissions or spurious signals can be detected. The threshold is then obtained by adding a margin to this noise level.

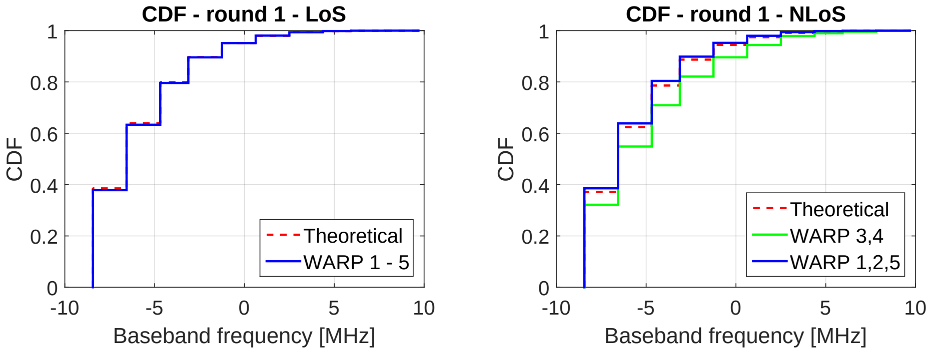
2.6 Tone detection - Robustness and coverage range

In order to assess the robustness of tone detection and the probability of erroneously considering an impulsive noise as a contention tone, we traced more than 60000 contentions between two nodes in Line-of-Sight (LoS) propagation conditions, as the number m of contention tones varies from 2 to 16. For each contention round we traced the indices of the tones randomly selected for transmission and the indices of the tones detected at each node.



(a) (b)

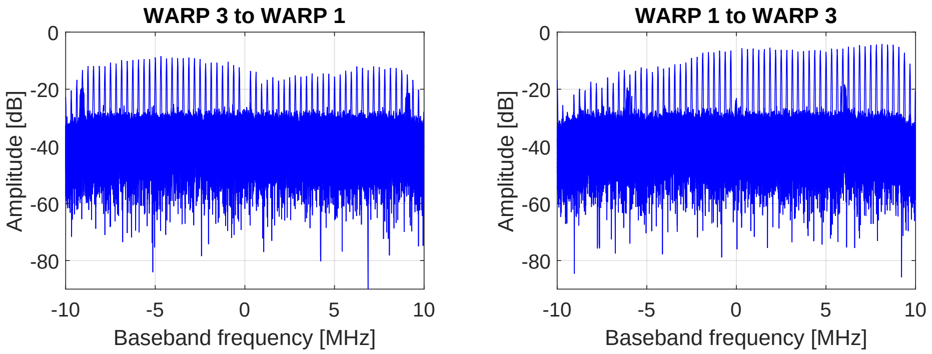
Figure 2.6. False detection made by the detection algorithm with variable threshold tuned as a function of the background noise at two WARP stations. WARP 1 (a) and WARP 2 (b).



(a) (b)

Figure 2.7. Cumulative Distribution Function of tone transmissions observed by five contending stations in Line Of Sight (a) and No Line Of Sight (b) propagation conditions among the stations.

Fig. 2.6 shows the detection errors in terms of false negatives, i.e., tones not detected at the frequency index selected by the contending station, and false positives, i.e., tones detected at a frequency index where no tone has been transmitted. From the figure, it is evident that false negatives are almost zero, while the probability of false positives can be relevant as the number of contention tones increases. However, the effects of false positives on the overall contention mechanism can be negligible. Indeed, a contention station can take a wrong decision of leaving the contention only if a false tone is detected at a frequency index lower than the one used for transmission. If the wrong decision implies that no station attempts a channel access, a new contention process can be restarted right after a DIFS time (with a small waste of channel time).



(a) (b)

Figure 2.8. Effects of selective fading on tone detection: channel response between station 3 and station 1 in two different link directions.

Fig. 2.7 shows the cumulative distribution of the tone detected at the lowest frequency, in a contention round with $m = 11$ and five contending stations. In Fig. 2.7(a), under LoS propagation, we can see that the curve almost coincides with the one resulting from ideal detection, i.e., with the CDF of the minimum of five variables uniformly distributed in a range of 11 possible values. In Fig. 2.7(b), when one node is moved to a different room and one node has been hidden by means of an obstacle, we found that distributions observed by each node do not coincide anymore. For example, we can clearly observe that stations 3 and 4 see a probability equal to 0.32 that the first contention round ends with a transmission on the first tone, while the other stations see a probability value equal to 0.38. Indeed, in some cases, nodes 3 and 4 are not able to hear the tones sent by the other stations. The possibility to hear tones also depends on the specific frequency selected by the far station for contention. In absence of a dominant propagation path, the channel can be selective in frequency and therefore the tone transmissions performed by the far station that can be detected by the other contending nodes are no more uniformly distributed among the available frequencies. These considerations are quantified in Fig. 2.8(a) and Fig. 2.8(b), where we visualize the channel response between station 3 and station 1 in two different link directions. Because tone attenuations strongly depend on the sub-carrier index and the channel is not symmetrical, the contention process is no more fair. For example, since transmitted tones by station 1 with low sub-carrier indexes are highly attenuated at the receiver of the station 3 (as depicted in Fig. 2.8(b)), the winning probability in the first contention round of the station 3 is higher.

Table 2.3. Detection of tones and data frames.

	Distance [m]	Noise [dBm]	RX Tone [#]	RSSI Tone [Average dBm]	RX Frame [#]	RSSI Frame [Average dBm]
Low TX Power	5	-94.362	1000	-70.490	0	-93.816
	10	-94.367	981	-84.140	0	-94.333
	15	-94.364	881	-86.901	0	-94.335
Medium TX Power	5	-94.372	1000	-55.122	1000	-88.129
	10	-94.347	1000	-70.330	407	-92.955
	15	-94.355	1000	-66.789	326	-93.154
High TX Power	5	-94.362	1000	-49.777	1000	-75.198
	10	-94.346	1000	-65.571	1000	-89.081
	15	-94.365	999	-64.210	1000	-88.391

Apart from the selective channel behavior, which can bias the results of the contention mechanism, a final aspect to be considered is the potential presence of hidden nodes, i.e., nodes which are not able to hear each other during contention but can transmit towards a common receiver. The occurrence of hidden nodes can be significantly mitigated by the fact that tone transmissions, which concentrate the transmission power in a very narrow band, have a detection range much higher than the one resulting for data transmission. Therefore, a station A, which does not hear data frames sent by a station B, that interferes with A's receiver, is generally able to hear tones sent by B. To demonstrate this phenomenon, we monitored the number of data frames and tones received by a reference station placed at different distances from another reference station transmitting at 0 dBm (low power), 11 dBm (medium power) or 22 dBm (high power). Table 2.3 summarizes the results, by showing that tones detection can be easier than data frame detection.

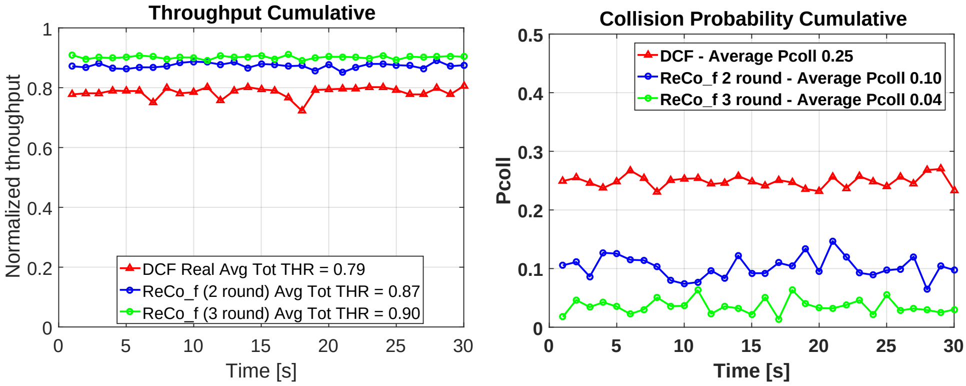
To mitigate the problem of hidden nodes, the ReCo mechanism can be extended, for example by using tones for coding additional information such as the duration of the data frame transmission to be used for virtual carrier sensing. ReCo can also be adapted by having the AP echo the lowest frequency tone it receives in a round, as we investigated in Sec. 2.12, where we defined a version of ReCo, called ReCHo. This mechanism realizes effectively an RTS/CTS signaling implemented through tones.

2.7 Experimental validation

We run several performance tests by limiting our observations to the scenario in which all the stations work in Line-of-Sight. In this context, tone contention is not degraded by selective fading, but only by errors due to the

Table 2.4. Numerical values of parameters used in the experiment.

δ	Back-off slot	$9\ \mu s$
W_{ch}	Radio channel bandwidth	$20\ MHz$
R	Bit rate	$6\ Mbps$
n	Number of stations (WARPs)	5
L	MAC PDU payload	1500 bytes
	SIFS and DIFS	$16\ \mu s$ and $34\ \mu s$
CW_{min}	Minimum contention window	15
CW_{max}	Maximum contention window	1023
s	Number of rounds for ReCo	2 or 3
m	Number of tones for ReCo	8
	ReCo contention phase time	$s \cdot \delta$

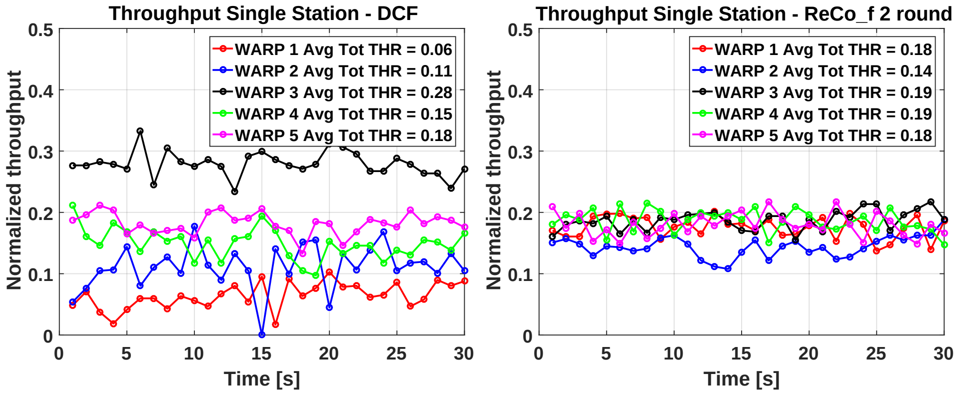


(a) (b)

Figure 2.9. Experimental throughput (a) and collision probability (b) results in case of legacy DCF (red curve), ReCo with 2 rounds (blue curve) and ReCo with 3 rounds (green curve) with 5 contending stations.

detection algorithm of the tones. Source rates have been configured for guaranteeing saturation conditions with data packets of 1500 *bytes* and a data transmission rate of 6 $Mbps$. The duration of each experiment has been set to 30 seconds. Tab. 2.4 summarizes the experimental parameters and gives their numerical values.

Fig. 2.9(a) shows the total normalized throughput achieved with 5 contending stations in case of legacy DCF and in case of ReCo with two ($s = 2$) or three ($s = 3$) contention rounds and a number of contention tones m equal to 11. Contention tones have been selected according to the list summarized in Table 2.2. From the figure, we see that the normalized throughput is 79% for DCF, 87% for ReCo with $s = 2$ and 90% for ReCo with $s = 3$.



(a) (b)

Figure 2.10. Per-station throughput results in case of legacy DCF (a) and ReCo with 2 round (b). COV of the access gap J: $C_J(legacyDCF) = 1.18$; $C_J(ReCo) = 0.91$.

Note that ReCo spends a fixed contention time for each channel access, equal to the product of the number of contention rounds and the duration of a contention tone (i.e. 2 or 3 times 9 μs). The DCF contention time depends on the number of contending stations and on the minimum contention window CW_{min}. Although for 5 stations and $CW_{min} = 15$ the average contention time for DCF is comparable with ReCo (about 3.1 back-off slots of 9 μs), ReCo efficiency is higher because of the reduction of the collision rate (as shown in Fig. 2.9(b)). Indeed, the average collision probability perceived with DCF is 0.25 (consistent with the well known Bianchi's result [29]), while such a value is reduced to 0.10 and 0.04 for ReCo with, respectively, $s = 2$ and $s = 3$.

Finally, as summarized in Fig. 2.10, we observe that ReCo improves the per-station fairness in comparison with DCF, because the average throughput perceived by each station exhibits a variability lower than DCF, even with two contention rounds. More into details, in our experiment we found that per-station normalized throughout achieved under legacy DCF varies in the range $[0.06, 0.28]$ between the worst and best performing station, achieve a COV of the access gap of 1.18, while for ReCo such a range is reduced to $[0.14 - 0.19]$ and the COV of the access gap is reduced to 0.91.

2.8 Analysis of coexistence issues with legacy DCF

Although the adoption of ReCo can significantly improve the efficiency of random contention, an important aspect to be addressed is related to coexistence with other legacy WLANs operating in the same channel. Consider for simplicity two cells only, each one operating with a different access protocol (ReCo and DCF). There are two main scenarios that can be considered: (i) *compatibility*: the two cells significantly overlap and most of the transmissions performed in one cell can be detected in the other one; (ii) *interference*: the two cells partially overlap and only tone transmissions (whose power density is higher than data frame) can be detected by the legacy nodes.

Compatibility. ReCo consecutive transmissions are separated by a DIFS time plus a few fixed-size contention rounds used for tone transmissions, during which legacy DCF stations detect the medium as busy. Therefore, in case of overlapping between nodes employing ReCo and DCF protocols, DCF legacy nodes will be indefinitely prevented from accessing the channel and therefore their throughput will be reduced down to zero. Being the number of orthogonal channels limited, the most suitable approach for tackling the compatibility issue is using TDMA (Time Division Multiple Access) between ReCo nodes and legacy DCF nodes. Different solutions can be envisioned. On one side, it is possible to activate ReCo only in a portion of the beacon interval, by announcing in the beacon frame a time interval open to ReCo nodes similarly to a contention-free periods used for polling. On the other side, rather than resuming the contention a DIFS interval after the ACK transmission, it is possible to adopt an extended AIFS space (e.g. a DIFS plus a number of backoff slot) for permitting the access of legacy stations between consecutive ReCo channel accesses. For example, a ReCo station could let go k idle back-off slots before starting tone contention. This way IEEE 802.11 stations that are competing for access could count down up to k back-off slots for each ReCo station access and eventually hit 0 and transmit. By modulating k one could trade-off the efficiency of the overall system with the access time granted to IEEE 802.11 stations. In both cases, different time shares between ReCo and DCF can be achieved, by modulating the fraction of time granted to ReCo access or the size of the AIFS time to be used by ReCo nodes. In other words, ReCo tone transmissions act similarly to trigger frames of IEEE 802.11ax in order to schedule the next channel access grant. Solutions for guaranteeing compatibility with legacy stations can be inspired by solutions envisioned for IEEE 802.11ax. Being the number of orthogonal channels limited, the most suitable approach for coexistence is using TDMA (Time Division Multiple Access) between ReCo nodes and legacy DCF nodes, for example exploiting inter-cell NAV as in the IEEE 802.11ax case [34]. We will briefly analyze the IEEE 802.11ax standard and a possible extension of ReCo to a Multi-User transmission system in the next sections.

Interference. When ReCo frames are not detected by the coexisting DCF nodes, there is no risk that DCF nodes are indefinitely prevented from accessing the channel. However, we cannot exclude that tone transmissions are received by DCF nodes or interfere with DCF frame transmissions because, as quantified in table 2.3, they can be detected at higher distances than frame transmissions. Note that interference generated by ReCo frames whose power in lower than the carrier sensing threshold is equivalent to the one due to any other coexisting DCF cell in similar position and therefore it is not analyzed as a ReCo coexistence issue.

Legacy DCF stations react to the detection of a tone, sent by a ReCo station, by activating the receiver chain and by freezing the back-off counter in case they are in back-off. Since a tone transmission lasts a few OFDM symbols, the back-off count-down is enlarged of a limited extra time (equal to the tone duration and an additional DIFS time), with a modest impact on the throughput. Moreover, in case of collisions, the narrow-band interference generated by the tone does not prevent the correct reception of a legacy frame, thanks to channel coding. To quantify such statements, we run an experiment in which a link at 6 Mbps between two commercial DCF cards is interfered by tone transmissions performed, at regular time intervals, by a close ReCo node, for which frame transmissions following the random contention have been disabled. For better enlightening the effect of the interference, the transmitter card is configured with a fixed contention window of 80 slots, corresponding to the time interval between two consecutive tones. Two monitoring USRP nodes are placed close to the DCF transmitter card and receiver card for estimating the interference power of the tones and the resulting SIR, as well as monitoring the inter-frame spaces between consecutive transmissions.

Table 2.5 summarizes the throughput results and the packet loss of the legacy link measured during the daily hours (i.e. in presence of other interference sources) with and without the tone transmissions for packet payload of 1470 byte. We set-up two interference experiments, with ReCo interference affecting the transmitter or the receiver only. A last experiment refers to the coexistence case, in which ReCo frame transmissions are detected by DCF nodes. In the first interference experiment, we can observe that tone transmissions marginally affects the throughput results. However, the back-off counter of the transmitter station experiences some extra freezes, as depicted in Fig. 2.11 in the cumulative distribution of the inter-frame spaces between consecutive transmissions. Because of the environment interference, the inter-frame spaces in absence of tone transmissions is not fixed to 80

INTERFERENCE AT THE TRANSMITTER SIDE		
	w/o ReCo	ReCo SIR 1.64dB
Throughput [Mbps]	μ =3.8, σ^2=0.0048	μ =3.6, σ^2=0.0171
Packet Loss [%]	μ =7.92, σ^2=2.21	μ =7.54, σ^2=2.36
INTERFERENCE AT THE RECEIVER SIDE		
	w/o ReCo	ReCo SIR -9.75dB
Throughput [Mbps]	μ =3.8, σ^2=0.0048	μ =3.78, σ^2=0.0021
Packet Loss [%]	μ =7.92, σ^2= 2.21	μ =7.88, σ^2=2.14
COMPATIBILITY SCENARIO		
	w/o ReCo	ReCo SIR -0.42dB
Throughput [Mbps]	μ =3.8,σ^2=0.0048	μ =0.0047, σ^2=3.46e-05
Packet Loss [%]	μ =7.92, σ^2=2.21	μ =0, σ^2=0

Table 2.5. Performance in term of average values (μ) and vairances (σ^2).

slots, but results in the cumulative distribution of the blue curve. When ReCo is active, the minimum space between transmissions grows of about 6 slots. More specifically, 4 slots are due to the DIFS interval and 2 slots are due to the tone duration (i.e. 2 slots are required to reset the receiving chain of the transmitter activated by the tone).

Also in the second interference experiment, when the interference acts on the receiver, despite the fact that during a frame transmission performed at 6 Mbps (lasting about 2ms), there are multiple tones transmissions spaced of 700 μs (about 78 slots), the throughput is marginally affected by the interference.

Finally, in the closing experiment we activated frame transmissions on the ReCo node and increased the interfering power. When DCF legacy cards detect ReCo frame transmissions, they are not able to access the channel anymore and the throughput is reduced down to zero.

2.9 IEEE 802.11ax Overview

The IEEE Standards Association (IEEE-SA) approved 802.11ax in March 2014 [35]. The scope of the IEEE 802.11ax amendment is to define innovative mechanisms at both the PHY and MAC layer for improving the the capacity of WLANs in scenarios with high density of nodes. The main innovation considered in this extension is the adoption of a PHY layer based

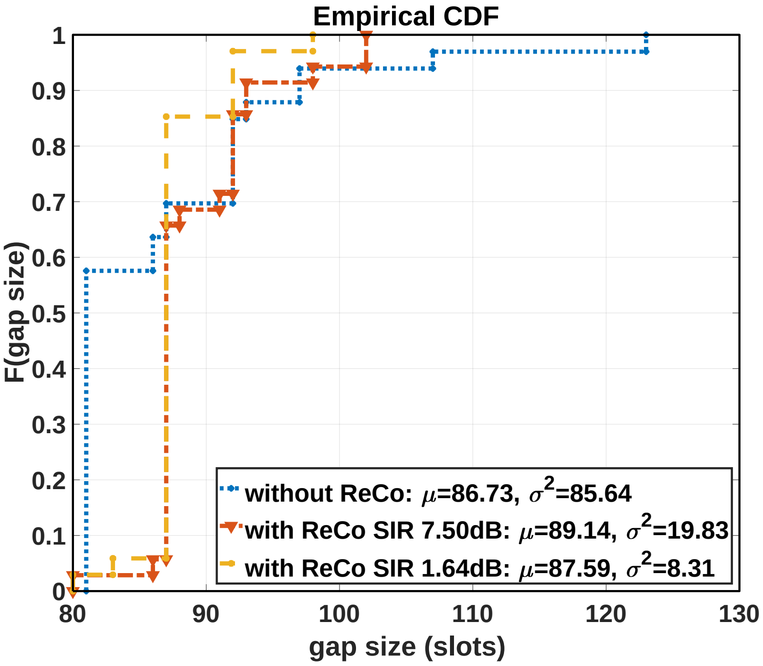
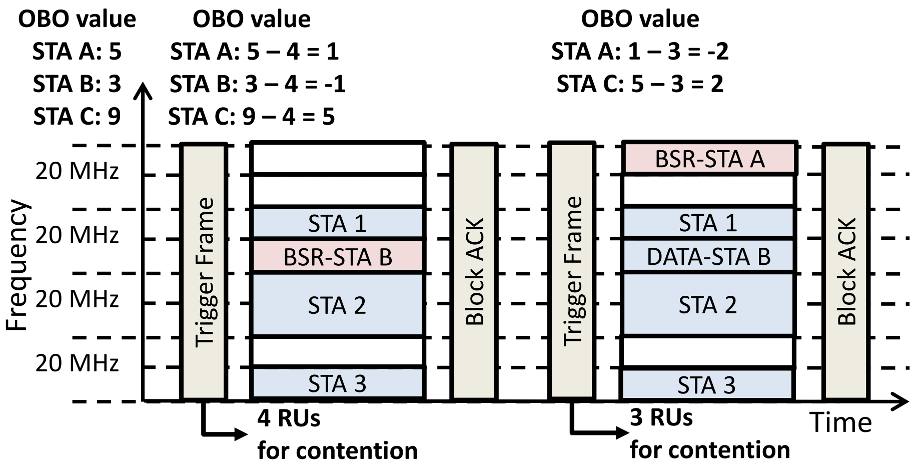


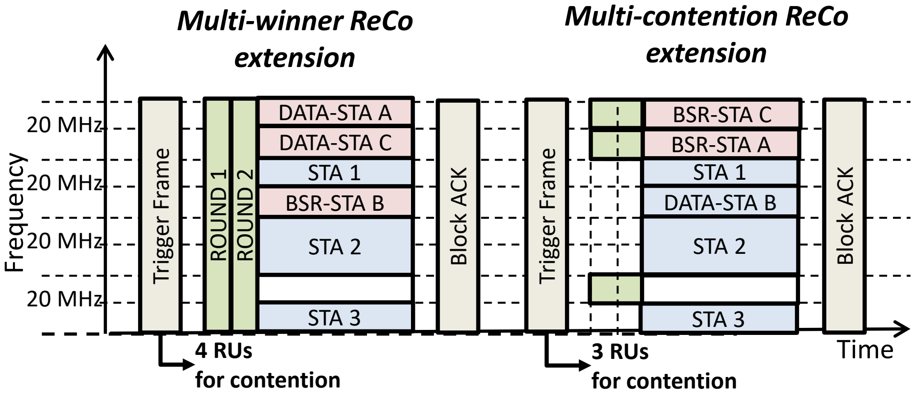
Figure 2.11. CDF of the number of slots varying SIR.

on OFDMA, which allows a fine-grained access to the channel resources organized into multiple sub-channels called Resource Units (RUs). Indeed, in OFDMA, the availability of OFDM multiple tones can be exploited for supporting multiple access, i.e. for enabling con-current transmissions to/from multiple stations. To this purpose, the AP needs to coordinate multiple stations providing a reference synchronisation signal and a precise scheduling of resource allocations, similarly to what happens in cellular systems. Scheduling of uplink transmission grants has been considered for optimizing the channel efficiency. However, random access cannot be completely avoided for several reasons, such as sending signalling information by new stations, notifying buffer status for facilitating AP scheduling decisions, or transmitting intermittent traffic. The random access mechanism defined in IEEE 802.11ax is called OFDMA Backo-OFF (OBO) [36] and implements a variant of multi-channel slotted Aloha by means of a back-off mechanism performed in the frequency domain among multiple resource units.

Although OFDMA is widely used in cellular networks, it is a significant novelty for WLAN networks, especially for uplink transmissions. For providing a coordination mechanism between stations, a trigger frame is sent by the AP before any uplink Multi-User (MU) transmission. The trigger frame has the role to provide a reference signal for synchronising over time



(a)



(b)

Figure 2.12. An example of uplink multi-user transmissions in IEEE 802.11ax networks under the OBO mechanism (a) or ReCo (b). RUs accessed by means of Random Access are colored in red, while reserved RUs are colored in blue.

the transmissions performed by independent stations (which start a SIFS after the end of the trigger frame) and identifying the right power level to be used. OFDMA in IEEE 802.11ax is frame-based, i.e. various tones grouped in Resource Units (RUs) are assigned to a given station for the entire frame duration. The trigger frame also specifies which RU is pre-allocated to a given station and which other units can be randomly accessed by all the stations. Acknowledgement frames are sent by the AP a SIFS after the end of the MU transmissions [37].

Fig. 2.12(a) shows an example of two OFDMA uplink transmissions in IEEE 802.11ax network, in which the channel bandwidth is organized into 8 RUs. Different colors are used for distinguishing the RUs accessed by means of scheduling (in blue) and random access (in red).

The contention process for utilizing the RUs left to contention implements a variant of multi-channel slotted Aloha called OFDMA Back-Off (OBO). Each transmitting station chooses a random value, called OBO value, in a contention window of size OCW. The contention window follows the usual exponential increment rule as a function of the transmission outcomes (i.e if the ACK is not received), and can vary between OCW_{min} and OCW_{max}, values that are signaled by the AP in the trigger frame. At each contention chance following the transmission of a trigger frame, the back-off counter is decremented by the number of RUs available in that contention chance. In other words, differently from legacy DCF, the back-off counter is not decremented over time as a function of the channel sensing results (i.e. when the channel is idle), but it is updated right after the transmission of the trigger frame. If the residual back-off counter is lower than the number of available RUs, the station can attempt a transmission by randomly choosing a RU among the ones devoted to contention. Otherwise, the station decreases the back-off counter and waits for the next contention chance. Fig. 2.12(a) shows the back-off counter updates performed by the stations involved in the example. The figure also shows that in a random transmission attempt stations can transmit either a data frame or a special control frame, called Buffer Status Report (BSR), to inform the AP on the status of the queues and ask for scheduled (reserved) RUs for subsequent transmissions. The details of the scheme can be found in [37].

2.10 ReCo for multi-user transmissions

Although the IEEE 802.11ax random scheme is very simple and its adoption can be limited to the transmission of the BSR frames, it is evident that significant channel wastes can arise for two main reasons: (i) lack of utilization of some RUs, which remain empty during the back-off countdown or are not selected by the stations that complete their countdown; (ii) collisions due to the selection of the same RU by multiple stations. For example, in the first uplink MU transmission of Fig. 2.12(a), only one of the four RUs available for random access is utilized for transmission, despite the fact that three stations are contending.

Our proposed solution can significantly reduce the waste of RUs at the expense of slightly increasing the duration of MU transmissions of a few extra symbols, as depicted in Fig. 2.12(b). During these extra symbols, the whole channel bandwidth is organized in tones (suitably spaced), to be used

Algorithm 2 Pseudo-code of the contention phase algorithm, ReCo for IEEE 802.11ax - MW ReCo

```
$r = 0$;
$c = 0$;
$dropout = FALSE$
while $(r < s)\&(dropout == FALSE)$ do
    $r = r + 1$;
    $y =$ extract_random_number$(1, m_r, \bar{q}^{(r)})$;
    transmit_burst$(r, f_y)$;
    if $y > RU_{free}$ then
        for $p = 1$ to $y - 1$ do
            $c = c +$ isbusy_channel$(r, f_p)$;
        end for
        if $c \geq RU_{free}$ then
            $dropout = TRUE$
        end if
    end if
end while
```

for the multi-round contention process. We remark that implementing ReCo in the frequency domain is possible without supporting full-duplex capability, i.e., without adding extra complexity to the stations as discussed in Sec. 2.5. Since in this case the contention process involves the allocation of multiple RUs rather than a single transmission grant, simple extensions of the basic ReCo scheme can be envisioned. In particular, we considered two possible solutions, which are also summarized in Alg. 2 and Alg. 3 by using the same convention (functions and variables) described in Sec. 2.3.

Multi-Winner (MW) extension. In order to allocate multiple RUs, it is possible to identify multiple winners at each contention round. Let RU_{free} be the number of RUs available for contention. At each round, each station decides to go to the next round if no more than $RU_{free} - 1$ tones are transmitted at a frequency lower than the tone selected by the station itself. Stations winning the last round take orderly the available RUs according to the position of the tones transmitted in the last round (starting from the lowest frequency).

Multi-Contention (MC) extension. In this case multiple allocations of RUs are achieved by performing multiple parallel contention processes for each RU. In other words, rather than using all the tones available for contention, at the first contention round stations randomly select a RU and pick one of the tones in the corresponding bandwidth. If the contention round is successful, i.e., no other station picks a tone in the RU bandwidth at a lower

Algorithm 3 Pseudo-code of the contention phase algorithm, ReCo for IEEE 802.11ax - MC ReCo

```
$r = 0$;
$k =$ extract_random_number$(1, RU_{free}, [\frac{1}{RU_{free}}, \frac{1}{RU_{free}}, \dots, \frac{1}{RU_{free}}])$;
$dropout = FALSE$
while $(r < s) \& (dropout == FALSE)$ do
    $r = r + 1$;
    $y =$ extract_random_number$(1, m_r, \bar{q}_k^{(r)})$;
    transmit_burst$(r, f_{y,k})$;
    if $y > 1$ then
        $lose =$ isbusy_channel$(r, [f_{1,k}, f_{2,k}, \dots, f_{y-1,k}])$;
        if $lose$ then
            $k =$ rand_isfree_otherchannel$(k)$;
            if $k == 0$ then
                $dropout = TRUE$
            end if
        end if
    end if
end while
```

frequency, the station moves to the next contention round in the same RU. Otherwise, rather than simply dropping out of the contention, the station can move to another RU in which no tone has been detected. After the last contention round, transmissions are finally attempted by the stations winning the contention in each RU. In the listing of Alg. 3, along the functions and variables described in Sec. 2.3, we rely on ad additional function:

- `rand_isfree_otherchannel`: generates integer samples uniformly distributed among the indexes of free RUs, i.e., in which no tone was detected during the previous round. Returns 0 if there are no free RUs.

Fig. 2.12(b) depicts the MU frames resulting from the two proposed access solutions. If we assume a number of contending stations higher than RU_{free}, the probability of leaving an empty RU is very low. The collision probability can be set at any desired level by configuring a suitable number of contention rounds. Therefore, the proposed solution can be more effective than the OBO scheme. However, differently from the frequency-domain contention implemented in OBO, these schemes strongly rely on the channel sensing capability of the stations. It is therefore necessary to add robustness to the method, as we investigated in Sec. 2.12, where we defined a version of ReCo, called ReCHo.

2.10.1 Simulation Results

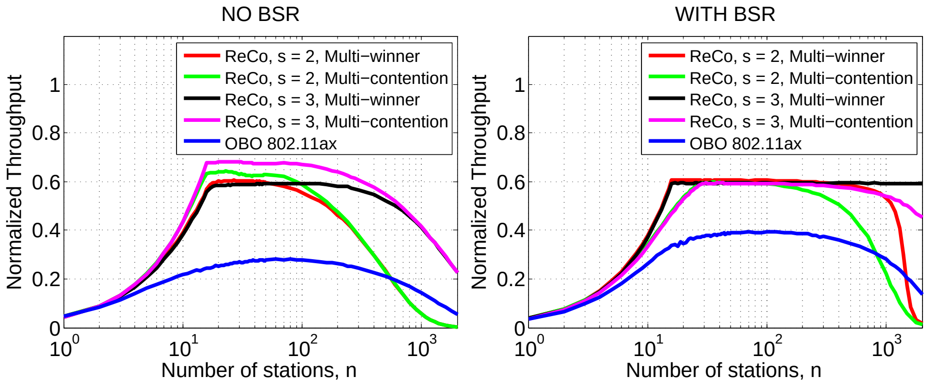
In order to compare the performance of the proposed ReCo variants with the OBO protocol, we implemented a simplified MATLAB OFDMA model, in which we completely abstracted the PHY into a number of available RUs (neglecting all the issues related to channel propagation models, synchronization errors, interference, imperfect channel sensing, etc., which we will analyze in the next chapter). More into details, we considered a bandwidth of 80 MHz, organized into 16 RUs and a single Spatial Stream, leading to an air bit rate of 600.5 $Mbps$ (≈ 37.5 $Mbps$ per-channel). Assuming that usual settings for $SIFS = 16$ μs and $DIFS = 34$ μs, for each uplink MU transmission the overheads due to the inter-frame spaces and to the transmission of the trigger frame and final acknowledgements correspond to 90 μs, while the overhead due to each ReCo contention round is set to 12.8 μs, i.e., to the maximum symbol duration. For the OBO protocol, the minimum contention window is 16 and the maximum contention window is 1024.

The activity phase is run by assigning one RU to each winning station for a time duration as specified in the Trigger frame (just as in IEEE 802.11ax). The issue of assigning a specific RU to each winning station can be solved analogously to OBO. In case of MC, k groups of contention tones are identified if k RU are opened up for contention. RU i is associated with group i, $i = 1, \ldots, k$. In case of MW, RU i is associated with the i-th ranked winning station. We run simulations in different network scenarios as the number of contending stations varies from 1 to 2000. We consider that the contention process can be used either for directly transmitting data frames or for transmitting the BSR control frames leading to channel reservations. In the first case, all the contending stations work with greedy traffic sources and saturated buffers; in the second case, after the transmission of a successful BSR frame, stations leave the contention until they receive a number of reserved transmission grants which varies uniformly from 1 to 10. Each transmission grant in one RU allows the transmission of a single frame, whose size is set to a constant value of 1500 bytes. Tab. 2.6 summarizes the simulation parameters and gives their numerical values.

Fig. 2.13 shows the normalized throughput achieved by OBO and by different versions of ReCo with 2 or 3 contention rounds in the two above cases: direct transmission of data frames on the RUs left to random access (NO BSR case) and usage of random access RUs for performing reservations (BSR case). The number of tones is $m = 64$. All the considered access mechanisms work on the whole set of 16 RUs. From the figures we can draw

Table 2.6. Numerical values of parameters used in the simulation.

W_{ch}	Radio channel bandwidth	80 MHz
RUs	Number of Resource Units	16
R	Bit rate	600.5 $Mbps$
L	MAC PDU payload	1500 bytes
T_{oh}	PHY/MAC overhead time plus inter-frame spacings	90 μs
BSR_v	BSR value	from 1 to 10
CW_{min}	Minimum contention window	15
CW_{max}	Maximum contention window	1023
s	Number of rounds for ReCo	2 or 3
m	Number of tones for ReCo	64
	ReCo contention phase time	$s \cdot 12.8$ μs
	Perfect channel sensing	



(a) (b)

Figure 2.13. Performance comparison of different access schemes in a OFDMA PHY with 16 RUs under direct frame transmissions (a) or BSR reservations (b).

two interesting observations. First, ReCo with $s = 3$ rounds outperforms OBO largely. For most values of n, also $s = 2$ contention rounds are enough for ReCo to peforma much better than OBO. Second, the MW scheme is better than the MC one when using BSR frames for reservations, while the MC scheme outperforms the MW one in case of direct frame transmissions. We justify these results by considering that the number of RUs available for contention in each uplink MU transmission can be much lower than the total number of RUs in case we use BSR frames for reservations. Indeed, in case of high load conditions, most RUs are used for scheduling transmissions of previously reserved stations and only a few RUs (e.g. one or two) are free for new reservations. If only few RUs are open for contention, running a single contention over all tones to elect multiple winners is a better strategy than having separate contention on each open RU, thus using a much smaller number of tones. For example, with a single RU left to random access, it is possible to minimize the collision probability by using the whole set m of available tones between n contending stations rather than one sub-set of $m/16$ tones between an average number of $n/16$ contending stations.

2.11 Impact of imperfect channel sensing

As seen so far, ReCo promises a low collision probability even for a large number of competing stations and few rounds, provided that carrier sensing works ideally. We explore here what happens when a more realistic propagation model is employed and the physical carrier sensing is modelled accounting for practical issues.

Imperfect tone detection can lead to non-uniform views of contention results among the stations. In general, either false positives and negatives are possible. A false positive can be the result of a noise spike or impulsive noise. False negative arise due to selective path loss, that kills some tones at a station receiver. In case that the tone erased by the channel at station S receiver is at a lower frequency than the tone selected by S, station S will miss the inhibition signal and will mistakenly move to the next contention round. the effects of these errors have been mentioned in the Sec 2.6.

Let us consider stations associated to an AP. For the association to be successfully performed, the level of power that a station must receive from the AP shall be greater than a prescribed threshold (the Carrier Detect Threshold, CDT). When the AP (a station) is transmitting, the stations (the AP) can sense a busy channel provided the signal level it receives is bigger than the Defer Threshold (DT). This works since the typical setting of the thresholds is such that $DT \leq CDT$ and we consider transmission over the entire bandwidth (i.e., frequency-selective fading is not a concern). When using sub-carrier signalling, frequency selective fading can harm reception. Therefore, proper threshold setting is required.

Let P_{tx} be the transmission power level, $G_d(x)$ the deterministic path gain at distance x between the transmitter and the receiver, G_s the path gain due to shadowing (obstacles), $G_f(f)$ the path gain due to frequency selective fading at frequency f. For the sake of a simple notation, we omit the argument f whenever there is no ambiguity. The power level received for a given tone at frequency f_0 transmitted by a device at distance x can be modelled as $P_{rx}(x) = P_{tx}G_d(x)G_sG_f(f_0)$.

In the following we assume that G_s is log-normal shadowing with unit mean and standard deviation $\sigma_S = 7\ dB$, accounting for obstacles, G_f is frequency selective Rayleigh fading with mean 1, and $G_d(x)$ follows a multi-slope power law given by:

$$G_d(x) = \begin{cases} \kappa \left(\frac{d_0}{d}\right)^{\alpha_1} & d_0 \leq d \leq d_c \\ \kappa \left(\frac{d_0}{d_c}\right)^{\alpha_1} \left(\frac{d_c}{d}\right)^{\alpha_2} & d \geq d_c \end{cases} \tag{2.16}$$

For the 5 GHz band of WiFi it is $\kappa \approx -47\ dB$. d_c represents a cut-off distance, where the power law exponent changes from $\alpha_1 = 2$ to a bigger value α_2. To model an indoor environment, we let $d_c = 5\ m$, $\alpha_2 = 3.5$ and we consider an additional path loss of 3 dB for each crossed wall, assuming there is one wall every 5 m. This is a simple model that aims at capturing an office building or a block of apartment flats. As for outdoor models, we let $d_c = 40\ m$ and $\alpha_2 = 4$.

Stations that associates successfully with the AP must be able to detect the carrier of the AP with a power level of at least CDT. Since detecting a signalling tone requires the received power level to exceed the threshold $P_{th} = DT$ and DT is several dB below CDT, stations detect tones emitted by the AP (and vice-versa) with very high reliability. Things are quite different as regards the possibility that stations hear to one another during the contention phase. With ReCo, contention depends critically on the fact that any station can detect reliably the tones sent by other contending stations. However, stations associated to a same AP can well be hidden to each other.

Fig. 2.14 plots the probability of missing a signalling tone as a function of the AP coverage radius, obtained with the outdoor radio channel model defined above. The power threshold is set to $P_{th} = DT = -88\ dBm$, i.e., equal to the Defer Threshold DT. With this setting, the false detection probability is negligible. The figure highlights that, if the threshold is set so that false detection is essentially ruled out, then the probability that a signalling tone might be missed is quite high, exceeding 0.1 for large coverage areas.

In order to gain insight into the effect of imperfect channel sensing on ReCo, we have implemented a MATLAB simulation, in which we can specify a PHY model and a network topology between different nodes implementing the ReCo access scheme. For sake of simplicity, let us consider the channel access problem on a given RU, then using the version of ReCo described in

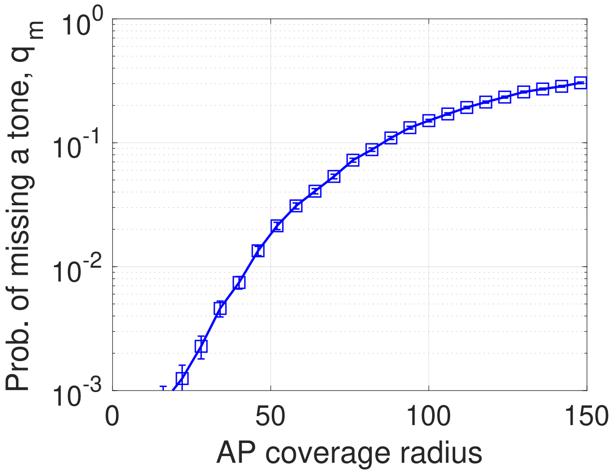


Figure 2.14. Probability q_m that a station misses a tone as a function of the AP coverage radius $R[m]$, under the constraint that the probability of false detection of a tone be 10^{-6}.

Table 2.7. Numerical values of parameters used in the simulation.

P_N	Noise floor power level	$-91\ dBm$
W_{ch}	Radio channel bandwidth	$20\ MHz$
n_{sc}	Number of sub-carriers	52
DT	Defer Threshold	$2 \cdot P_N$
CDT	Carrier Detect Threshold	$-82\ dBm$
P_{tx}	Transmission power level	$20\ dBm$
σ_S	Shadowing standard deviation	$5\ dB$
d_c	Cut-off distance of the path loss	$5\ m,\ 40\ m$
α_1	path loss exponent at distances $d < d_c$	2
α_2	path loss exponent at distances $d > d_c$	$3.5,\ 4$
A_w	path loss of a wall	$3\ dB$
R	Radius of the WLAN area	

sec. 2.3, with a bandwidth of 20 MHz. Our reference topology is given by an AP located at $(x_0, y_0) = (0,0)$ and n stations uniformly scattered around the AP, within a maximum distance R from the AP. Tab. 2.7 summarizes the simulation parameters and gives their numerical values.

A station receiving an average power level less than the *Carrier Detect Threshold* (CDT) from the AP is considered to be in outage and excluded from the WLAN (in other words, it cannot associate with the AP). The CDT is set to $-82\ dBm$ for a bandwidth of 20 MHz. Although all associated stations are in the AP range, some stations can be hidden to each other, thus missing each other signalling tones, as discussed at the end of the previous Section.

One important parameter for assessing the performance of the scheme is the number of stations surviving at each contention round, i.e., the number of stations transmitting a tone at the lowest frequency or missing the detection of a tone transmitted at a frequency lower than the one the station has chosen. In case of ideal channel sensing, it is very easy to find such a number of surviving stations. Let n be the number of backlogged contending stations at the beginning of the contention phase, s be the number of rounds and m be the number of frequencies. Let q_i denote the probability that a station picks the i-th frequency, $i = 1, \dots, m$. Let also $G_i = \sum_{j=i}^{m} q_j$ be the corresponding Complementary Cumulative Distribution Function (CCDF). The probability $P_{k,h}$ that h stations survive after a single contention round, given that k stations are contending at the beginning of that round, is

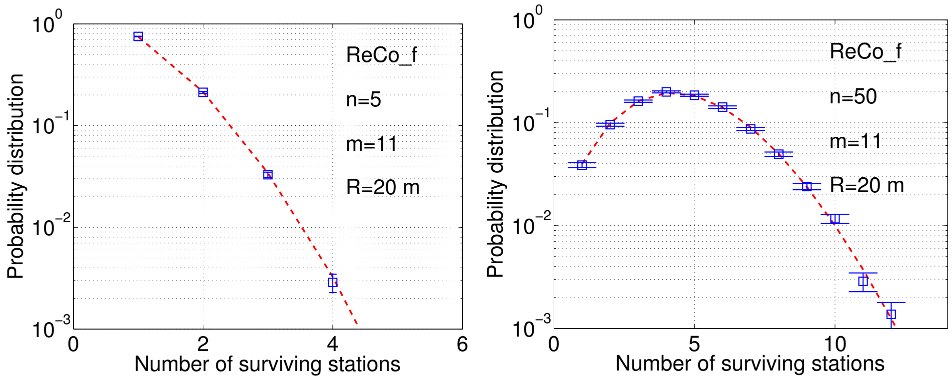
$$P_{k,h} = \sum_{i=1}^{m-1} \binom{k}{h} q_i^h G_{i+1}^{k-h}, \qquad h = 1, \dots, k-1 \tag{2.17}$$

This result holds under the assumption of *perfect channel sensing*, i.e., all stations can detect the tone signalled by any other station and no station makes a false detection.

We can find the number of surviving stations after multiple rounds by defining the $n \times n$ matrix $\mathbf{P}$, whose k-th row entries are $P_{k,h}$, for $h = 1, \dots, k$, and 0 for $h = k+1, \dots, n$ $(k = 1, \dots, n)$. If we denote with $\mathbf{x}(t)$ the probability vector at time t, whose i-th component $(i = 1, \dots n)$ represents the probability to have i surviving stations, the vector evolution at time $t+1$ can be obtained as $x(t)P$. Therefore, the number of stations that survive through s contention rounds (winning stations) is $\mathbf{x}(s) = \mathbf{x}(0)\mathbf{P}^s$, being $\mathbf{x}(0) = [0, 0, \dots 1]$.

Fig. 2.15 shows the probability distribution of the number of stations surviving one contention round for $m = 11$ signalling tones in a bandwidth of 20 MHz, the same value used in the testbed experiments. Square markers correspond to simulation results (including the 95% confidence intervals), while the dashed red line is the theoretical probability distribution derived above. We consider two scenarios, both based on the indoor radio channel model: a small WLAN with $n = 5$ stations (Fig. 2.15(a)) and a crowded WLAN with $n = 50$ stations (Fig. 2.15(b)). In all simulations we set $P_{th} = DT$, so that false detection has a negligible probability.

It is apparent that the analytical model matches the outcome of the simulations in spite of the radio channel impairments accounted for by the simulation model. In no one of the considered scenarios does the event ‘surviving stations = 0’ occur. All stations dropping out could occur in case a station



(a) (b)

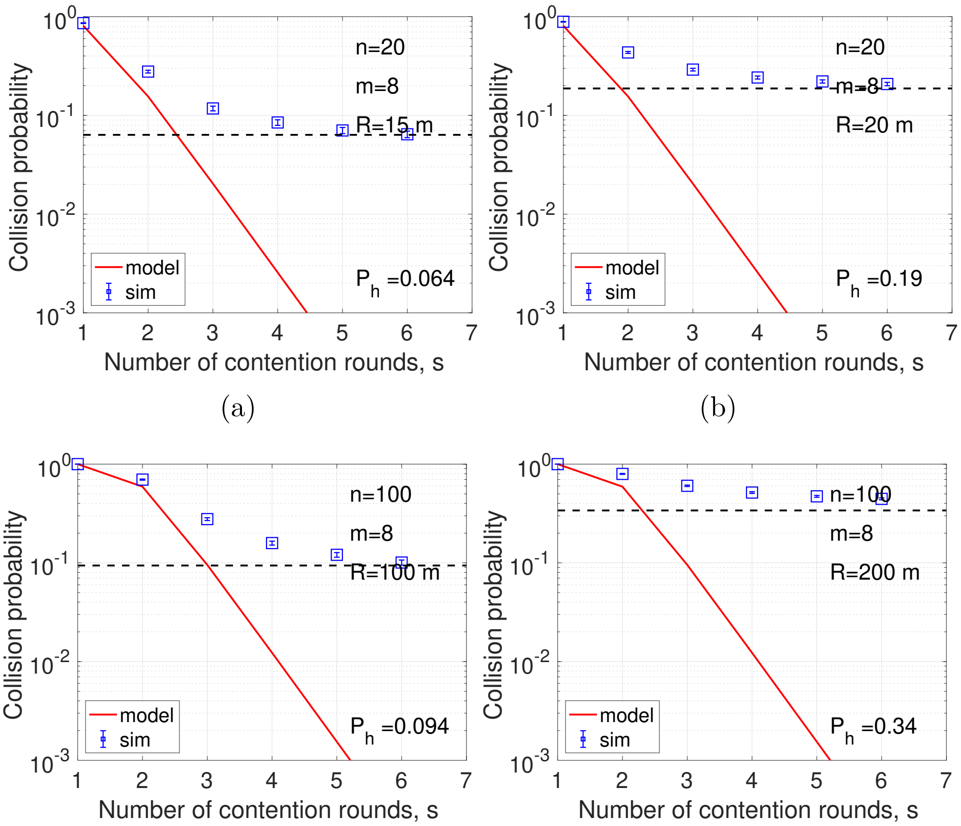
Figure 2.15. Probability distribution of the number of stations winning a single contention round (square marks: simulations, with 95% confidence intervals; dashed line: analytical model). The maximum distance of a station from the AP is $R = 20\ m$, $m = 11$. (a) $n = 5$ stations; (b) $n = 50$ stations.

mistook a spike of noise and/or self-interference for a busy tone. With the set threshold $P_{th} = DT$, this event is essentially ruled out.

While the first round of the access contention is not affected sensitively by physical layer impairments, the collision probability is definitely affected by the stochastic propagation channel. In ideal sensing conditions, the collision probability p_c can be expressed as the probability that the number of surviving stations after s rounds is higher than 1, and therefore it can be easily obtained from $\mathbf{x}(s)$. the collision probability p_c is bounded in Eq. 2.3, hence it decays exponentially with s, for a given m.

Fig. 2.16 shows the collision probability as a function of the number s of contention rounds for $m = 8$ and for indoor WLAN models with $n = 20$ stations, scattered within a distance $R = 15\ m$ (Fig. 2.16(a)) and $R = 20\ m$ (Fig. 2.16(b)) from the AP, and for relatively large outdoor WLAN models with $n = 100$ stations, scattered within a distance $R = 100\ m$ (Fig. 2.16(c)) and $R = 200\ m$ (Fig. 2.16(d)) from the AP.

The solid line represents the analytical model results that assume ideal channel sensing. The square markers (with 95% confidence intervals) represent the outcome of simulations of ReCo. The probability that a couple of stations be hidden to each other (i.e., the average power level received when one transmits and the other receives be less than DT) is annotated on each graph and marked by a horizontal dashed line. It is apparent that the analytical model predicts the collision probability correctly both with few and many stations, up to the point where the collision probability level



(a) (b)

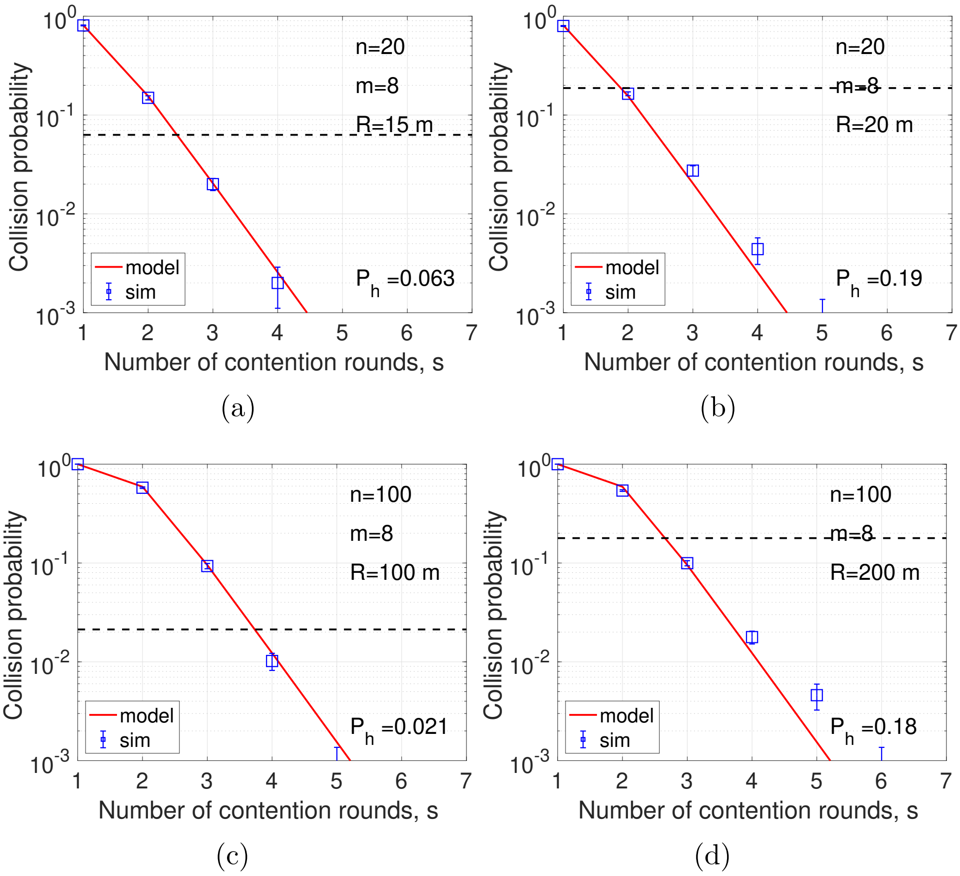
(c) (d)

Figure 2.16. ReCo collision probability as a function of the number of contention rounds s for $m = 8$. Top plots refer to an indoor channel with $n = 20$ stations. Bottom plots refer to an outdoor channel with $n = 100$ stations. The dashed line represents the probability that two stations picked at random are hidden to each other.

falls below the probability of having hidden stations. Below that level, the analytical model is optimistic. The actual level of collision probability exhibits a floor, due to false negatives, i.e., to stations missing the busy tones of other stations hidden to them. The strong degradation of ReCo performance under imperfect channel sensing could be somewhat mitigated by trying to adjust the threshold P_{th}, even if this comes at the cost of having to fine tune a network access parameter for each given AP configuration.

2.12 Echo mechanism and ReCHo

In the previous section, we observed that ReCo performance can be degraded in presence of imperfect channel sensing. For infrastructure networks, we can



(a) (b)

(c) (d)

Figure 2.17. ReCHo collision probability obtained in the same cases of Fig. 2.16, but with the tone-echo mechanism activated at the AP

exploit the role of the AP to improve the robustness of the access procedure with respect to collisions. The key idea is to have the AP re-transmit the least frequency tone among those transmitted by the stations, in order to increase the detection range of tones. The duration of the contention phase must be extended to make room for the extra-signalling of the AP. Specifically, we enlarge the contention phase to $2s$ mini-slots in case of s contention rounds. One contention round consists of two adjacent mini-slots. In the first one, stations (possibly including the AP) choose a random tone among the m possible tones and transmit it. In the immediately following mini-slot of the same round, stations remain silent and the AP transmits a tone on the lowest frequency it has received in the first mini-slot of the contention round. In other words, the AP *echoes* the least tone it has heard during the first part of the round. For this reason, this new access scheme is dubbed ReCHo.

Since sensing is reliable when involving the AP (i.e., stations sensing the channel when the AP is transmitting or vice-versa), we expect that the AP will detect the tones transmitted by the stations reliably during the first half of the contention round and, conversely, stations will detect the tone echoed by the AP reliably during the second part of the contention round.

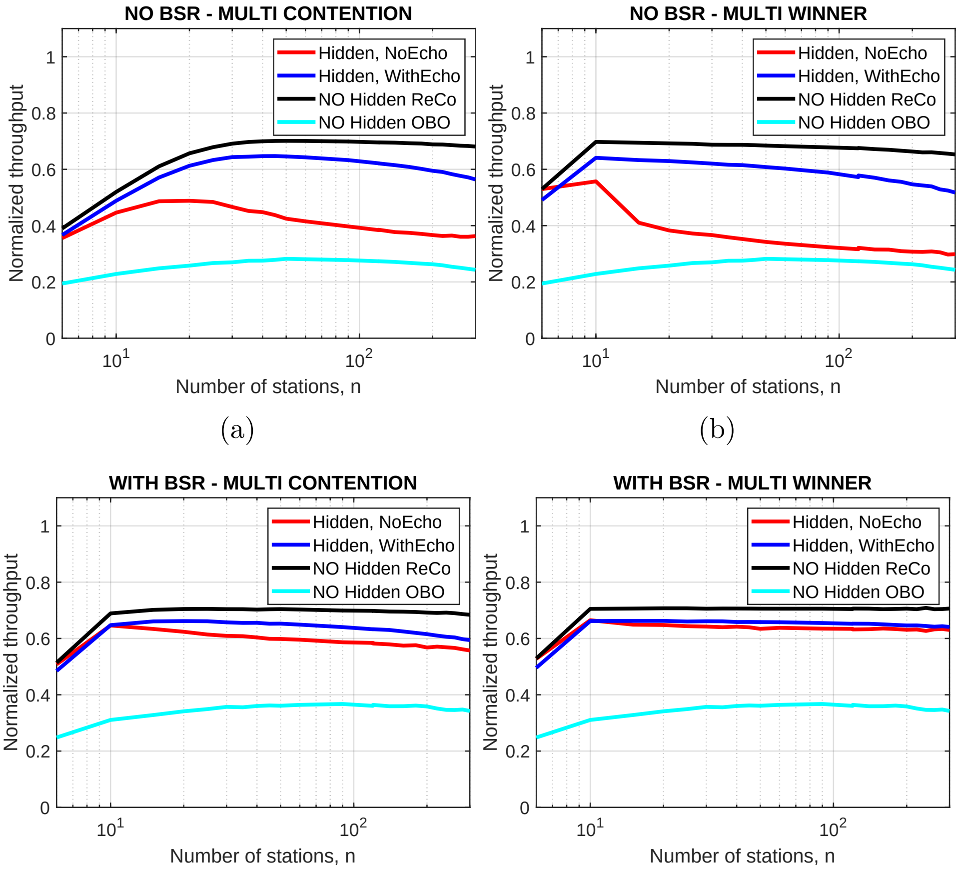
Note that, if uplink and downlink transmissions are managed according to time duplexing, as envisaged in IEEE 802.11ax with trigger frames, during the uplink phase only stations send tones in the first mini-slot of a contention round and only the AP echoes the lowest tone in the second part of the round. Thus, in these cases, it is not even required that stations perform tone reception at the same time as they are transmitting their own tone.

Fig. 2.17 shows the collision probability obtained in the same cases of Fig. 2.16, but with the tone-echo mechanism activated at the AP. It appears that tone echoing removes most of the fading effect. Still, for low levels of the collision probability (around 10^{-3} or lower), the performance obtained with the simulations are somewhat worse than those predicted according to the analytical model that assumes perfect Clear Channel Assessment(CCA) operations.

2.12.1 ReCHo performance applied to 802.11ax

In order to compare the performance of the proposed ReCo variants with the OBO protocol, we extended the MATLAB simulator used for the assessment of the impact of imperfect channel sensing with a complete MAC layer. ReCHo planned that the AP retransmit the least frequent tone among those transmitted by the stations. To apply ReCHo to MU transmissions provided by the emerging IEEE 802.11ax standard, a slight modification is required. The AP retransmit the highest tone among the lowest RU_{free} tones in case of Multi-Winner and retransmit many tones, the lowests in the RU_{free} contentions, in case of Multi-Contention.

We consider almost the same parameters used in Subsec. 2.10.1 (so, the same parameters of tab. 2.6), with some differences, we report them all for simplicity. Bandwidth of 80 MHz, organized into 16 RUs and a single spatial stream (8 RUs free for contention and 8 RUs perpetually busy), bit rate of 600.5 $Mbps$ ($\approx$ 37.5 $Mbps$ per-channel), $SIFS = 16\ \mu s$, $DIFS = 34\ \mu s$, payload of 1500 bytes, overhead due to the inter-frame spaces and to the transmission of the trigger frame and final acknowledgements are to 90 μs, overhead due to each contention round is set to 12.8 μs. For the OBO protocol, the minimum contention window is 16 and the maximum contention window is 1024. We also consider in the simulation both the MW and MC



(a) (b)

(c) (d)

Figure 2.18. Performance comparison of different access schemes in case multi contention and multi winner when BSR is disabled (a)(b) or BSR is enabled (c)(d).

schemes for contending on multiple RUs. We run simulations in different load conditions, varying the number of contending stations from 5 to 110. We perform simulations with and without the BSR control frames. BSR buffer frame varies uniformly from 1 to 6. Moreover, we consider two scenarios: i) hidden nodes, where the stations are scattered over a circular ring around the AP, between a distance of 10 m and 20 m from the AP; ii) no hidden nodes, where nodes are scattered over a circle of radius 5 m centered at the AP.

Fig. 2.18 shows the normalized throughput achieved by OBO and by different versions of tone-based contention schemes (MC and MW), for $m = 64$ and $s = 3$. With no hidden nodes (black curves), ReCo achieves a significant performance improvement in comparison to OBO, thanks to the capability of supporting a higher channel efficiency (i.e. avoiding wastes of RUs). When a channel model incorporating hidden nodes is used, throughput performance degrade as indicated by the red curves. However, the performance are still higher than the ones perceived by OBO. If the tone-based contention is extended with the echo mechanism proposed in ReCHo, despite of the increased contention overhead, the overall throughput gets close to the ideal channel results, as indicated by the blue curves. Also in presence of BSR reservations for multiple transmissions, the ReCHO scheme outperforms OBO. However, the performance difference between the scenarios with and without hidden nodes is reduced, because the utilisation of BSR frames reduces the number RUs left open for random contentions.

Chapter 3

Long Range Technology

3.1 Introduction

LoRa provides different possibilities to orthogonalize transmissions as much as possible – Carrier Frequency (CF), Spreading Factor (SF), Bandwidth (BW), Coding Rate (CR) – and provides simultaneous collision free communications. However, despite the robustness of the LoRa PHY[1] patented by Semtech, in WAN scenarios where multiple gateways can be installed, the scalability of this technology is still under investigation [38]. Current studies are mostly based on the assumption that using multiple transmission channels and spreading factors leads to a system that can be considered as the simple super-position of independent (single channel, single SF) subsystems [39]. This is actually a strong simplification, especially because the SFs adopted by LoRa are pseudo-orthogonal [40] and therefore, in near-far conditions, collisions can prevent the correct reception of the overlapping transmissions using different SFs.

For characterizing these phenomena, in this Chapter we provide two main contributions: a link-level characterization of LoRa modulation and then, exploiting such link-level properties, we provide a complete cell model study of multi-link LoRa systems. Regarding the first aspect, we characterize LoRa modulation experimentally, showing that collisions between packets of different SFs can indeed cause packet loss. We modified the software transceiver

---

[1] Semtech. LoRa Modulation Basics. AN1200.22, Revision 2, May 2015. https://www.semtech.com/images/datasheet/an1200.22.pdf

presented in [41] to generate synthesized LoRa modulated packets and transmit them through the well-known USRP software-defined-radio (SDR) platform. This transceiver is used to emulate, in a controlled and repeatable manner, collisions produced by different devices: the modulated LoRa signals are first generated in software, then summed together (with tunable power level difference) to replicate a given super-position of LoRa signals and finally the obtained combined radio signal is transmitted over the air. We use this traffic generator to experimentally characterize the performance of a commercial LoRa device, under intra-SF and inter-SF collisions caused by multiple simultaneously active LoRa links. We quantify the power difference for which capture effects and packet loss occur, for all combinations of SFs. Our experimental results show that the co-channel rejection thresholds are on average an order of magnitude higher than the theoretical ones presented in [42], with values as high as -8 dB. These poor co-channel rejection thresholds might be insufficient in common LoRa application scenarios (the received power of two radio signals can easily differ by tens of dB), thus contradicting the common belief that pseudo-orthogonal SFs can be considered as orthogonal in practice.

The second main contribution of this Chapter regards the capacity analysis of a LoRa cell under realistic link behaviors: we propose a simple yet accurate analytical framework to model the performance of LoRa cells, deriving the aggregated capacity and data extraction rate of a LoRa cell working on a single frequency with one or multiple gateways. The framework has been built as a generalization of the Aloha model (the channel access protocol used in LoRa), by taking into account the heterogeneous probabilities of intra-SF and inter-SF collisions, due to the specific position of the target ED (which translates in a specific received power at the gateway). The models provide excellent results, closely following the simulations obtained with LoRaSim [38] and with our own custom Matlab simulator. Our analysis demonstrates that capture effects and imperfect orthogonality of SFs can significantly affect the cell capacity. In particular, we show that more robust SFs, usually envisioned for EDs experiencing strong channel attenuations, can be severely affected by inter-SF interference and therefore, their usage could consume a large fraction of cell resources without real benefits. Also, we show that power control and packet fragmentation can be counterproductive. Finally, we quantify the performance increase obtained by deploying multiple gateways and we show that it might be best to deploy them at the edge of the cell more than on a regular grid.

The rest of the Chapter is organized as follows. After a brief review of literature work about LoRa in §3.2, we provide a background description of LoRa modulation and a characterization of link-level performance in §3.3. The analysis of cell capacity is presented in §3.4, while in §3.5 we extend our model in case of non-uniform allocation of SFs, power control and packet fragmentation. Finally, we analyze the capacity improvements achievable with multiple gateways and the performance impact of topology in §3.6.

3.2 Related Work

Since LoRa is a fairly new technology, relatively few works exist in the literature. Link-level studies are mainly based on the experimental characterization of the coverage and on the rejection of interference properties of Semtech's patented LoRa PHY [43]. The paper in [42] quantifies the Signal-to-Interference-Ratio (SIR) power thresholds needed to reject interference caused by LoRa signals modulated with different SFs. However, the presented theoretical results are very different from our experimental ones. In [40], the performance of LoRa is compared to ultra-narrowband technologies (such as Sigfox), where it is shown that ultra-narrowband has a greater coverage although LoRa networks are less sensitive to interference.

Studies at the cell level are based on the characterization of the link behavior. In [38], after demonstrating the capture phenomena between LoRa frames, authors quantify the capacity of a cell in simulation[2]. The simulator assumes that a 6 dB power ratio between the collision packets is needed for channel captures and that different SFs can be considered completely orthogonal [39]. This last hypothesis is a strong simplification, since LoRa SFs are pseudo-orthogonal [40] and inter-SF collisions can appear in near-far scenarios. In [44], the authors propose a solution to improve LoRa performance in high density scenario, but the details of the model used are not provided.

Capture effects can significantly increase the performance of wireless systems, because the strongest received signal might be correctly demodulated even in case of collision. Several approaches to model this phenomenon have been presented estimating interference power, collision times, channel fading, etc. Given the number of interfering frames and their power, it is possible to estimate channel captures when the SIR is greater than the capture threshold [45]. Alternatively, the highest possible interference level can be mapped to

---

[2] http://www.lancaster.ac.uk/scc/sites/lora/

a vulnerability range from which interfering signals do not affect packet reception [46]. Despite the simplicity this approach, this model provides good results in network with stable conditions. Thus, in this chapter we generalize the concept of vulnerability in both intra-SF and inter-SF collisions.

3.3 Dissecting LoRa

In this section, we provide a characterization of LoRa link-level performance which will then be exploited in the next sections to develop our cell-level model of LoRa systems.

3.3.1 Modulation and Demodulation

LoRa modulation is derived from *Chirp Spread Spectrum (CSS)*, which makes use of *chirp signals*, i.e. frequency-modulated signals obtained when the modulating signal varies linearly in the range $[f_0, f_1]$ (upchirp) or $[f_1, f_0]$ (downchirp) in a symbol time T. Binary modulations, mapping 0/1 information bits in upchirps/downchirps, have been demonstrated to be very robust against in-band or out-band interference[3]. LoRa employs a M-ary modulation scheme based on chirps, in which symbols are obtained by considering different circular shifts of the basic upchirp signal. The temporal shifts, characterizing each symbol, are slotted into multiples of time $T_{chip} = 1/BW$, called chip, being $BW = f_1 - f_0$ the bandwidth of the signal. It results that the modulating signal for a generic n-th LoRa symbol can be expressed as:

$$f(t) = \begin{cases} f_1 + k(t - n \cdot T_{chip}) & \text{for } 0 \leqslant t \leqslant n \cdot T_{chip} \\ f_0 + k(t - n \cdot T_{chip}) & \text{for } n \cdot T_{chip} < t \leqslant T \end{cases}$$

where $k = (f_1 - f_0)/T$ is the slope of the frequency variations. The total number of symbols (coding i information bits) is chosen equal to 2^i, where i is called spreading factor. The symbol duration T required for representing any possible shift is

$$T = 2^i \cdot T_{chip} = \frac{2^i}{BW} \tag{3.1}$$

It follows that, for a fixed bandwidth, the symbol period and the temporal occupancy of the signal increase with larger SFs. Fig. 3.1 shows the mod-

---

[3] Semtech. LoRa Modulation Basics. AN1200.22, Revision 2, May 2015. https://www.semtech.com/images/datasheet/an1200.22.pdf

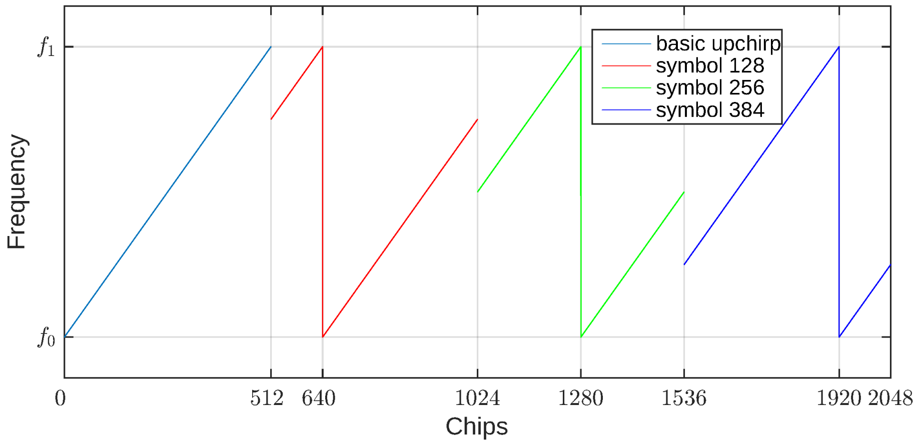


Figure 3.1. Modulating signal with $SF = 9$ for one basic upchirp and three symbols: 128, 256 and 384.

ulating signal used for a basic upchirp and three examples of circular shifts obtained for a SF equal to 9: the symbol time is 512 T_{chip}, while the three exemplary shifts code the symbols 128, 256 and 384.

The preamble of any LoRa frame is obtained by sending a sequence of at least eight upchirps followed by two coded upchirps, used for network identification (sync word), and two and a quarter base downchirps. Payload data are then sent by using the M-ary modulation symbols. LoRa provides three BW settings (125, 250 or 500 kHz) and seven different SF values (from 6 to 12). In general, a larger bandwidth translates in a data rate increase and a receiver sensitivity deterioration. Conversely, higher spreading factors can be used for improving the link robustness at the cost of a lower data rate. For demodulation, the received LoRa signal is synchronously multiplied to the base downchirp. This results in a signal comprising only two frequencies: $f_n = -kn \cdot T_{chip}$ and $f_n - BW = -(f_1 - f_0) - kn \cdot T_{chip}$. Both frequencies will be aliased to the same frequency f_n by downsampling at the rate BW. The estimated symbol index $\hat{n}$ corresponds to the position of the peak at the output of an iFFT, as described in [42].

An interesting feature of LoRa modulation is the orthogonality of signals modulated under different spreading factors, which can be exploited for enabling multiple concurrent transmissions. Although perfect orthogonality is guaranteed only in case of exact synchronization of the transmitters, the cross-energy between two signals modulated with different spreading factors is almost zero, regardless of the starting of the symbol times. Then the cross-energy is not exactly equal to zero although it reaches very low values.

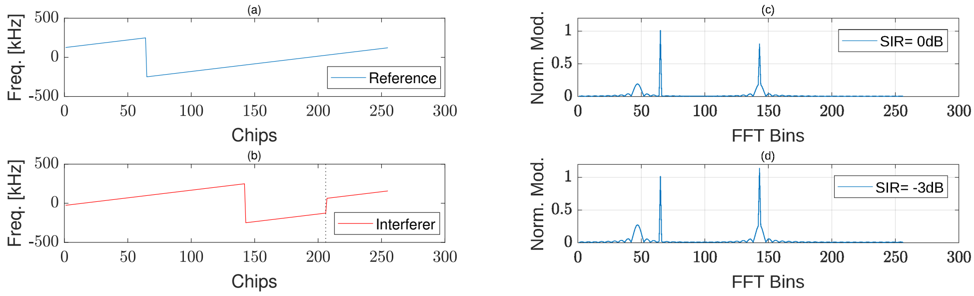


Figure 3.2. An example of capture effect within signals modulated with same SF 8. A LoRa reference symbol (a) and two partially overlapping interfering symbols (b) are received at different SIR levels. The iFFT output after multiplication with the base SF 8 downchirp and downsampling shows the highest peak for the perfectly synchronized reference symbol and two lower peaks for the partially overlapping symbols (c) but a SIR of -3 dB is enough to overcome the reference signal (d).

In general, considering two LoRa signals modulated with different spreading factors, say s_1 and s_2, the cross-energy between them is not equal to zero:

$$E_{s_1,s_2}(\tau) = \int_0^T s_1(t) \cdot s_2(t-\tau)^* dt \simeq 0 \tag{3.2}$$

where T is the symbol period of the signal with the highest spreading factor.

In case of collisions with other LoRa symbols, we can distinguish two different scenarios, depending on the interfering spreading factor SF_{int}. First, if the SF_{int} is the same as the one the receiver is listening for, the above receiver will observe multiple peaks at the output of the iFFT. Indeed, assuming that the two transmissions are received at the same power and that the reference signal is perfectly synchronized with the receiver, the iFFT will show a maximum peak corresponding to the reference symbol and two smaller peaks corresponding to two partially overlapping interference symbols, with different height depending on the transmitted symbols and on the offset with the receiver. For example, Fig. 3.2 shows two signals modulated with same SF 8 and bandwidth 500 kHz: the reference symbol (Fig. 3.2(a)) and two partially overlapping interfering symbols (Fig. 3.2(b)). As depicted in Fig. 3.2(c), when the signals are received with the same power, the iFFT output after multiplication with the base downchirp and downsampling, shows the highest peak for the synchronized reference symbol (index $\hat{n} = 64$) and two lower peaks for the partially overlapping symbols (index $\hat{n_1} = 96$ and $\hat{n_2} = 192$, with shift of $0.2T$ – i.e. 51.2). However, a Signal to Interference

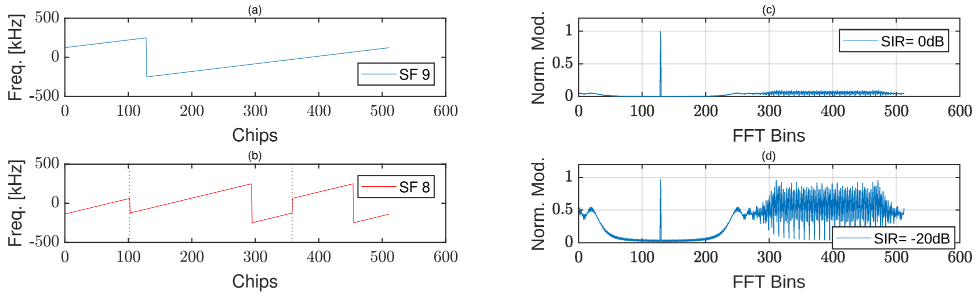


Figure 3.3. An example of collision between signals modulated with different SF. A LoRa symbol modulated with SF equal to 9 (a) and two overlapping and circularly shifted interfering symbols with SF 8 (b) are simultaneously received at different SIR levels. The iFFT output after multiplication with the base SF 9 downchirp and downsampling shows a clear peak when the two signals have the same power (c) while this is not the case when the SIR is too low (d).

Ratio (SIR) of -3dB is enough for the interfering signal to overcome the reference signal and "capture" the channel (Fig. 3.2(d)). This means that LoRa exhibits a very high capture probability within the same SF.

Second, when the SF_{int} is different from the one the receiver is interested in, after multiplication with the base downchirp and downsampling, the interfering signal will still be a chirped waveform, resulting in a wide-band spectrum with low spectral density. An example is illustrated in Fig. 3.3, where one signal modulated with SF equal to 9 (Fig. 3.3(a)) is overlapped to two symbols modulated with SF 8 (Fig. 3.3(b)), circularly shifted to desynchronize them with the reference symbol (the dotted lines represent the boundaries of the symbols). At the receiver, when the two signals are received with same power, the iFFT output after multiplication with the base SF 9 downchirp and downsampling shows a clear peak corresponding to the reference SF 9 symbol index $\hat{n} = 128$ (Fig. 3.3(c)), while this is not the case when the SIR is too low because of the non perfect SF orthogonality (Fig. 3.3(d)). In this scenario, the co-channel rejection is much higher (≈ -20dB in the figure).

3.3.2 PHY Coding

Up to now, we have neglected the impact of bit coding schemes. Indeed, the patented LoRa PHY includes several mechanisms to make the system more robust to interference. After transmitting the preamble, both header and payload bits of LoRa frames are mapped to symbols by a pipeline of processing operations, which include: Hamming coding[4], whitening, shuffling & interleaving, and gray coding. These operations have been specifically designed for increasing robustness towards synchronization errors or narrowband interference, which can be a serious issue for CSS-based modulations. In fact, in case of synchronization errors or narrowband interference, the receiver described in the previous section will most probably mistake the transmitted symbol, mapped to frequency f_n after the iFFT, for one of the immediately adjacent symbols. Since gray coding ensures that adjacent symbols are mapped to bit patterns differing in one position only, the receiver is able to identify the less reliable bits (at most two bits) of each received symbol. The purpose of the LoRa interleaver is spreading unreliable bits among several codewords, thus enabling even the 4/5 Hamming code (consisting in a simple parity check) in exhibiting a significant channel coding gain.

In order to understand if Gray coding has an impact also on inter-SF interference, we tried to characterize the distance between the transmitted symbol and the decoded one in presence of inter-SF collisions. To this purpose, we use our MATLAB implementation with Gray encoding and quantified such distance in our simulation. For example, Fig. 3.4 shows the histogram of the Hamming distance of the decoded symbol from the transmitted one, when a LoRa transmission at SF equal to 12 (with Gray encoding enabled) is interfered by another transmission with SF equal to 8 and SIR=-24dB. From the figure it is clear that the error distance probability approximates a Binomial distribution (and is not concentrated around the adjacent symbol). Thus, LoRa PHY coding mechanisms can mitigate synchronization errors but cannot protect from collisions.

3.3.3 Imperfect Orthogonality Quantify

To quantify the co-channel rejection, including the impact of PHY coding, we implemented a LoRa modulator and demodulator in MATLAB, based on [42] and [47]. We performed a number of simulations for testing the reception of two overlapping transmissions modulated with different SFs, after Hamming coding at rate 4/7, interleaving and Gray encoding. Our goal is identifying a SIR threshold below which the demodulation of the received frame is affected

---

[4] The Hamming codes used in LoRa have a coding rate between 4/5 and 4/8, and can reveal or correct at most one error on the AWGN channel.

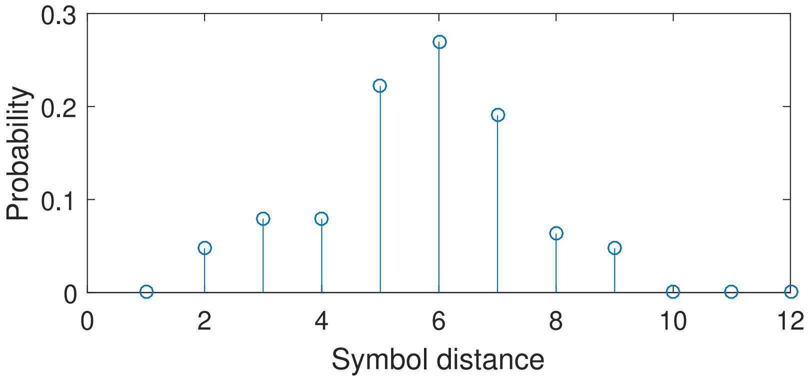


Figure 3.4. Histogram of the error distance when a LoRa transmission at $SF = 12$ (with Gray encoding enabled) is interfered by another transmission with different $SF = 8$ at SIR=-24dB.

by errors. In each simulation run, we created an overlapped signal by summing the reference frame, modulated with a reference spreading factor SF_{ref}, with a number of random interfering symbols, modulated with a different spreading factor SF_{int} (with an equivalent time on air). We assumed the transmitter to be perfectly synchronized with the receiver, while the interference frame is randomly shifted in time for de-synchronizing the interfering symbols. The number of interfering symbols for each reference symbol can be calculated as $N_{\text{int}} = 2^{SF_{\text{ref}}-SF_{\text{int}}}$. The amplitude A_{ref} of the reference signal is set to one, whereas the amplitude A_{int} of the interferer is a tunable value depending on the SIR, i.e. $A_{\text{int}} = \sqrt{10^{-SIR/10}} \cdot A_{\text{ref}}$. The resulting combined signal has been then processed by the MATLAB demodulator, in absence of noise on the channel. For each simulation run, we randomly generated interfered packets until the occurrence of 100 total error events. Packets are transmitted with SF_{ref} and include 20 Bytes of data and a zero padding up to an integer number of interleaving blocks. This signal is interfered by a random LoRa-like signal modulated with SF_{int}, with a random time-offset and a SIR increasing from -30dB with 1dB steps. A Bit Error Rate (BER) statistic has then been obtained by comparing the demodulated bits with the modulated ones.

Table 3.1 summarizes the SIR thresholds leading to a BER of approximately 1%. In the table, we also consider the case when the interfering

Table 3.1. SIR thresholds in MATLAB simulations.

SF_{ref} \ SF_{int}	7	8	9	10	11	12
7	0	-11	-13	-14	-14	-14
8	-13	0	-14	-16	-17	-17
9	-17	-16	0	-17	-19	-20
10	-19	-19	-19	0	-20	-22
11	-22	-22	-22	-22	0	-23
12	-24	-24	-25	-25	-25	0

signal has the same SF of the reference signal. As also documented in the Semtech specifications, LoRa modulations achieve a very high probability of capture effects even with low SIR values (0dB for the different SFs in our simulations, versus 6dB specified by Semtech).

For validating the thresholds found with the MATLAB simulator, we performed a number of experiments on real LoRa links. To this purpose, we used a Semtech SX1272 transceiver, controlled by an Arduino Yun, for characterizing the behavior of a commercial receiver in presence of collisions. We implemented a LoRa synthesizer able to encode, modulate and generate the I/Q samples of a real LoRa packet, which can be easily transmitted over the air with a USRP B210 board through GNU radio. With this LoRa synthesizer, we generated two traces (one for the interferer and one for the reference LoRa link) for each combination of SFs, composed of a stream of 20 byte-long packets (for the reference SF) and adjusting the payload length of the interfering SF to match the length of the reference signals. The offset of each interfering packet, overlapped in time to the packets of the reference link, has been randomly selected within a window which guarantees that the two packets collide for at least one symbol. We filled the payload of all frames with randomly generated bytes, except for the two bytes that specify the destination address and the payload length. In particular, we assigned the destination address of the SX1272 receiver only to the packets of the reference link. This allows the receiver to discard the interfering packets when they are modulated with the same SF of the reference ones. Finally, we scaled the amplitude of the interfering packet stream to achieve the desired SIR and added it to the reference stream. For each couple of SF_{ref} and SF_{int}, the resulting combined stream was transmitted through the USRP towards the SX1272, thus emulating the traffic generated by two different transmitters.

The results of the experiments are summarized in table 3.2, for a subset

Table 3.2. SIR thresholds with SX1272 transceiver.

SF_{int} SF_{ref}	7	8	9	10	11	12
7	1	-8	-9	-9	-9	-9
8	-11	1	-11	-12	-13	-13
9	-15	-13	1	-13	-14	-15
10	-19	-18	-17	1	-17	-18
11	-22	-22	-21	-20	1	-20
12	-25	-25	-25	-24	-23	1

of reference and interfering SF combinations. The table shows that the SIR thresholds for correct demodulation are similar to the ones obtained in MATLAB simulations and very different (over 10 dB – an orders of magnitude) lower than the ones in [42], with values as low as -8 dB. In table 1 of [42], the lower triangular part follows the law $10 \cdot \log_{10}(2^{SF_{\text{ref}}})$, i.e. the SIR thresholds are equal to the spreading gain of a matched filter receiver over an AWGN channel. However, this result is unrealistic, because the receiver does not work by comparing the mean squares of the signals and the interfering signal is not a white process. Such power difference between two radio signals can easily appear in common LoRa application scenarios, thus contradicting the common belief that different SFs can be considered as orthogonal in practice.

3.4 LoRa Single Gateway

The brief description of the LoRa PHY presented in section 3.3 enlightens two important aspects that have to be considered for studying the real capacity of LoRa cells: i) the possibility of correctly receiving a packet, in case of collision with other packets modulated with the same SF; ii) the possibility that multiple SFs are not exactly orthogonal and therefore do not work as independent multiple channels. In this section we show that both these aspects have a strong effect on the uplink cell capacity, because of the simple access scheme used in LoRa, which is basically a non-slotted Aloha mechanism (without carrier sense). We derive some simple expressions for predicting LoRa uplink capacity in presence of a single gateway, in terms of average throughput, by generalizing the classical Aloha results in presence of channel captures and imperfect orthogonality between SFs. We also compare our capacity models with simulation results obtained by using a custom Matlab simulator, which we also validated against the LoRaSim simulator

Table 3.3. Numerical values of simulation parameters.

Carrier Frequency	868.0 MHz
Transmission Power	14 dBm
Bandwidth	500 kHz
Code Rate	4/5
Message size	20 Bytes
Message Period	90000 ms
Number of EDs	[50-2000]
Path loss attenuation exponent	4
Simulation time	9000 s

used by the authors of [38] (which we warmly thank for publishing the source code). Unless specified otherwise, the parameters used for configuring the reference LoRa cell are summarized in table 3.3.

3.4.1 Ideal Cell Capacity

LoRa cells work as non-slotted Aloha systems. Under Poisson packet arrivals, the throughput of an ideal non-slotted Aloha cell is $G \cdot e^{-2G}$, being G the normalized load offered in the cell, i.e. the amount of data transmitted in the unit time by the EDs over nominal channel capacity. The Data Extraction Rate (DER), i.e. the probability of correctly receiving a packet transmission – a typical parameter for characterizing LoRa systems–, is given by e^{-2G}. In ideal conditions, since different SFs are available, the system works as the super-position of multiple coexisting (but independent) Aloha systems, each one experiencing the load due to the EDs employing a given SF equal to i (with $i \in \{7, 12\}$).

Let $G_i = \lambda_i \cdot ToA_i$ be the load offered in the cell sub-system working with SF i, which depends on the packet arrival rate λ_i and packet transmission time ToA_i (also called Time on Air or airtime). The ToA_i values change significantly from one SF to another. Indeed, the time interval required for transmitting a packet is given by the sum of the preamble time, which lasts n_{ph} symbol times T, and the payload transmission time. Since each symbol codes i bits and a channel coding with rate $CR = 4/(4 + RDD)$ is applied (with redundancy bits $RDD = 1, ..., 4$), the time ToA_i required for transmitting a frame long P bytes with SF i can be expressed as $(n_{ph} + \lceil \frac{P \cdot 8}{i \cdot 4} \rceil \cdot (4 + RDD)) \cdot T$. Thus, the total uplink capacity results equal to $\sum_{i=7}^{12} G_i e^{-2G_i}$ and can be dramatically reduced (down to zero) as the loads G_i increase (up

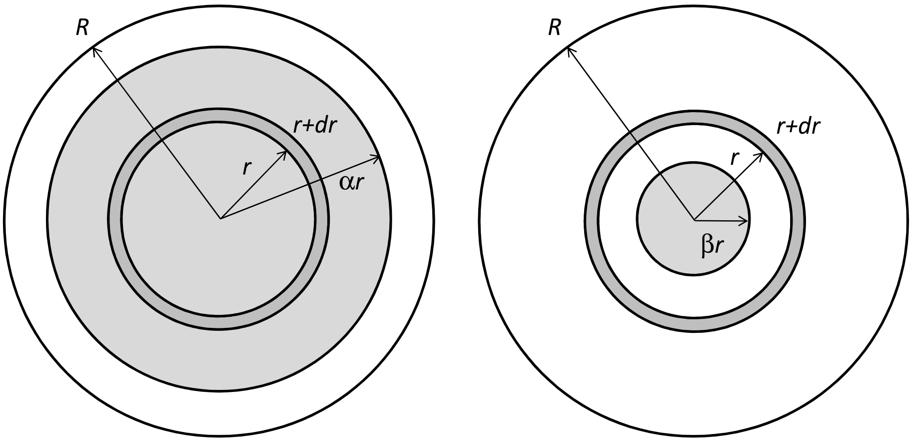


Figure 3.5. Traffic competing with receivers placed in the circular ring between r and $r + dr$ (dark gray area): distribution of intra-SF (left cell) and inter-SF (right cell) competing load.

to infinity).

Obviously, in a real cell the number n_i of EDs working on a given SF i is generally high but finite, and λ_i can be evaluated as $n_i \cdot s$, being s the source rate of each ED (which we assume to be constant for all devices). In absence of capture effects and inter-SF interference, the cell capacity is affected solely by the number of EDs configured on each SF, but it does not depend on the spatial distribution of the EDs within the cell (provided that all EDs are in the coverage range of the gateway).

3.4.2 Channel Captures

Consider first the case when collisions are only due to frames using the same SF (i.e. different SFs are perfectly orthogonal) and the spatial distribution of the devices is uniform in the whole cell area. As discussed in §3.3, LoRa modulation is very robust to interfering signals, and therefore it is very likely that the frames colliding at a given gateway result in the correct reception of the strongest one. For quantifying the performance improvements due to these events, we assume that in most practical cases a target ED can be interfered by a single colliding signal at a time. As we will show, this assumption is reasonable when the cell works in stable, non-congested conditions.

Capture effects have been observed when the reception power of the colliding signal is sufficiently lower than the power of the target ED (i.e. the

SIR of the target ED in dB is higher than a positive threshold, which we experimentally found equal to 1 dB, as shown in table 3.2). For a transmitter located at distance r from the gateway, neglecting the effect of random fading and assuming uniform transmission power among EDs and an attenuation law of type $r^{-\eta}$, this capture condition can be mapped into the placement of the interfering ED in a circular ring delimited by a distance αr, with $\alpha = 10^{SIR/10\eta} > 1$, and the cell radius R. An exemplary representation of this area in which the interfering ED does not prevent the reception of the target ED is shown in the left-most cell of Fig. 3.5 in white. Obviously, when αr is higher than R, such a region does not exist, because the transmitter is too far from the gateway and captures cannot occur. It follows that a target ED employing a given SF i is actually competing with a fraction of the total load G_i, which corresponds to the ratio between the area of the circle of radius $\min(\alpha r, R)$ and the total area of the cell. The smaller the α coefficient, the smaller the real competing load is. Note that in case of severe attenuation, higher values of η translate to lower values of α and higher chances of channel captures. Without loss of generality, in our numerical experiments we use $\eta = 4$.

To model the performance of LoRa in presence of channel captures, we extend the basic non-slotted Aloha model exploiting the above considerations: assuming for simplicity that all frames have fixed size with transmission time ToA_i, the throughput $S_c(G_i)$ in presence of captures obtained using SF i can be quantified by considering that the load offered in the circular ring between distance r and $r + dr$ will compete with the fraction $\min(\alpha^2 \cdot r^2/R^2, 1)$ of G_i:

$$S_c(G_i) = 2\pi \int_0^R \delta_i e^{-2\min(\frac{\alpha^2 r^2}{R^2},1)\cdot G_i} r \cdot dr \tag{3.3}$$

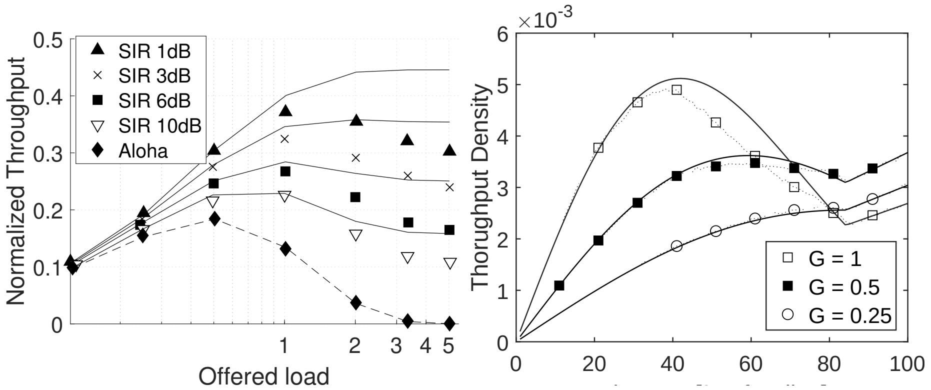
where $\delta_i = G_i/(\pi R^2)$ is the density of load offered to SF i. It results:

$$S_c(G_i) = \frac{1}{2\alpha^2}\left(1 - e^{-2\cdot G_i}\right) + G_i\left(1 - \frac{1}{\alpha^2}\right)e^{-2\cdot G_i}$$

and the DER can then be obtained as $S_c(G_i)/G_i$.

In the previous derivation, we generically refer to a uniform load density δ_i, while in real cells we have a finite number n_i of EDs usually placed at fixed positions. However, we can generally consider that n_i is sufficiently high and the throughput derivation can refer to the average results obtained in different realizations of node placements.

Fig. 3.6-a shows the throughput curves obtained by equation 3.3 as a function of the offered load G_i, for several SIR values (i.e. α values). In the

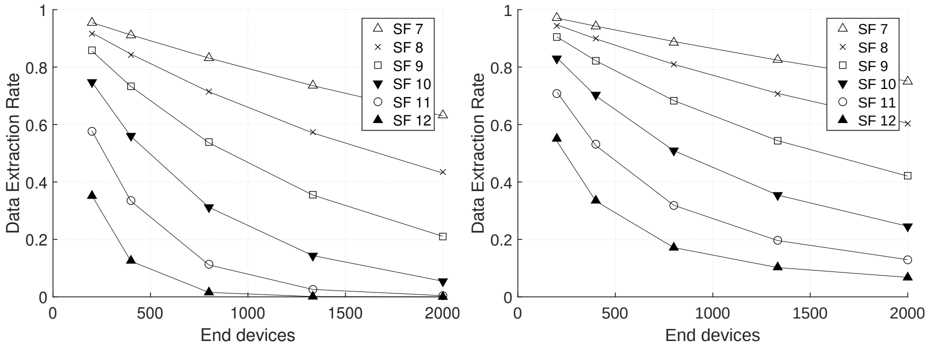


(a) (b)

Figure 3.6. Impact of single packet interference approximation on throughput with channel captures: theoretical model (lines) and simulation results including multiple packet interference (markers).

figure, the lines correspond to the analytical model (theoretical results), while the marker points represent the simulation results obtained with our Matlab simulator, in a simulation run of 9000 seconds. For deriving the average throughput results in simulation, we varied the offered load by adjusting the source rate of 1000 EDs and we randomly generated the position of each ED at each transmission attempt. The figure clearly shows that capture effects can significantly increase the maximum Aloha efficiency (up to about 300% for a capture SIR of 1dB when the offered load G_i is 1). Note that, our model works well in non-congested operating scenarios ($G_i < 1$), while it diverges from simulations in highly congested conditions, in which collisions involve multiple transmitted frames. Indeed, the asymptotic capacity of real systems tends to zero with the increase of the traffic load, while the model asymptotic value is different from zero (namely, it is equal to $1/2\alpha^2$).

To better visualize the effects of the load on the capacity approximation, Fig.3.6-b shows the throughput density (i.e. the integral argument in equation 3.3) achieved by EDs uniformly placed within the cell, as a function of the distance from the gateway (from 0 to 100% of the cell radius). Different curves refer to different G_i values: simulation results are shown with a dotted line and points, while our model results are plotted with solid lines. When the distance is higher than R/α and no capture effect is possible, we can easily recognize that the throughput density follows a linear distribution, because all EDs have the same success probability and the number of EDs



(a) Aloha (no capture)

(b) Capture

Figure 3.7. Simulation (markers) and analytical model (lines) results for channel capture effect.

considered in the integral grows proportionally to the distance. This last segments of the curve coincide with the throughput density of Aloha whose complete distribution is a linear distribution of the same slope in the whole interval $[0, R]$. For smaller distances, the capture effects can significantly increase the throughput density of Aloha. The figure also shows how our model overestimates the capture probability as the normalized load G_i increases: indeed, for $G_i = 0.25$ the points are perfectly overlapped with the solid line, while for $G_i = 1$ there is a region of the cell in which the real throughput is smaller than the one predicted.

Finally, we evaluated the DER achieved in a cell where all the nodes are configured on the same SF with and without the capture effects (for a capture SIR of 1 dB). Fig. 3.7 shows the DER results as a function of the number n of EDs configured on each SF, ignoring inter-SF interference. Each curve refers to an independent cell sub-system with the same number of EDs but different load conditions. The results demonstrate that the DER can increase significantly thanks to channel captures. For example for $n = 800$ EDs, using SF 10 the DER increases from about 0.32 in case of pure Aloha to about 0.52 in presence of captures. Obviously, for a fixed number of EDs n, lower SFs have the best DER because the resulting offered load is lower.

3.4.3 Imperfect Orthogonality Model

In order to quantify the impact of imperfect orthogonality among SFs on the cell capacity, we reasoned similarly to the previous section, by considering a single interfering signal at time. Because of imperfect orthogonality, a target ED working on SF i at a generic distance r will compete not only with the load G_i offered to the same SF, but also with a fraction of the load G_{-i} working with a SF different from i. Such a fraction corresponds to the EDs closer to the gateway, which generate an interfering signal whose power is much higher than the desired signal and exceeds the rejection capability of the LoRa receiver. Our experimental results showed that the SIR threshold under which interference rejection does not work is almost independent on the SF used by the interfering signal. Therefore, the minimum required SIR value can be mapped into the placement of the interfering ED in a cell sub-region delimited by a radius $\beta_i \cdot r$, with $\beta_i = 10^{SIR/10\eta} < 1$, as shown in Fig. 3.5 (right).

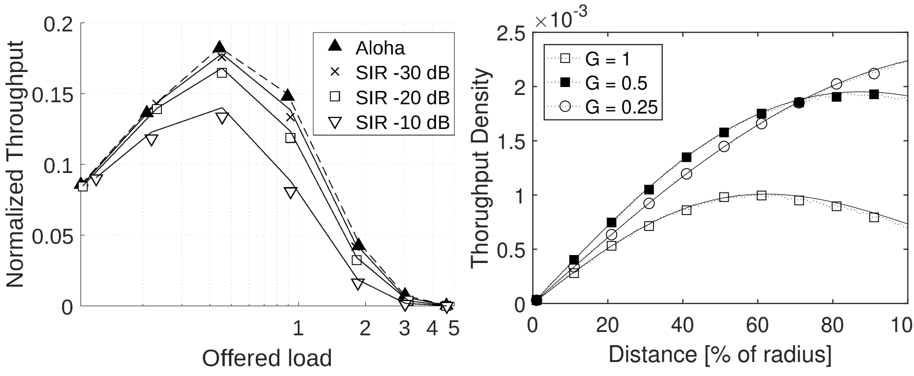
The analysis of collisions between frames transmitted with different SFs requires taking into account that frames have heterogeneous transmission times, even if they have a fixed size P. Consider first a simple scenario in which channel captures with frames transmitted with the same SFs are not possible. In this case, interfering signals are given by the totality of transmissions performed with SF i and with the fraction $\beta_i r^2/R^2$ of transmissions performed with other SFs. The success probability of a target ED employing SF i depends on the probability of finding the channel free from other interfering signals when starting frame transmission and during the following time interval ToA_i. The probability Pr_s of finding the channel idle at the starting time of frame transmission is given by the probability that no packet arrival at SF i is originated within a previous interval lasting ToA_i, while a fraction $\beta_i r^2/R^2$ of other packets employing a different SF $k \neq i$ has not started a transmission in an interval corresponding to the relevant frame transmission time ToA_k. Such a probability can be expressed as:

$$Pr_s(r) = e^{-\lambda_i ToA_i - \frac{\beta_i r^2}{R^2}\sum_{k\neq i}\lambda_k ToA_k} = e^{-G_i} \cdot e^{-\frac{\beta_i r^2}{R^2} G_{-i}}$$

being $G_{-i} = \sum_{k\neq i} G_k$ the total load offered by SFs different from i. The probability Pr_e that no other interfering signal is started until the end of the transmission time ToA_i can be expressed as:

$$Pr_e(r) = e^{-(\lambda_i + \frac{\beta_i r^2}{R^2}\sum_{k\neq i}\lambda_k) ToA_i} = e^{-G_i} \cdot e^{-\frac{\beta_i r^2}{R^2} G^*_{-i}}$$

with $G^*_{-i} = \sum_{k\neq i} \lambda_k \cdot ToA_i = \lambda_{-i} ToA_i$. The total throughput obtained on sub-channel i with imperfect orthogonality between SFs can be computed by integrating the success probability experienced at each distance r over all



(a) (b)

Figure 3.8. Impact of single packet interference approximation on throughput with non orthogonal SFs: model approximation (lines) and simulation results including multiple packet interference (markers).

possible distances as:

$$S_{qo}(G_i, G_{-i}) = 2\pi \int_0^R \delta_i Pr_s(r) \cdot Pr_e(r) r \cdot dr =$$

$$= e^{-2G_i} \cdot 2\pi \int_0^R \delta_i e^{-\frac{\beta_i^2 r^2}{R^2}(G_{-i}+G_{-i}^*)} r \cdot dr \tag{3.4}$$

It results:

$$S_{qo}(G_i, G_{-i}) = G_i e^{-2\cdot G_i} \frac{1 - e^{-\beta_i^2 (G_{-i}+G_{-i}^*)}}{\beta_i^2 \cdot (G_{-i} + G_{-i}^*)} \tag{3.5}$$

which is obviously smaller than the ideal orthogonal case $S_o(G_i, G_{-i}) = G_i e^{-2\cdot G_i}$.

Finally, if we want to take into account both the capture effects and the imperfect orthogonality of SFs, we can follow the same approach discussed so far and specify that the competing load for each target ED working on SF i is the sum of a fraction $l = \min(\alpha^2 r^2/R^2, 1)$ of the intra-SF load and a fraction $\beta_i^2 r^2/R^2$ of the inter-SF load:

$$S_{qo,c}(G_i, G_{-i}) = 2\pi \int_0^R \delta_i e^{-\frac{\beta_i^2 r^2}{R^2}(G_{-i}+G_{-i}^*)} e^{-2lG_i} r \cdot dr$$

which leads to:

$$S_{qo,c}(G_i, G_{-i}) = \frac{G_i e^{-2G_i} \left(1 - e^{-2G_i - (G_{-i}+G_{-i}^*)\alpha^2/\beta_i^2}\right)}{2\alpha^2 G_i + \beta_i^2 (G_{-i} + G_{-i}^*)} +$$

$$\frac{G_i e^{-2G_i}}{\beta_i^2(G_{-i}+G^*_{-i})}\left(e^{(G_{-i}+G^*_{-i})\alpha^2/\beta_i^2} - e^{(G_{-i}+G^*_{-i})\beta_i^2}\right) \tag{3.6}$$

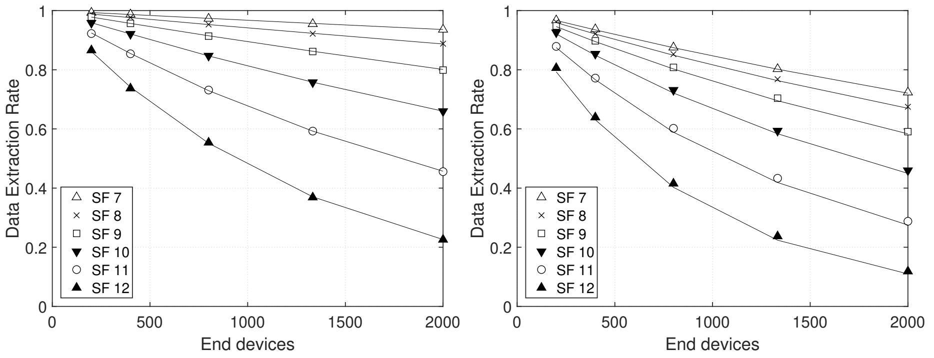
The average DER can then be computed by dividing the throughput with the total offered load.

To show the impact of inter-SF collisions, fig. 3.8-a shows the theoretical throughput curves (lines) obtained by equation 3.5 (no captures) for different values of the rejection SIR (i.e. β_i values), together with simulation results (points), when two different SFs coexist in the same cell (namely, SF 7 and SF 9). The curves refer to the throughput of EDs configured on SF 7. Also in this case, for deriving the average throughput results in simulation, we considered a fixed number of 1000 EDs (half configured on SF 7 and half on SF 9), varying the ED positions at each transmission attempt and tuning the source rate with increasing load. As shown in the figure, despite the pseudo-orthogonality of the SFs, the throughput can indeed be severely affected compared to the ideal Aloha without inter-SF collisions (almost 50% reduction in case of congested scenarios, i.e. $G_i \approx 1$). Fig. 3.8-b shows the throughput density (i.e. the integral argument in equation 3.4) achieved by EDs using 2 SFs with a rejection SIR of -10 dB and uniformly placed within the cell, as a function of the distance from the gateway (from 0 to 100% of the cell radius). Different curves refer to different G_i values: for low load conditions and small distances, the throughput density follows an almost linear distribution typical of Aloha with uniform EDs. However, for higher load conditions the curves can significantly deviate from a linear function, especially at high distances where inter-SF interference can be more critical.

Fig. 3.9 shows the DER results as a function of the total number N of EDs active in the cell, under the assumption that such a number is equally shared between different SFs (i.e. each SF is assigned to $n_i \approx N/6$ EDs) and same source rate s for all EDs. Again, markers represent simulations and lines the analytical results, which are still remarkably close to the simulations. From the figure, it is clear that the impact of non-orthogonality can be severe. The performance deteriorates quickly as the number of EDs increases.

3.5 Cell configurations

In this section, we discuss the impact of different cell configurations that can be considered for optimizing the cell capacity when multiple SFs are available. For sake of presentation, we consider two SFs only, namely SF



(a) Orthogonal SFs (b) Interfering SFs

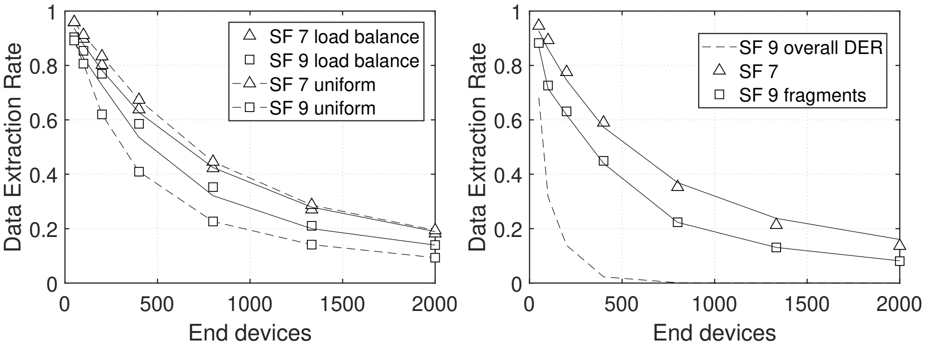
Figure 3.9. Simulation (markers) and analytical model (lines) results for interfering SFs.

$i <$ SF j, but generalizations to multiple SFs are straightforward. From our previous considerations, it is evident that LoRa cell performance is a function of the arrival rates of packets working on SF i and SF j, but also of their placement within the cell, because the inter-SF and intra-SF interference power experienced in case of collisions also depends on the relative position of the EDs.

3.5.1 Load balancing

Assuming that a total arrival rate of Λ pk/s is uniformly distributed within the cell, the arrival rates λ_i and λ_j experienced in each SF (with $\lambda_i + \lambda_j = \Lambda$) depend on the SF selected for each packet transmission. The selection of the certain SF may be a local decision, such as the selection of the highest possible rate compatible with the link budget available at a given spatial position, or may be extended with load balancing considerations. If the link budget at the cell border is enough for transmitting at the highest rate corresponding to SF i, the first approach would lead to $\lambda_i = \Lambda$ and $\lambda_j = 0$. Load balancing can significantly improve the overall cell capacity by reducing the load on SF i and by exploiting the additional capacity available on SF j.

A balancing solution devised to provide fair performance to devices working on different SFs is equalizing the offered load $G_i = G_j$. Note that this is different from equalizing the number of transmissions performed at each SF, because packet transmission times vary as a function of the employed SF. Specifically, load balancing is achieved for $\lambda_i \cdot ToA_i = \lambda_j \cdot ToA_j$, i.e.



(a) Load balance (b) Fragmentation

Figure 3.10. Impact of load balancing (left) and fragmentation (right) to increase fairness among the SFs.

$\lambda_i = \Lambda \cdot \frac{ToA_j}{ToA_j+ToA_i}$. This implies that in a real cell with a finite number N of EDs, if all EDs work with uniform source rates, the same proportion is applied for deriving n_i as $N \cdot \frac{ToA_j}{ToA_j+ToA_i}$. However, perfect load balancing is not always feasible, because some EDs distant from the gateway can be forced to work on the most robust SF for guaranteeing that the received power is above the reception threshold.

Fig. 3.10(a) compares the DER obtained in a cell with a total number N of EDs and two SFs available (namely, SF 7 and SF 9). Simulation results are shown with markers and model results with lines. The figure shows results for different criteria on SF allocations: equally sharing the number of EDs between SF 7 and SF 9 (dashed lines) or allocating a number of EDs proportionally to the relevant airtimes, with n_7 about four times of n_9 (solid lines). From the figure, it is evident that the second choice, i.e. load balancing, can be an effective solution for providing a similar DER to all devices, regardless of the allocated SF. However, DER performance are not exactly the same because inter-SF interference is not symmetrical due to different rejection thresholds (i.e. β_i and β_j coefficients) and transmission times (which result in $G^*_{-i} < G^*_{-j}$ even in case of load balancing). Overall, the most robust SF j suffers an higher inter-SF competing load.

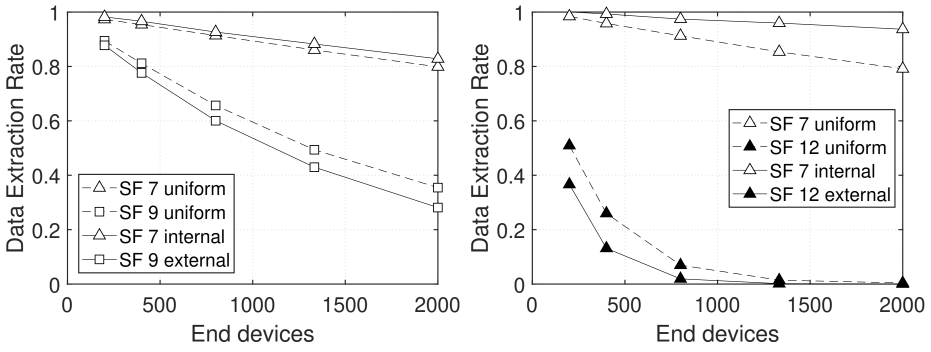
An additional mechanism to be considered for improving the fairness of the system could be the use of fragmentation. The idea is to equalize the airtimes of packets transmitted at different SFs. Obviously, in such a case, the arrival rates of fragments grow proportionally to the number of per-packet fragments. Fig. 3.10(b) shows the performance results in the same scenario described for Fig. 3.10(a), with n_7 about four times n_9, and 4 fragments are used for packets transmitted with SF 9. From the figure, it is clear that fragmentation is not effective for equalizing the performance of the EDs, due to the additional overhead added to each fragment. Additionally, we have to consider that now four fragments are required for reassembling a single packet at SF 9. Thus, the overall DER (dashed line) is worse than the previous case. This is also due to the fact that LoRa technology does not easily support selective re-transmissions of corrupted fragments, because the downlink channel from the gateway to the EDs would result congested by the transmission of the feedback messages. Therefore, fragmentation without selective re-transmissions does not bring benefits to the network.

3.5.2 Spatial allocations

For a fixed number of EDs n_i (or n_j) to be configured on SF i (or SF j), different allocation choices are possible. EDs working on the same SF can be selected uniformly within the whole cell area or can be restricted to a specific area of the cell. According to the position of the selected nodes, each allocation policy can be mapped into the opportunistic definitions of the $\delta_i(r)$ and $\delta_j(r)$ functions. In order to predict the cell capacity resulting from a specific allocation, we can generalize previous throughput derivations for dealing with non uniform load density functions.

As an illustrative example of model extension under generic $\delta_i(r)$ functions, we consider the case in which nodes employing different SFs are placed in different sub-regions of the cell, rather than being uniformly spread in the whole area. A choice could be allocating the most robust SF to far users, placed in a circular ring between distance $d > 0$ and R, and the less robust SF to users closer to the gateway within a maximum distance d, in order to maximize the reception margin of each ED. The distance d can be chosen for achieving the desired load balancing. In such a case, the density functions of nodes employing different SFs can be defined as: $\delta_i(r) = G_i/(\pi d^2)$ when $r \leq d$ (and 0 otherwise), and $\delta_j(r) = G_j/(\pi R^2 - \pi d^2)$ when $r < d$, and 0 otherwise.

Allocating EDs in circular rings have different implications for both the inter-SF and the intra-SF interference. Regarding the first aspect, by neglecting the fading effects, it never happens that users employing spreading factor SF i are interfered by users employing SF j with a higher interfering power (being these users deterministically located at higher distances). In other words, the throughput achieved on spreading factor SF i is given by



(a) (b)

Figure 3.11. Performance of SF9 (a) and SF12 (b) when competing with SF7. Comparison between uniform distribution of EDs (dashed lines) and when higher SFs are allocated to far away EDs (solid lines). The inter-SF SIR threshold is 10 dB and $\eta = 4$.

$S_i = G_i \cdot e^{-2\cdot G_i}$, i.e. the density of potentially interfering signals is equal to zero as in case of perfect orthogonality. Conversely, users employing spreading factor SF j are more likely affected by interference generated by users employing spreading factor SF i, because in case of collisions the interfering signals are concentrated in the cell area closer to the gateway, which results in a interference density higher than the previous case.

Being $\delta_{-i}(r) = \sum_{j \neq i} \delta_j(r)$ and $\delta^*_{-i}(r) = \sum_{j \neq i} \delta_j(r) \frac{ToA_i}{ToA_j}$, the throughput in absence of channel captures becomes:

$$S_{qo}(i) = e^{-2\pi \int_0^R 2\delta_i(r) r \cdot dr} \cdot 2\pi \int_0^R \delta_i(r) \left[e^{-2\pi \int_0^{\beta_i r} (\delta_{-i}(t) + \delta^*_{-i}(t)) t \cdot dt} \right] r \cdot dr \tag{3.7}$$

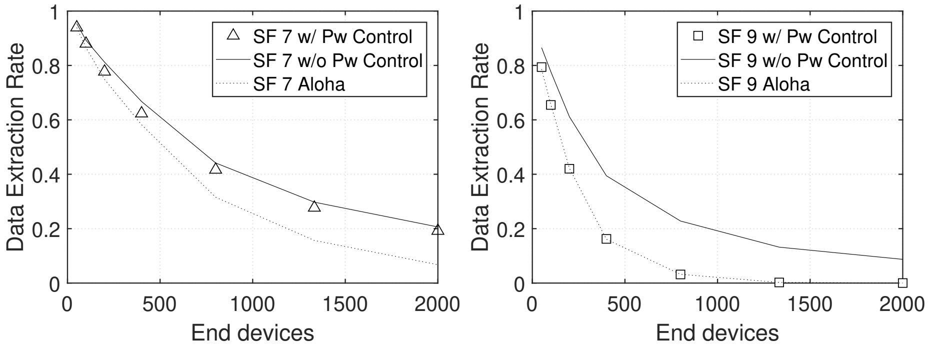
Figures 3.11(a) and 3.11(b) compare the performance of SF 7 when competing with SF 9 and SF 12 respectively, in scenarios with uniform allocation of SFs in the cell (dashed lines) or when far distance EDs are using the highest SF 9 and SF 12, while SF 7 is used for EDs close to the gateway. The figures show that the performance of the higher SFs is deteriorated when allocating them to faraway EDs, while SF 7 performance improves because of the absence of inter-SF collisions. This, demonstrates that allocating higher SFs to far distance EDs is detrimental more than beneficial, because the performances are highly affected by closer devices while fading has a much lower impact. In other words, although higher SFs improve the robustness to fading and allow longer distances, this comes at the cost of an increased airtime of the transmitted frames and, since LoRa uses the Aloha access protocol, collisions arise quickly when increasing the offered load. Clearly, using low SFs for long distances might cause unnecessary retries, or even no packet delivery at all in case the link budget is not sufficient.

SF allocation has also an impact on the capture probability. For a given number n_i of EDs working on SF i, the highest capture probability is achieved when nodes are spread in the whole cell (rather than in a circular ring), because this choice corresponds to the spreading of the RSSI values of potentially colliding signals (which may result in the correct reception of the strongest signal). In the limit case in which all nodes working on SF i are at the same distance from the gateway, no capture effect can occur.

3.5.3 Power control

Another important configuration parameter of LoRa cells is the transmission power of EDs, which can be tuned by means of specific control messages sent by the gateway. The message specifies the power to be used in terms of a reduction margin to be applied to the maximum possible power (which may vary in different countries); the reduction margin is coded in steps of -2dB from 0 to -14dB. The tuning of the transmission power can be considered for reducing the energy consumption of devices which are close to the gateway, but also for mitigating the impact of inter-SF interference, Indeed, orthogonality of different SFs can be guaranteed in case the difference between the reception powers of EDs working at different SFs is lower than the minimum margin in table I (about -8dB).

To model the impact of power control on the LoRa cell performance, we can consider that each power reduction applied to a specific ED is equivalent to moving the device at an higher distance from the gateway. If η is the propagation loss coefficient, every step of -2dB corresponds to a distance increment of a factor equal to $\gamma = 10^{2/10\eta}$. In case power control is used for equalizing the reception power of the EDs, taking into account that only -14dB are available at steps of -2dB and the distance r_{min} of the closer ED, we could equivalently consider that no ED is placed at a distance lower than $d = 10^{14/10\eta} \cdot r_{min}$ and that EDs originally placed within the circular area of radius d are moved in the circular ring between d and $d \cdot \gamma$. In other words, power control can be evaluated again by working on the definition of the load density function $\delta_i(r)$ as follows: 0 when $r < d$, $G_i/(\pi R^2) \cdot [1 + 1/(\gamma^2 - 1)]$ when $r \in [d, d \cdot \gamma]$, and $G_i/\pi R^2$ when $r > d \cdot \gamma$. Equivalently, we can define $\delta_j(r)$ and derive the cell capacity S_{qo} of each SF applying equation 3.7.



(a) SF7 (b) SF9

Figure 3.12. DER with two interfering SFs with and without using transmission power control.

As an illustrative example, we consider a cell using only SF 7 and SF 9, with half of the EDs using SF 7 are placed close to the gateway and the others using SF 9 are far away from the gateway. For clearness of presentation, Fig. 3.12 reports separately in two plots the DER achieved by the two SFs, although in the experiment both SFs are used as explained above. From the figure, it is evident that when equalizing the received power, the DER of SF 9 is close to the Aloha model (no channel captures) and much lower than the DER obtained without power control, while for SF 7 there is almost no change. This means that tuning the transmitted power of the EDs in order to equalize the received power at the gateway is detrimental more than beneficial for performance because, in case of collision between packets of the same SF, it reduces the probability of capturing the channel.

3.6 LoRa Multiple Gateway

Thanks to the capture effect, the capacity of a LoRa cell can be increased by deploying multiple gateways. Indeed, each gateway sees a given ED with a different distance and therefore, in case of collisions, experiences a different power ratio between the strongest received packet and the interfering signals. When the power ratios are higher than the capture threshold, the collisions can result in the correct reception of a number of packets equal at most to the number of gateways. Obviously, it may also happen that the same packet is correctly received by multiple gateways, but all the packets are forwarded to a common network server, which discards duplicated packets.

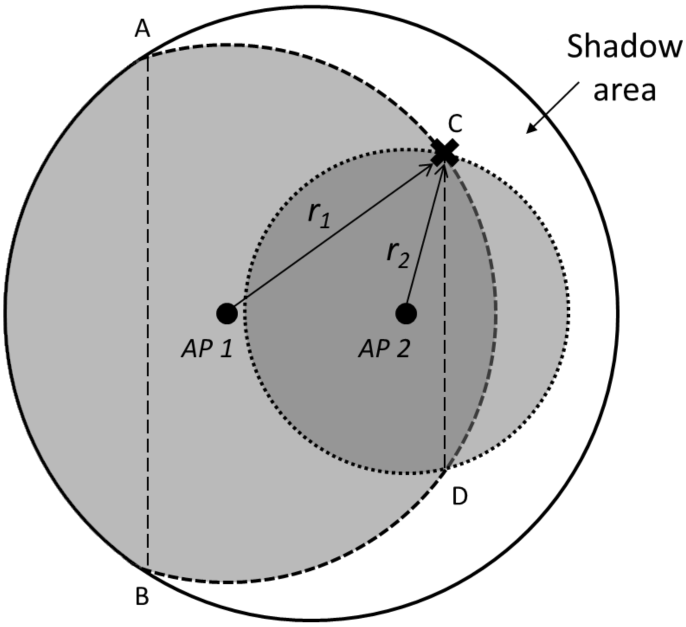
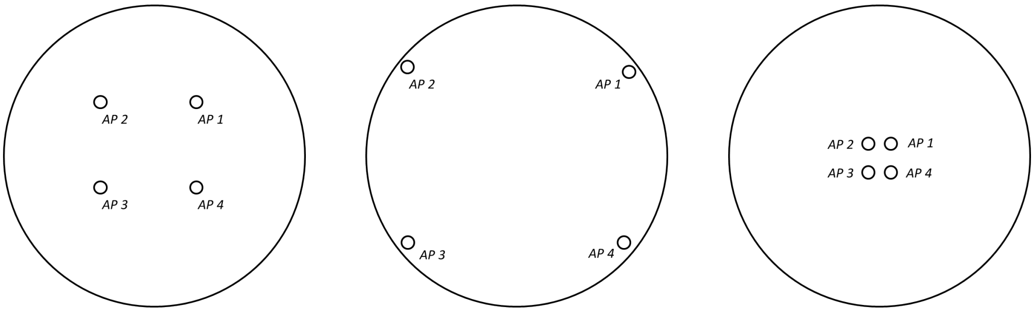


Figure 3.13. Competing area (gray) in presence of multiple gateways and "shadow area" created by the ED in position C.

For deriving the average cell capacity in presence of multiple gateways, it is necessary to specify the positions of the gateways within the cell, because the capture probability depends on the relative distance between the potential receivers (i.e. the gateways) and the transmitters. Differently from the single gateway case, even under the assumption that only a colliding signal is experienced at a given time, the EDs competing with a target transmitter are different for each gateway and distributed in regions which cannot be modeled as simple circle areas. For example, Fig. 3.13 shows the scenario of a cell with two gateways, $\alpha = 1$ and uniform distribution of EDs: the target transmitter in position C experiences different competing loads at each gateway, which are proportional to the intersection area between a circle of radius αr_m and the cell, being r_m the distance between the ED and a generic gateway $m = 1{,}2$. Although the areas can be computed as circular areas delimited by the chords AB and CD, it is not easy to generalize the approach proposed in §3.4, because the performance of the target transmitter cannot be averaged as a function of the distance from the cell center, being also affected by the direction (i.e. by the specific position within the cell). The figure also shows a fraction of the cell from which no other ED can transmit successfully a packet overlapping with the target transmitter, being the target transmitter



(a) Grid (b) Edge (c) Center

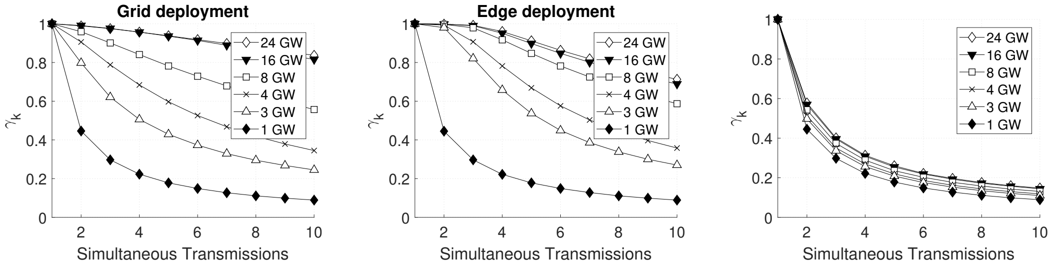
Figure 3.14. Multiple gateway topology used in the experiments: example with 4 gateways.

closer to all the available gateways.

For a specific placement of the gateways, we modeled the geometry of the cell under uniform distribution of the EDs and complete coverage (i.e. assuming that EDs can transmit to a gateway placed at distance $2 \cdot R$) by numerically evaluating the average probability γ_k that a target ED successfully transmits its packet to at least one gateway, in presence of $k \geq 0$ interfering EDs. By considering one interfering signal at time as in the single gateway derivation, such a probability has been computed by averaging (on all possible transmitter positions) the probability that, for at least one gateway, k interfering EDs are at distances higher than α times the distance of the target transmitter. The analysis of the capture probability in presence of k interfering EDs allows us the decoupling between the geometric effects (due to the specific gateway placement) and the interference probability (due to the cell load). In other words, assuming that the γ_k coefficients are known, the cell capacity in presence of multiple gateways can be computed for any possible load G_i as:

$$S_c(i) = G_i \cdot \left[\sum_{k=0}^{\infty} \left(\gamma_k \frac{(2G_i)^k \, e^{-2G_i}}{k!} \right) \right]. \tag{3.8}$$

Note that, when $SIR = 0$ dB (i.e. $\alpha = 1$) and a single gateway is placed in the center of the cell, then $\gamma_k = 1/(k+1)$, because for any placement of $k+1$ EDs, only one ED will result closer to the single gateway (assuming negligible the probability of extracting two EDs on the same distance to the gateway). In this condition, it is easy to see that equation (3.8) gives the same result of equation (3.3).

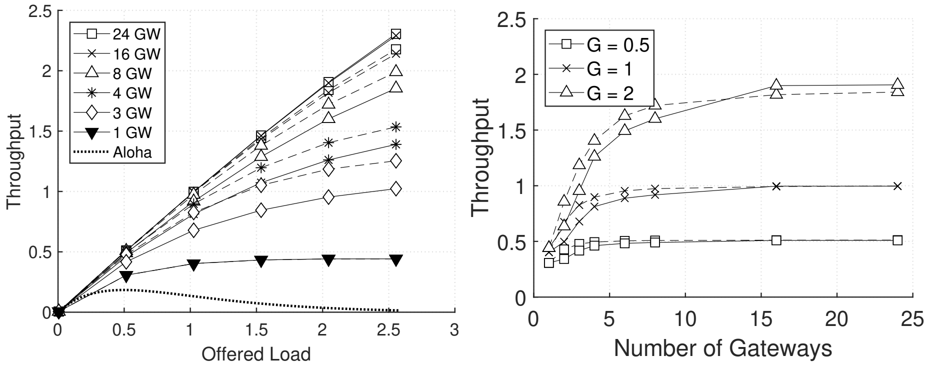


(a) Grid (b) Edge (c) Center

Figure 3.15. Impact of gateway deployment: values of γ_k with $\eta = 4$, SIR= 1dB and different number of gateways on a grid deployment (a), at the cell edge (b), or when the gateways are all concentrated close to the center of the cell(c).

Obviously, the γ_k coefficients depend not only on the number of gateways but also on their specific position within the cell. Since a closed form derivation is not generally possible, we evaluated the γ_k coefficients numerically for specific gateway positions. To this purpose, we randomly generated $k+1$ EDs uniformly distributed in the cell, quantified the ratio of EDs whose distance r from at least one gateway was α times smaller than the distance between the other k EDs and the same gateway, and averaged results over multiple random placements. In our evaluations, we analyzed different deployment strategies: placing the gateways at a regular grid, on the cell edge, or concentrated very close to the cell center. For the grid deployments, we kept a central symmetry towards the cell center, by equally spacing the gateways in the two cell dimensions; for the edge deployment, we equally spaced the gateways along the cell circumference, while for the last setting we created a small grid of size $R/10$.

Capture probability. Fig. 3.15 shows the γ_k results numerically obtained by placing a varying number of gateways with the three deployment strategies explained above: from the figure, it is easy to see how γ_k coefficients are strongly dependent on the gateway placements. This is particularly evident in the last case: when the gateways are too close to each other, the performance does not improve in comparison with the single gateway cell. Indeed, the space diversity between gateways is poor and offers little opportunities for increasing the channel captures. For the other deployment solutions, we also observe that coefficients result generally higher when the gateways are placed on the cell edge rather than on a regular grid. Only when the number of gateways is high (e.g. 16 or 24) the grid topology is better, although the



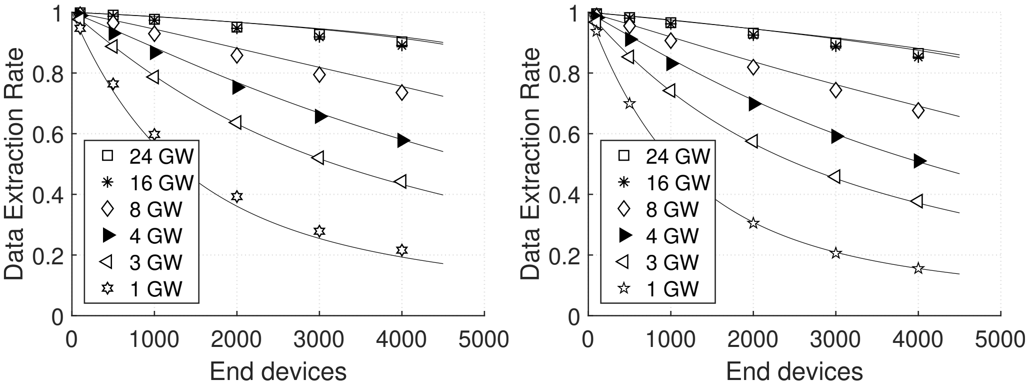
(a) (b)

Figure 3.16. Grid deployment (solid) vs. edge deployment (dashed) when varying the offered load or the number of gateways.

improvement on the capture probability is marginal. Although not intuitive, this result can be justified by considering that (as long as coverage is guaranteed) the more the gateways are sparse, the more channel captures can be achieved. For example, for $k = 1$ (one interferer, 2 simultaneous transmissions) and 3 gateways, the average number of transmissions correctly received by at least one gateway is 0.999 for gateways deployed on the cell edge and 0.804 for gateways placed in a regular grid. As depicted in Fig. 3.13, on a grid topology a transmitting ED can create a "shadow area" towards the cell edge, impeding any other ED in this area to successfully capture the channel.

Throughput. Fig. 3.16 quantifies the cell throughput by using equation 3.8 and the γ_k coefficients derived in our numerical evaluations, for the grid (solide lines) and edge (dashed lines) deployment strategies. The figures clearly show that increasing the number of gateways allows to achieve a throughput almost equal to the offered load. Obviously, higher capture probabilities are mapped into higher throughput results. For example, with 3 gateways the throughput achieved with the edge deployment is about 25% higher than the one achieved with the regular grid, while for a number of gateways equal or higher than 16 the throughput gain of the grid deployment is about 5%. We can expect that these capacity differences under different gateway deployments can be reduced in case of multi-cell systems.

In order to validate our model, Fig. 3.17 compares our throughput predictions with simulation results. For space reason, we only show results for the regular grid case (but conclusions are similar for the edge deployment case). The figures have been obtained by considering the availability of one



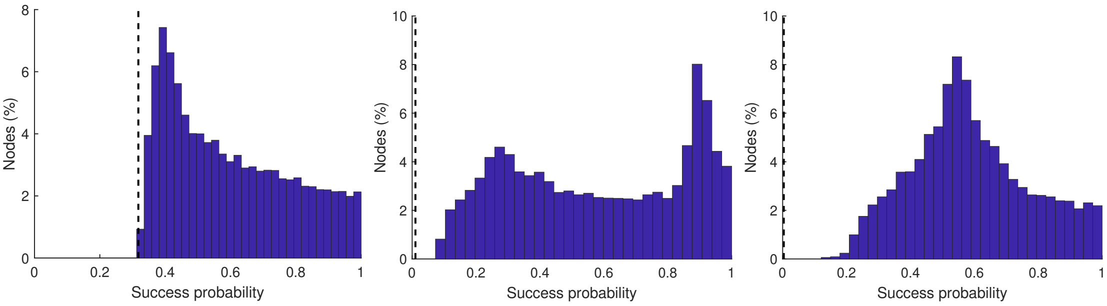
(a) SIR=1dB.

(b) SIR=3dB.

Figure 3.17. DER using multiple gateways in a grid topology: model (lines) and simulation results (points) for SIR values of 1 and 3 dB.

SF only (namely, SF7) and two different capture thresholds (1 and 3 dB). The figure allows to draw some interesting conclusions. First, despite of the simplification assumptions used for modeling the capture effects, the model is in agreement with simulations in all the considered scenarios, thus validating equation 3.8 and the γ_k derivation. Second, when the capture effects are very common, a cell with M gateways and a number of devices equal to N provides similar performance of M independent systems loaded with N/M devices. For example, the DER achieved with 4 gateways and 2000 EDs results slightly less than 0.8 and comparable with the one achieved by a single gateway with 500 EDs or 8 gateways and 4000 EDs. Third, the capture threshold has a much lower impact than in a single gateway case, thanks to the increased capture possibilities provided by the spatial diversity of the gateways. Finally, although results always improve as the number of gateways increases, at a given point the improvements are marginal (DER results with 16 or 24 gateways are almost comparable).

Fairness. The position of the gateways has also a significant impact not only on the average cell performance, but also on the fairness of the network, i.e. on the spreading of the DER results achieved by different EDs. Fig. 3.18 shows the success probability histogram of the ED's transmissions in various scenarios characterized by *same* DER of 0.6 but different number of gateways, EDs and deployment strategy. In particular, the figure shows what happens in a cell with 1 gateway in the center or with 4 gateways, in a regular grid or at the cell edge. To obtain the same DER of 0.6, the number of EDs



(a) 1 GW Center (b) 4 GW Grid (c) 4 GW Edge

Figure 3.18. Impact of gateway deployment on fairness: distribution of the success probability among EDs in a cell with 1 gateway in the center (a) and with 4 gateways in a regular grid (b) or at the cell edge (c).

was 1000, 4000 and 4900 respectively and the corresponding offered load was 0.57, 2.29 and 2.81 (these values were obtained experimentally to equalize the capacity and achieve equal DER for all scenarios). From the figure, it is clear that with only one gateway most of the EDs have low performances, while in a grid topology the distribution is bimodal, with a group of EDs that transmit with very high success probability (higher than 80%) and another group which suffer low performance (around 30% successful frames). Instead, the edge deployment tends to be more fair, in the sense that most of the EDs experience a success probability around 0.6, corresponding to the cell average DER. The figure also shows the minimum theoretical success probability of the pure Aloha system (given by e^{-2G} and marked by the dashed lines), which is obviously different for the three scenarios because of the different load levels. Note that the capture effects can improve the Aloha results for some devices, without penalizing the ones with low capture probabilities.

Chapter 4

Signal transmitter localization using low-cost SDR receivers

4.1 Introduction

Tracking of indoor positions is one of the key enabling technologies for various applications. The most famous location system is the Global Positioning System (GPS), but it is well known that it does not work well in indoor environments, because the satellite signal energy is easily affected by the obstacles of an indoor environment, and it gets down rapidly below the receiving threshold [48]. In wireless indoor localization systems, distance-based positioning methods that use Time of Arrival (TOA) [49], Time Difference of Arrival (TDOA, [7, 8]) [50], Angle of Arrival (AoA) [51] and Received Signal Strength (RSS) [52] are often used. The methods AOA and RSS are affected by the multipath effect of indoor wireless signal transmission, which greatly degrades performance [53]. Compared with TOA, TDOA localization method only requires clock synchronization between receivers, and does not need to add additional clock synchronization equipment on the transmitter.

The objective of this work is to provide a method for localization of the transmitter based on TDOA, drastically reducing errors by using appropriate signal processing techniques. This method can be used for creating an experimental indoor localization system based on low-cost SDR receivers, to then later integrate it into the ElectroSense [9, 10] network, which is a crowdsourcing initiative to collect and analyse spectrum data. The data collected can be of whatever transmitter, even unknown. Therefore, not having information on the transmitter clock, and being in an indoor environment, for the reasons explained before, we decided to use the TDOA technique. Although the TDOA of a signal received at a spatially distributed sensors array is one of widely used schemes for transmitter localization, there are some aspects which can be improved.

The remainder of this Chapter is organized as follows: after literature review in §4.2, we introduce the transmitter localization system in §4.3, also describing the problems of the TDOA and our solutions. A first version of the system is described in §4.4, where the experimental results are presented. Finally, we discuss about studies on possible future improvements in §4.5 and §4.6.

4.2 Related work

Indoor transmitter localization systems have been researched for several decades. A number of methods exist, varying in approaches, technologies, location accuracy, infrastructure, signals used, etc. Our method is based on a TDOA approach, it can work on a wide spectrum and with whatever type of radio signal (ie whatever modulation), targets an accuracy of the order of the meter and it is based on a low-cost infrastructure, i.e. it requires a minimum of three low-cost receivers, for a total cost less than a hundred dollars.

The most accurate results in field on indoor localization are achieved by laser based methods. The best systems achieves an accuracy of few cm using few lasers and multiple high-end cameras. These systems cost thousands of dollars. However, low-cost high-resolution cameras exist, such as smartphones featuring Google's Project Tango hardware that have an accuracy of a few centimeters [54]. At the expense of low cost and high resolution, they require dedicated hardware and a training phase with very long times, especially for large rooms or even a whole building. Still in the field of light localization techniques, approaches based on RSS and AOA have been used. An example of RSS based localization with light is shown in [55], where higher accuracy is achieve compared to WiFi RSS techniques [56]. Performing AOA localization is even more accurate and can yield a positioning error of tens cm [57].

Some indoor localization methods are not based on traditional signals (WiFi, GSM, 5G, Bluetooth, etc.), but rather on ultrasound, such as SpiderBat [58], or on RFID-based technologies, of which we find an overview in [59].

Both systems are relatively inexpensive, but require dedicated hardware. Indeed, due to the short range of both systems, many devices are required for a localization system serving a large areas.

Although the GPS does not work well in indoor, some works use the same principle as GPS localization, for example employing signals transmitted by an aircraft. For safety reasons, airplanes repeatedly transmit their location, like GPS satellites. Their position is transmitted with signals called ADS-B, that are strong enough to be received indoors, do not require dedicated hardware to be received. In [60] an indoor localization system exploiting aircraft signals is implemented. Their system however needs a training phase with long times and achieves an accuracy lower than GPS (approximately 25 meters).

The most popular works in field of indoor localization are achieved by exploit traditional wireless signals. Numerous solutions demonstrate that Bluetooth [61, 62, 63], Wifi [64], ZigBee [65], GSM [66, 67, 68], are feasible to room-level localization. These solutions show lower accuracy and require a training phase with long times, unlike our method which, as we will show, requires a very fast training phase.

Finally, at the expense of simplicity, which is a advantage of our work, other solutions exploit multiple signals or technologies for improving accuracy. Considering that the environment may contain more kinds of signals, they can be exploited together with WiFi for improving indoor localization. In addition to the WiFi signal, Chen et. al. in [69] also utilize three kinds of signals including TV, FM and AM. Fang et. al. in [70] utilize DVB-T, GSM, FM and WiFi signals. Carvalho et. al. in [71] utilize DVB-T, FM and WiFi signals for realize an indoor localization system based on Software Defined Radio (SDR). WiFi-based indoor localization methods can also be supported by a sensor. Sensor assisted localization methods are particularly favored in smartphone applications, because basically all of these devices feature a gyroscope, a compass and an accelerometer. For example, cars driving into a tunnel will lose GPS signals, but based on measurements of the current speed, their position can be estimated until the tunnel ends [72].

4.3 Transmitter localization System

Our transmitter localization system, as mentioned, is based on widely employed technique known as TDOA. A signal emitted by an unknown transmitter, and received by multiple receivers located in different positions, will arrive at different times due to the different distances between transmitter and receivers. This difference in arrival time is called TDOA and can be measured between each pair of receivers. A TDOA value, hereinafter referred T_v, can be expressed as a distance by multiplication with propagation speed. This distance represents how much the transmitter is distant from the pair of receivers considered. Finally it is possible to apply multilateration, ie to calculate the region of space in which the unknown transmitter is located, it will be a hyperbola, because the region of space that is x from a receivers away and $x \pm T_v$ on the other is precisely a hyperbola. With minimum 3 receivers, then 3 pairs, then also 3 hyperbolas, it is possible to obtain, through the intersection of the hyperbolas, the position of the transmitter. The more receivers you have available, the higher the resolution. With 4 receivers properly positioned, it is already possible to have information on the height of the transmitter, thus making 3-D localization possible.

Although the basics of TDOA are quite simple, there are several aspects to be addressed: i) Local Oscillator (LO) offset correction; ii) Synchronization between devices; iii) Signal Processing for TDOA.

4.3.1 Local Oscillator offset correction

It is important that the receivers are tuned to exactly the same frequency and that they have the same sampling rate. It is known that the SDR receivers have not a perfect accurate local oscillator (LO) [73]. As we will show in Sec. 4.4, for experimental validation of the system we use low-cost Software Defined Radio (SDR), called RTL-SDR, in particular RTL2832U, which have a LO offset declared in the specifications of ±1 ppm. The main side effects of the LO offset are represented by a central frequency and a sampling rate that are not exactly those desired, effects that therefore need to be corrected.

The most popular software for the LO offset estimation are Kalibrate-RTL[1], rtl_test[2], LTE-Cell-Scanner[3] and LTESS-track[4]. We have chosen to use the most recent, LTESS-track, as in [73] it is shown that it has better performance than the others software. LTESS-track presents a very short processing times, less than 1 second, against about ten seconds for Kalibrate-RTL or even several minutes for rtl_test. In terms of errors in case of LO

---

[1] https://github.com/steve-m/kalibrate-rtl

[2] https://github.com/steve-m/librtlsdr

[3] https://github.com/Evrytania/LTE-Cell-Scanner

[4] https://github.com/electrosense/LTESS-track

with high offsets, LTESS-track present a more accurate estimate than LTE-Cell-Scanner and Kalibrate-RTL, that even does not work in these cases.

LTESS-track uses a LO offset evaluation method that relies on the Synchronization Signals (SS) transmitted by LTE base stations as reference. The output is the Local Oscillator Offset, hereinafter referred φ.

To correct the central frequency of the received signal, the property of frequency translation can be applied. Given f'_c real central frequency of signal received s and f_c desired central frequency, it is $f_c = (1+\varphi)f'_c$. Let $\Delta f_{LO} = \varphi f'_c$ the difference between the real and desired central frequency, caused by the LO offset, the correct signal is just:

$$s'\left(nT\right) = s\left(nT\right) \cdot e^{j2\pi\Delta f_{LO} nT} \tag{4.1}$$

where nT represents discrete time, so $t = nT$, n is the sample number and T is the sampling time, so $T = 1/f_s$.

To correct the imperfect sampling rate of the receiver, ideally, if the received signal were continuous, it would be enough to take the samples at the time instants corresponding to the desired sampling rate. So, in the continuous:

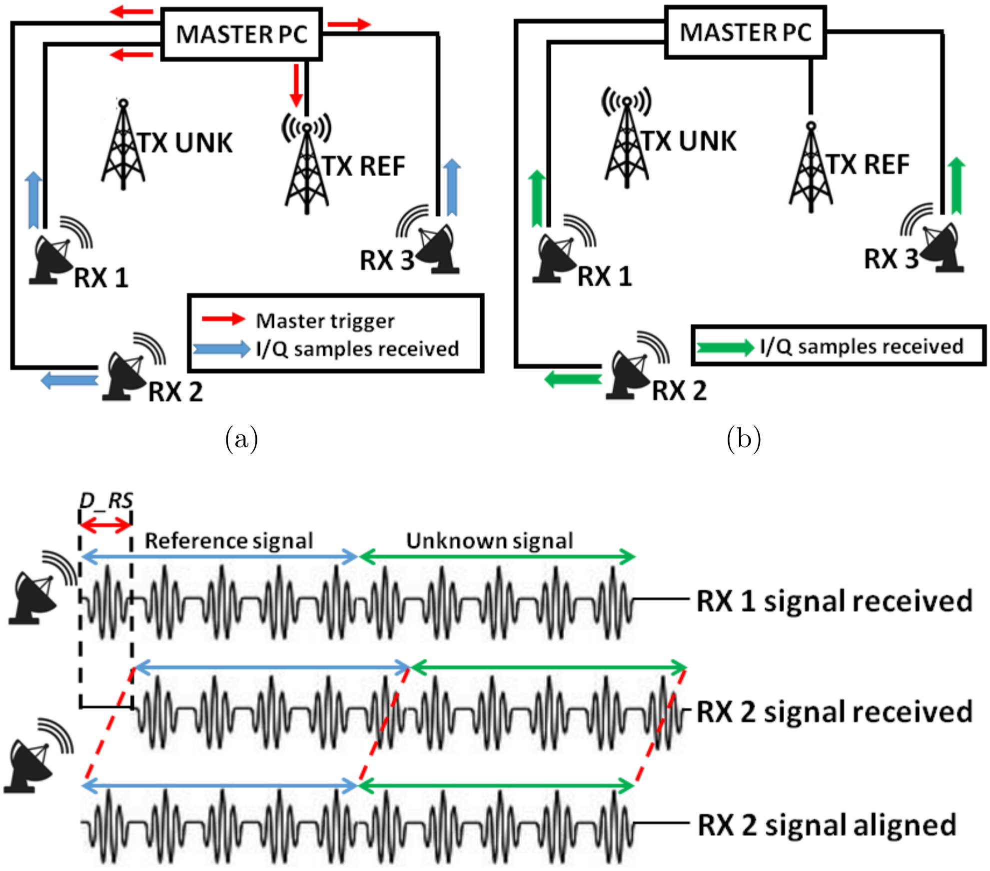
$$s'\left(t'\right) = s'\left(t'\right)|_{t'=t(1+\varphi)} \tag{4.2}$$

As the received signal is not continuous, we decided to perform upsampling operation of a sufficiently high N factor. Then, on the oversampled signal, we perform a downsampling operation by linear interpolation, in order to take the samples at the time instants desired.

4.3.2 Synchronization

A prerequisite for TDOA localization is the synchronization of the receivers. The difference of arrival times could be wrong due to the difference in receivers absolute time.

The idea is to synchronize the receivers roughly in the same time instant (*coarse time synchronization*), then align the received signals (*fine synchronization*). In other words, the idea is not to synchronize the receivers themselves, but rather the raw samples of the received signals. Fig. 4.1 shows an example of samples-based synchronization. The master PC sends a trigger signal to the receivers and reference transmitter (with known position), specifying the reference frequency. The transmitter transmits a Reference Signal (RS) and the receivers send the received I/Q samples to the master PC (Fig. 4.1(a)). Then, without interruptions, the receivers receive the Unknown Signal (US) transmitted by the unknown transmitter to be located,



(a) (b)

(c)

Figure 4.1. Samples-based synchronization. The master PC sends a trigger signal to the devices, the transmitter transmit a RS and the receivers send the received I/Q samples to the master PC (a). Without interruptions, the receivers receive the US (b). The master aligns the signals received from the different receivers on the time axis, using the estimate of the delay (c).

and then send the received I/Q samples to the master PC (Fig. 4.1(b)). Fig. 4.1(c) show the *fine synchronization* between two receivers. The master aligns the signals received from the different receivers on the time axis, using the estimate of the delay. Alignment is done in such a way that the delay of reference signal corresponds to the known distances between transmitter and receivers. The next subsections will show how the calculations are performed to find the delay between the signals, first for the sync and then for the distance of the unknown transmitter.

Let D_RS a delay value between the two received signals by the two RTL-SDR, in the part corresponding to the Reference Signal, it will be the

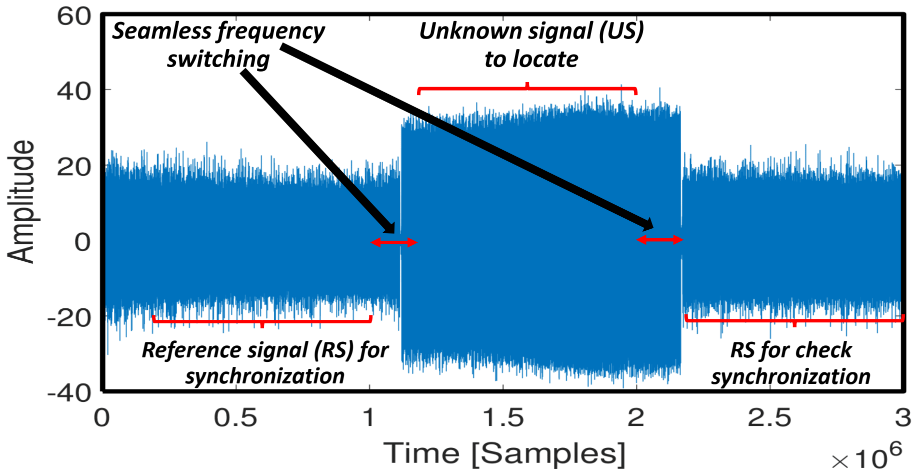


Figure 4.2. Real part of an received signal by RTL-SDR using *librtlsdr-2freq*. Guard interval: $200ms$.

sum of two times:

$$D\_RS = D_{start,rx1,rx2} + T_{v,Ptx\_ref,Prx1,Prx2} \tag{4.3}$$

where D_{start} represents the delay between the two receivers start time, while T_v represents the theoretically predicted TDOA value, calculated based on the speed of light and the relative distance between the reference transmitter and the two receivers ("P" stands for "Position"). Therefore, the rest of the delayed signal, ie s_1 or s_2 based on whether D_{start} is positive or negative, will be shifted back in time of $|D_{start}|$. In Fig. 4.1 $T_v = 0$, because the relative distance between the reference transmitter and the two receivers is the same.

A samples-based synchronization must be repeated for each reception. Accordingly, the transmitter to be located must transmit the signal at the same frequency of the reference transmitter, so that both signals can be received correctly by the RTL-SDRs. In order to allow the localization of signals on a wider spectrum, for example from $500kHz$ to $1.7GHz$ for RTL2832U (declared in the specifications), it is necessary to be able to switch frequencies without interruptions. For this purpose it's possible to use *librtlsdr-2freq*[5], which is a modified version of *librtlsdr* that use the *rtl_sdr* command. *Librtlsdr-2freq* allows the switching between two frequencies without loss of samples, basing on a branch of *librtlsdr* called *async_rearrangement*[6].

---

[5] https://github.com/DC9ST/librtlsdr-2freq

[6] https://github.com/mutability/librtlsdr/tree/async-rearrangements

Figure 4.2 shows an example of a signal received from the RTL-SDR in one experiment discussed in Sec. 4.4. The first part of the signal corresponds to the reception of the RS, used for synchronization. The second part corresponds to the US to be located. The third part will be used as additional verification (Reference Signal Check, RSC). Figure 4.2 also shows that it is necessary to consider a time interval useful for switching to a new frequency that, determined empirically in [74] and reproduced in the laboratory, resulting always less than $200ms$. A guard interval of $200ms$ will then be used, in which the received signal will not be analyzed. Seems to be there a signal within the guard interval time and so the time required for switching the frequency to be much less, but this will be discussed in Sec. 4.5.

4.3.3 Signal processing

The most suitable function for measuring the delay between two inputs precisely is the correlation function. Let s_1 and s_2 the signals received by two receivers and N is the total number of samples of the two signals, the correlation function (for real signals) is:

$$\left(|s_1| * |s_2|\right)[\tau] = \sum_{n=0}^{N-1} s_1\left(nT\right) s_2\left(\left(n+\tau\right)T\right) \tag{4.4}$$

The estimated delay, in samples, is the peak of the correlation function:

$$\hat{T}_v = \underset{\tau}{\operatorname{argmax}} \left(|s_1| * |s_2|\right)[\tau] \tag{4.5}$$

Eq. 4.4 and Eq. 4.5 are an example for real signals, but actually RTL-SDR, such as most radios, deliver I/Q samples at its output that can be interpreted as a complex signal, exhibiting a phase and amplitude information. It exists one correlation function for each on the following type of data: complex values, amplitudes, phases and phase differences. The Performance of the correlation function depends on many factors, such as signal bandwidth, filtering, noise, multipath, signal content, correlation method, signal length, and others. More details on the correlation operation can be found in [75], chapter 7. In [74] it is stated that the method using phase differences is mostly suitable for a wide range of different signals. Let $s_{dp}\left(k\right) = \measuredangle s\left(k\right) - \measuredangle s\left(k-1\right)$, we will use:

$$\left(|s_{1dp}| * |s_{2dp}|\right)[\tau] = \sum_{n=1}^{N-1} s_{1dp}\left(nT\right) s_{2dp}\left(\left(n+\tau\right)T\right) \tag{4.6}$$

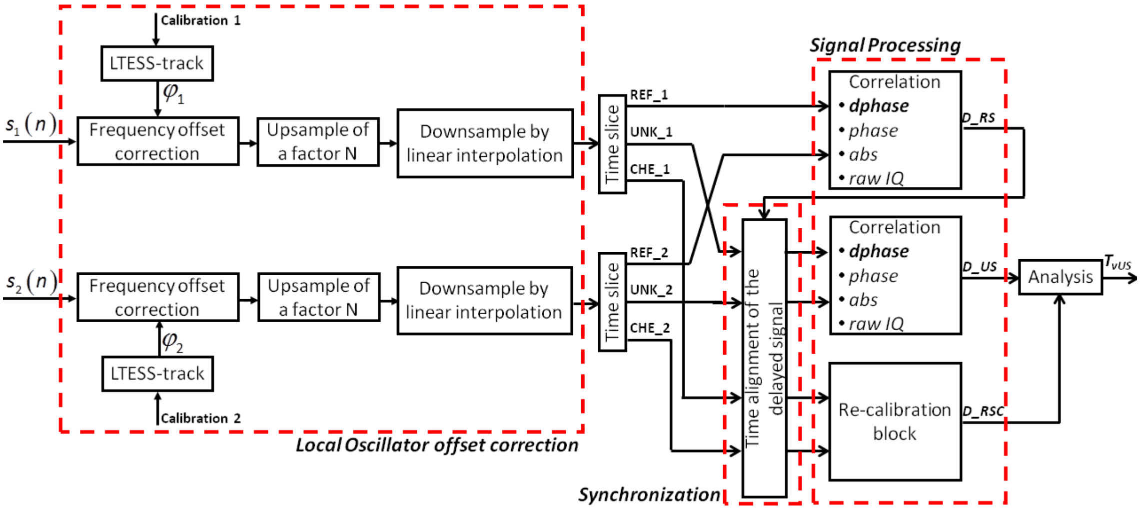


Figure 4.3. Block diagram of the system for the calculation of TDOA value between two receivers.

4.3.4 System overview

Fig. 4.3 shows the block diagram to calculate TDOA value between two receivers, highlighting the steps discussed before.

Local Oscillator offset correction and *Synchronization* are fully explained in Subsec. 4.3.1 and Subsec. 4.3.2.

Signal processing. As explained in Subsec. 4.3.3, it is formed by simple correlation operations. The blocks output is in Samples, D_RS will be useful for synchronization and D_US and D_RSC will be useful for *Analysis* block. They are delays between the two receivers, respectively the delay of reception of US and the theoretically predicted TDOA value relative to the reception of RSC.

Analysis block. We would expect that $D\_RSC = T_{v,Ptx\_ref,Prx1,Prx2}$ of eq. 4.3, because the reference transmitter for the two RS and RSC signals is the same. If it is not so, it indicates an imperfect synchronization operation, so the difference between the expected value and the real value will be used as a corrective value for T_{vUS}, as follows:

$$T_{vUS} = D\_US + \frac{D\_RSC - T_{v,Ptx\_ref,Prx1,Prx2}}{2} \tag{4.7}$$

As mentioned in Sec. 4.1, a TDOA value, T_{vUS}, can be converted to a

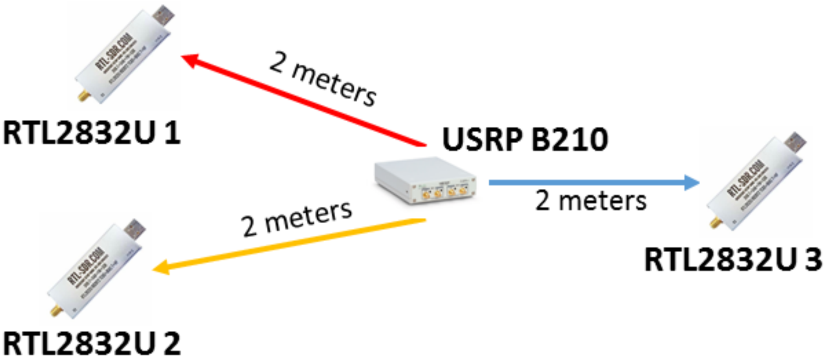


Figure 4.4. Setup.

distance with which is possible to calculate the hyperbola in which the unknown transmitter is located. Subsequently, through intersection with the hyperbolas generated by the pairs of receivers, it will allow the location of the transmitter.

4.4 Experimental validation

In order to test our transmitter localization system, an experiment was carried out using a simple setup shown in Fig. 4.4. The three RTL2832U are connected via USB to a master pc, 2 meters away from a USRP B210 that acts both as reference TX and as unknown TX. The experiment starts on one tests, for each RTL2832U, for the estimate of the LO offset with LTESS-track. Then, done 50 tests of the transmitter localization. For the LO offset (φ_x) estimate an RTL2832U receives 10^6 samples at 1.92 MS/s at the LTE frequency, ie 806 MHz. The 50 tests follow what is explained in Sec. 4.3, with the follow parameters, some of which are shown also in fig. 4.2, frequency RS and RSC 806 MHz, frequency US 244 MHz, sampling rate 2 MS/s, $3 \cdot 10^6$ samples received for each RTL-SDR for each test, guard interval 200 ms. This time was chosen because *librtlsdr-2freq* requires a time interval to switch to a new frequency, which was found to be limited superiorly by 200ms in the study empirically conducted both in [74] and in the laboratory. The all transmitted signals, RS, US and RSC, are DAB+ signals. As mentioned in Subsec. 4.3.3 we chose correlation version based on phase difference (4.6). Tab. 4.1 summarizes the experimental parameters.

The expected value of both D_US and D_RSC is zero, because they

Table 4.1. Parameters used in the experiment.

Sampling rate for φ_x estimate	$1.92 MS/s$
Number of tests for φ_x estimate	1 for $0.521s$
frequency RS and RSC	806 MHz
frequency US	244 MHz
Sampling rate	$2MS/s$
Number of tests	50 for $1.5s$
Guard interval	$200ms$
Correlation version	*Phase difference*

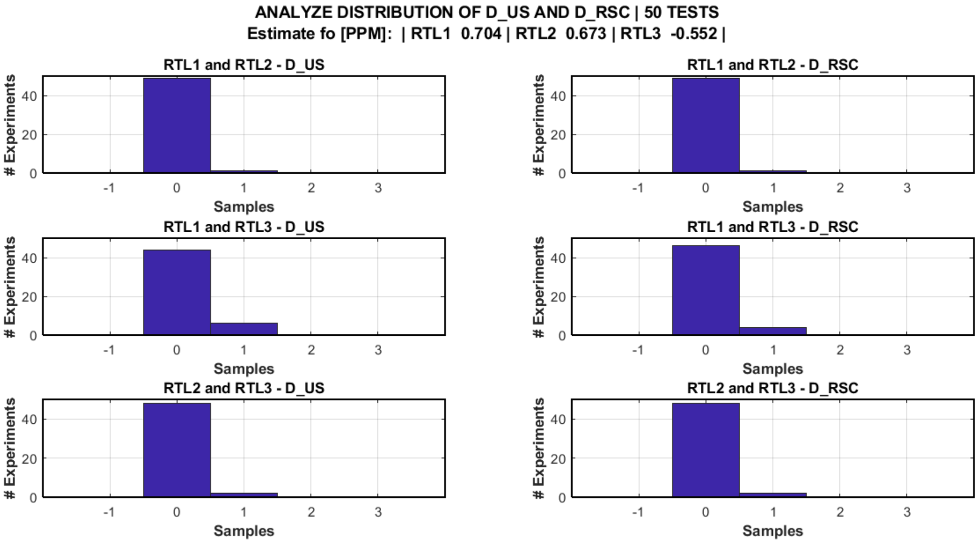


Figure 4.5. Distribution of the values of D_US and D_RSC.

represent the TDOA value of the US and of the RSC between two different receivers, that with this configuration should be zero. Having a TDOA value other than zero means, in this case, it means wrong estimate the position (ie of the hyperbola) of the transmitter of hundreds meters. To quantify this statement, we use the Eq. 4.7, noting that the TDOA value between two receivers at the same distance from the transmitter is zero (ie $T_{v,Ptx\_ref,Prx1,Prx2} = 0$), and convert the result in meters

$$T_{vUS}[m] = \frac{c \cdot T_{vUS}}{fs} = \frac{c \cdot (2 \cdot D\_US + D\_RSC)}{2 \cdot fs} \tag{4.8}$$

where c is the speed of light and fs is the sampling rate. As mentioned, T_{vUS} must be equal to zero. An error of one sample on the estimate of

D_US, with the data of our experiment, will therefore cause an error on the estimate of T_{vUS} of 150 meters (75 meters for an error of one sample on the estimate of D_RSC).

For each test and for each pair of receivers we will have a value of D_US and a value of D_RSC. The results are shown in fig. 4.5, which also shows the estimate of LO offset for each RTL-SDR.

The results show a Residual Error of 1 sample for less than 6% of the total values, show a high robustness of the system.

4.5 Processing time for synchronization

The experiment described in Sec. 4.4 use, as mentioned, $3 \cdot 10^6$ samples for each localization. Therefore, without considering the other times, such as processing times, data transmission time to the PC master and so on, at the sampling rate used are required 1.5 seconds of reception for each RTL-SDR. In a future integration of the localization system in the ElectroSense network, which has the goal to sense the entire spectrum and to make the data available in real-time, it is clear that this time is inefficient, because it would make the receiver unable to analyze the spectrum for a long time. Some solution, which can also be implemented together, to reduce the inefficiency could be the following.

Less frequent synchronization. In order to reduce the inefficiency time, it is possible to decrease the frequency of reception of the RS and the RSC, so as not to do a synchronization for each localization. In other words, increase the number of US signals between the RS and RSC signals, fixed to one for the experiment shown in Sec.4.4;

Reduce number of samples. It is the most intuitive solution. We have verified the minimum number of samples that can be used in a study that will be shown shortly;

Variable correlation window. The idea is to analyze a RS with a strong length the first few times and subsequently reduce it. In order to reduce the number of samples of RS, without getting worse the estimate of D_RS, it is possible to try to predict the trend of the D_RS by using linear regression methods (as a Kalman filter model, used in [76]). Then, with a priori knowledge of D_RS for the block ”Time alignment of the delayed signal”, adjusting the estimate using a RS transmission shorter.

The proposed solutions need a modification of the *librtlsdr-2freq* library. It currently only provides 3 reception parts, all with the same number of

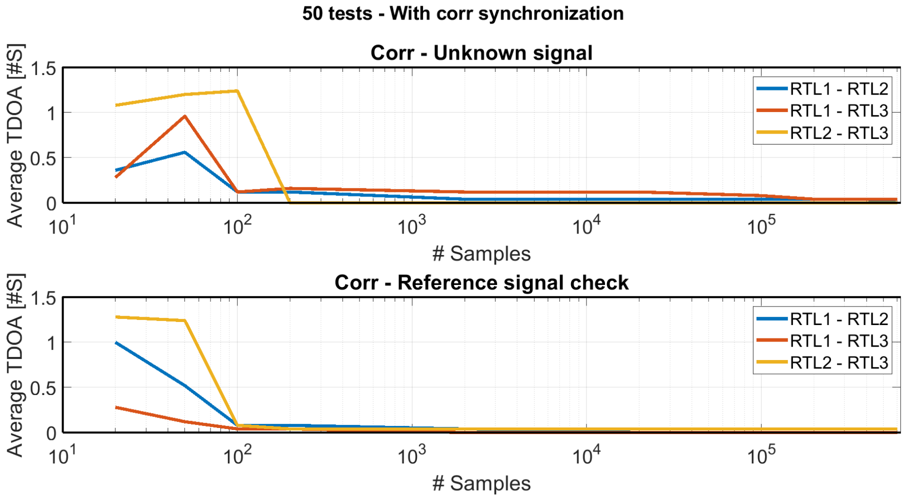


Figure 4.6. Average error (ideally $TDOA = 0$) to vary the length of the signal used for the correlation operation. Top on the figure shown the average, on 50 tests, of the value of D_US, bot instead shows D_RSC.

samples and with the first and the last part of signal, that in the original system correspond to receive the RS and RSC, on the same central frequency.

In order to understand how far the number of samples can be reduced, it is necessary to be able to analyze very short signals. If the starting time delay between two receivers (D_{start}) is greater than the time chosen for data analysis, the two received signals will have no common parts, making the correlation operation useless. So, very fine synchronization between the receivers is necessary, for example using protocols such as the Network Time Protocol (NTP) through internet, which gives a time synchronization of tens of milliseconds.

Using the same data as the experiment described in Sec. 4.4, we analyze the Residual Error as the signal length varies. We simulated the perfect time synchronization by taking the signal already aligned with the results of the first correlation (D_RS). The values of D_US and D_RSC can be calculate analyzing a portion of the raw samples I/Q received, truncating the signal at the number of samples desidered. The simulation results are shown in fig. 4.6, from which it can be deduced that, if we have a perfect time synchronization, we obtain an acceptable error also using a very low number of samples.

Although the study would indicate that only 200 samples could be used, the limits of *librtlsdr-2freq* must be considered. As mentioned in Sec. 4.3.2, it requires a time interval to switch to a new frequency. This means that even

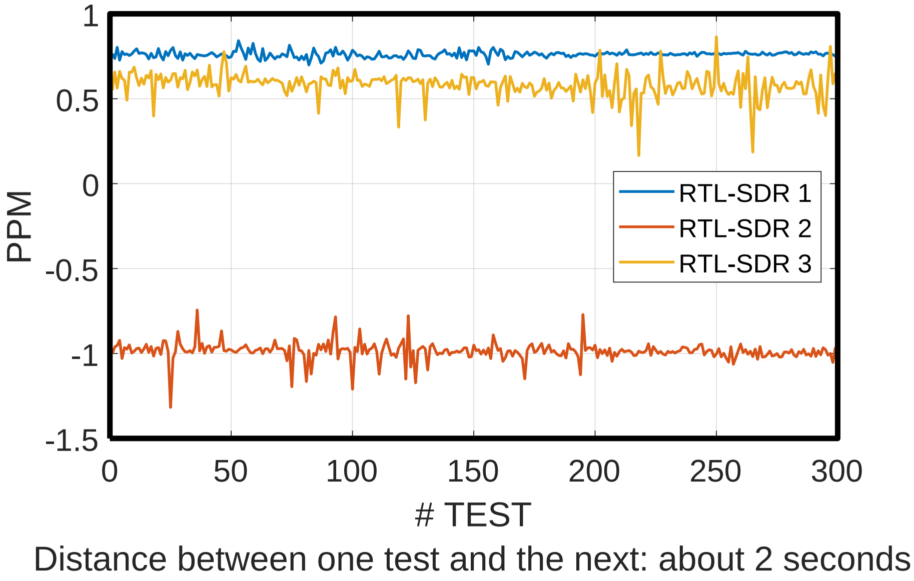


Figure 4.7. Stability of the Local Oscillator offset of 3 RTL-SDR located in different rooms.

if set an experiment with 200 samples, $0.1ms$ at $2MS/s$, for each frequency, would still need $200ms$ for switching frequency, unless there are future improvements to *librtlsdr-2freq*. This makes the switch operation slower than the signal to receive and causes the crashes of tool.

4.6 Rate and time for estimate of LO offset

The experiment described in the Sec. 4.4 provides, as mentioned, a preliminary tests, for each RTL-SDR, for the estimate of the LO offset with LTESS-track. Only one estimate is sufficient because, as we will see shortly, the stability of the LOs of the RTLs RTL2832U is quite high. In a long-term experiment it will be necessary to analyze the ratey of LO offset estimate. As in the previous subsection for D_US, we can try to predict the trend of the LO offset φ_x by using linear regression methods, as a Kalman filter or other. It could be useful to drastically reduce the rate of use of the LTESS-track, then to reduce the stops of localization experiments for the calibration of the LO.

We conducted a preliminary study. The experiment consists of 300 tests, for each RTL2832U, for estimate the LO offset with LTESS-track. The devices are connected to the master PC via ssh connection on a LAN. Each RTL2832U receives 10^6 Samples at $1.92MS/s$ for 300 times at the LTE frequency, ie $806MHz$. This samples will then be the LTESS-track input, which will then provide the φ_x value. The results are shown in Fig. 4.7.

The figure shown the results and prove a hight stability of the LO, with practically the same values over 300 tests, which correspond to about 10 minutes of tests (about $2s$ is the time to receive 10^6 samples at $1.92MS/s$ plus the time to start the next reception).

In order to remove the stops of localization experiments, we can use the signal transmitted by the reference transmitter (RS and RSC) as input for LTESS-track. We conducted an experiment by integrating the the LTESS-track code in the localization code. In order to use RS and RSC for LTESS-track input is necessary the two signals must be LTE signals. Then, the setup is the same of the experiment shown in Sec. 4.4, except that the reference transmitter is not a USRP B210, but the the closest LTE station. We have achieved similar results.

Chapter 5

Conclusions

In this thesis we exploited physical layer flexibility enabled by emerging wireless standards for coping with different application scenarios. For physical layer flexibility we do not mean necessarily using wireless nodes based on Software-Defined-Radio, but rather exposing advanced primitives and PHY-layer information for building flexible modulations, sending signalling mechanisms for controlling the access to the wireless channel and optimizing the network performance. Examples of these primitives are the transmission of signalling tones for physical layers based on OFDM, in order to implement innovative contention mechanisms or other signalling messages for mitigating hidden nodes. An example of PHY-layer information useful for optimizing the capacity of wireless systems is the airtime consumed by devices working at different SFs (even during collisions) in LoRa networks, while the difference in arrival times is used for localization. For each main contribution of the thesis, we can draw the following main conclusions.

First, in Chapter 2, we have explored the possibility of running multiple contention rounds in the frequency domain in random access networks. Performance enhancement is due mainly to the small contention time and the very low level of collision probability, reducing the two major sources of inefficiency of current WiFi MAC protocol versions. As for the practical feasibility of the frequency domain repeated contention rounds, we have demonstrated that repeated contentions can be implemented on current wireless technologies, in OFDM-based PHY layers. We give a full analytical evaluation of the collision probability and throughput of ReCo, including the case of non-homogeneous probability distributions for the selection of tones in each round. This evaluation provided some guidelines on the dimensioned of ReCo, that can be dimensioned with reliable and simple formulas and it does not require critical tuning or complex adaptive algorithms, e.g., as the number of stations varies. Eq. 2.3 provides a bound for collision probability in the case of homogeneous probability distributions, indicating how it is better to increase the number of rounds of ReCo to reduce the collision probability, rather than increase the number of tones. The study on non-homogeneous probability distributions on the choice of tones, provides an optimal probability distribution (eq. 2.11) that can be used in the first round to drastically reduce the number of collisions. Despite its simplicity, ReCo offers robust and close-to-ideal throughput performance. The proposed mechanism is very suitable for the emerging high-rate PHY of recent 802.11 extensions, such as 802.11ax. So, we extend the basic ReCo to multi-channel contention and apply it to the IEEE 802.11ax standard. We show that ReCo can significantly improve performance and, in order to increase robustness in presence of imperfect carrier sensing, we extended the idea of frequency-domain contention in a scheme called ReCHo, exploiting special echo signals sent by the APs. We prove the effectiveness of ReCHo to restore the ideal performance level promised by repeated contention, overcoming the effect of imperfect channel sensing completely. We quantified the performance benefits that can be achieved by exploiting tone-based contention with echo for accessing multiple RUs in 802.11ax networks, by comparing the throughput results under varying load conditions with the standard OBO scheme. Our simulations prove that tone-based contention significantly improves the OBO efficiency even in presence of hidden nodes.

Second, in Chapter 3, we have studied the impact of two peculiar characteristics of LoRa modulations, i.e. the high capture probability and the imperfect orthogonality between different SFs, on the overall cell capacity. We showed that the link-level performance of LoRa is deeply influenced by capture effects and by inter-SF collisions which can indeed cause loss if the interference power is strong enough. We then exploited this link-level analysis to model analytically the achievable network capacity in a typical LoRa cell. We showed that high SFs are severely affected by inter-SF interference and that the use of power control and packet fragmentation to compensate such problem may be counterproductive. Although deploying multiple gateways can mitigate the capacity loss and boost the occurrence of channel captures, the overall capacity increase becomes negligible after 16-24 gateways. Finally, when only a handful of gateways are present, the deployment should be as distant from the center as possible and not on a regular grid. We give important insights on LoRa technology with which is possible to provide new accurate guidelines for the correct design of future LoRa networks.

Finally, in Chapter 4, we defines a method for indoor localization based on TDOA and that using low cost SDR receivers. Using appropriate signal processing techniques for each step of TDOA, we have shown that our system achieves optimal results in a simple scenario. We have also performed a feasibility study on a way to reduce the time inefficiency of two steps of TDOA (*synchronization* and *LO offset correction*). Future works could be improving the *LO offset correction* step. For example by replacing equations 4.1 and 4.2 with other more efficient, such as equation (6) in [73]. Another example could be replace the operations of upsample and downsample with a faster method, which uses the FFT of the received samples in time, divides them into several parts, adds zeros between each sample and the other and finally the IFFT next for return in the time domain. Others future works could be improving the correlation operation, for example choose the correlation version based on the RS transmitted. Finally, the localization system will be tested in a realistic scenario and then integrated into the ElectroSense network, in order to enable technologies for various applications, such as to support do collaborative decoding of a wideband signal [76], or even to ease the beamforming technology in the IEEE 802.11ac and later standards networks.

Bibliography

[1] “Bglobal mobile data traffic forecast update, 2016–2021 white paper.”

[2] V. Jones and H. Sampath, “Emerging technologies for WLAN,” *IEEE Communications Magazine*, vol. 53, pp. 141–149, March 2015.

[3] B. Bellalta, L. Bononi, R. Bruno, and A. Kassler, “Next generation IEEE 802.11 Wireless Local Area Networks: Current status, future directions and open challenges ,” *Comp. Comm.*, vol. 75, pp. 1 – 25, 2016.

[4] M. S. Afaqui, E. Garcia-Villegas, and E. Lopez-Aguilera, “IEEE 802.11ax: Challenges and Requirements for Future High Efficiency WiFi,” *IEEE Wireless Communications*, vol. 24, pp. 130–137, June 2017.

[5] D. Deng, Y. Lin, X. Yang, J. Zhu, Y. Li, J. Luo, and K. Chen, “IEEE 802.11ax: Highly Efficient WLANs for Intelligent Information Infrastructure,” *IEEE Communications Magazine*, vol. 55, pp. 52–59, Dec 2017.

[6] E. Khorov, A. Kiryanov, A. Lyakhov, and G. Bianchi, “A tutorial on ieee 802.11ax high efficiency wlans,” *IEEE Communications Surveys Tutorials*, pp. 1–1, 2018.

[7] C. Knapp and G. Carter, “The generalized correlation method for estimation of time delay,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, pp. 320–327, August 1976.

[8] Y. Chan, R. Hattin, and J. Plant, “The least squares estimation of time delay and its use in signal detection,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 26, pp. 217–222, June 1978.

[9] S. Rajendran, R. Calvo-Palomino, M. Fuchs, B. Van den Bergh, H. Cordobes, D. Giustiniano, S. Pollin, and V. Lenders, “Electrosense: Open and big spectrum data,” *IEEE Communications Magazine*, vol. 56, pp. 210–217, Jan 2018.

[10] R. Calvo-Palomino, H. Cordobés, M. Engel, M. Fuchs, P. Jain, M. Liechti, S. Rajendran, M. Schäfer, B. V. den Bergh, S. Pollin, D. Giustiniano, and V. Lenders, “Electrosense+: Empowering people to decode

the radio spectrum," 2018.

[11] E. Fitzgerald and B. Landfeldt, "The failure of CSMA in emerging wireless network scenarios," in *2014 IFIP Wireless Days (WD)*, pp. 1–4, Nov 2014.

[12] A. Baiocchi, "Variability of service times and throughput efficiency trade-off in IEEE 802.11 DCF," in *21st International Teletraffic Congress, 2009 (ITC 21)*, pp. 1–8, Sept 2009.

[13] Z. Abichar and J. Chang, "A Medium Access Control Scheme for Wireless LANs with Constant-Time Contention," *IEEE Transactions on Mobile Computing*, vol. 10, pp. 191–204, Feb 2011.

[14] M. Gowda, N. Roy, R. Roy Choudhury, and S. Nelakuditi, "Backing out of Linear Backoff in Wireless Networks," in *Proceedings of the 1st ACM Workshop on Hot Topics in Wireless*, HotWireless '14, (New York, NY, USA), pp. 7–12, ACM, 2014.

[15] Y. Mao and L. Shen, "A first-round-bye based priority scheme for WLANs with two access categories," in *2015 Int. Conf. on Wireless Communications Signal Processing (WCSP)*, pp. 1–5, Oct 2015.

[16] X. Feng, J. Zhang, Q. Zhang, and B. Li, "Use your frequency wisely: Explore frequency domain for channel contention and ACK," in *INFOCOM, 2012 Proceedings IEEE*, pp. 549–557, March 2012.

[17] S. Sen, R. Roy Choudhury, and S. Nelakuditi, "No time to countdown: Migrating backoff to the frequency domain," in *Proc. of ACM MOBICOM'11*, pp. 241–252, 2011.

[18] D. Betsekas and R. Gallager, *Data networks.* 2nd Ed., Prentice Hall, Englewood Cliffs, NJ, 1992.

[19] W. Zame, J. Xu, and M. van der Schaar, "Winning the Lottery: Learning Perfect Coordination With Minimal Feedback," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, pp. 846–857, Oct 2013.

[20] A. Baiocchi, I. Tinnirello, D. Garlisi, and A. L. Valvo, "Random Access with Repeated Contentions for Emerging Wireless Technologies," in *IEEE INFOCOM 2017*, pp. 1–9, May 2017.

[21] L. Song, Y. Liao, K. Bian, L. Song, and Z. Han, "Cross-Layer Protocol Design for CSMA/CD in Full-Duplex WiFi Networks," *IEEE Communications Letters*, vol. 20, pp. 792–795, April 2016.

[22] Y. Zhang, L. Lazos, K. Chen, B. Hu, and S. Shivaramaiah, "Multi-Channel Medium Access without Control Channels: A Full Duplex MAC Design," *IEEE Transactions on Mobile Computing*, vol. 16, pp. 1032–1046, April 2017.

[23] A. Mutairi and S. Roy, "An OFDM-Aware Reservation Random Access Protocol for Interference Mitigation in OFDMA Femtocells," *IEEE Transactions on Communications*, vol. 63, pp. 301–310, Jan 2015.

[24] S. K. Fayaz, F. Zarinni, and S. Das, "Ez-Channel: A distributed {MAC} protocol for efficient channelization in wireless networks ," *Ad Hoc Networks*, vol. 31, pp. 34 – 44, Aug 2015.

[25] E. Magistretti, O. Gurewitz, and E. Knightly, "802.11ec: Collision Avoidance Without Control Messages," *IEEE/ACM Transactions on Networking*, vol. 22, pp. 1845–1858, Dec 2014.

[26] K. Tan, J. Fang, Y. Zhang, S. Chen, L. Shi, J. Zhang, and Y. Zhang, "Fine-grained Channel Access in Wireless LAN," *SIGCOMM Comput. Commun. Rev.*, vol. 40, pp. 147–158, Aug. 2010.

[27] M. Heusse, F. Rousseau, R. Guillier, and A. Duda, "Idle Sense: An Optimal Access Method for High Throughput and Fairness in Rate Diverse Wireless LANs," *SIGCOMM Comput. Commun. Rev.*, vol. 35, pp. 121–132, Aug 2005.

[28] Y. Grunenberger, M. Heusse, F. Rousseau, and A. Duda, "Experience with an implementation of the idle sense wireless access method," in *Proceedings of the 2007 ACM CoNEXT Conference*, CoNEXT '07, (New York, NY, USA), pp. 24:1–24:12, ACM, 2007.

[29] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 535–547, March 2000.

[30] A. Kumar, E. Altman, D. Miorandi, and M. Goyal, "New Insights From a Fixed-Point Analysis of Single Cell IEEE 802.11 WLANs," *IEEE/ACM Trans. on Networking*, vol. 15, pp. 588–601, June 2007.

[31] A. Baiocchi, "Variability of service times and throughput efficiency trade-off in IEEE 802.11 DCF," in *2009 21st International Teletraffic Congress*, pp. 1–8, Sep. 2009.

[32] S. Sen, R. R. Choudhury, and S. Nelakuditi, "Listen (on the frequency domain) before you talk," in *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks*, Hotnets-IX, (New York, NY, USA), pp. 16:1–16:6, ACM, 2010.

[33] D. Bharadia, E. McMilin, and S. Katti, "Full Duplex Radios," *SIGCOMM Comput. Commun. Rev.*, vol. 43, pp. 375–386, Aug 2013.

[34] M. S. Afaqui, E. Garcia-Villegas, and E. Lopez-Aguilera, "Ieee 802.11ax: Challenges and requirements for future high efficiency wifi," *IEEE Wireless Communications*, vol. 24, pp. 130–137, June 2017.

[35] D. J. Deng, K. C. Chen, and R. C. Cheng, "IEEE 802.11ax: Next generation wireless local area networks," *10th International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (QShine)*, 2014.

[36] L. Lanante, H. O. T. Uwai, Y. Nagao, M. Kurosaki, and C. Ghosh, "Performance analysis of the 802.11ax ul ofdma random access protocol in dense networks," in *2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, May 2017.

[37] "IEEE Proposed TGax draft specification. doc.: IEEE P802.11ax /D2.0, October 2017. Technical report," IEEE, 2017.

[38] M. Bor, U. Roedig, T. Voigt, and J. Alonso, "Do lora low-power wide-area networks scale?," 11 2016.

[39] B. Reynders and S. Pollin, "Chirp spread spectrum as a modulation technique for long range communication," in *2016 Symposium on Communications and Vehicular Technologies (SCVT)*, pp. 1–5, Nov 2016.

[40] B. Reynders, W. Meert, and S. Pollin, "Range and coexistence analysis of long range unlicensed communication," in *2016 23rd International Conference on Telecommunications (ICT)*, pp. 1–6, May 2016.

[41] M. Gucciardo, I. Tinnirello, and D. Garlisi, "Demo: A cell-level traffic generator for lora networks," pp. 474–476, 10 2017.

[42] C. Goursaud and J.-M. Gorce, "Dedicated networks for iot: Phy / mac state of the art and challenges," *EAI Endorsed Transactions on Internet of Things*, vol. 1, p. 150597, Oct 2015.

[43] A. Augustin, J. Yi, T. H. Clausen, and W. Townsley, "A study of lora: Long range low power networks for the internet of things," *Sensors*, vol. 16, p. 1466, 10 2016.

[44] D. Bankov, E. Khorov, and A. Lyakhov, "On the limits of lorawan channel access," pp. 10–14, 11 2016.

[45] M. Zorzi and R. R. Rao, "Capture and retransmission control in mobile radio," *IEEE J.Sel. A. Commun.*, vol. 12, pp. 1289–1298, Sept. 2006.

[46] D. J. Goodman and A. A. M. Saleh, "The near/far effect in local aloha radio communications," *IEEE Transactions on Vehicular Technology*, vol. 36, pp. 19–27, Feb 1987.

[47] O. Bernard, A. Seller, and N. Sornin, "Low power long range transmitter," *European Patent Application EP 2763321 A1*, Aug 2014.

[48] K. Pahlavan, Xinrong Li, and J. P. Makela, "Indoor geolocation science and technology," *IEEE Communications Magazine*, vol. 40, pp. 112–118, Feb 2002.

[49] C. Xu, J. He, X. Zhang, P. Tseng, and S. Duan, "Toward near-ground

localization: Modeling and applications for toa ranging error," *IEEE Transactions on Antennas and Propagation*, vol. 65, pp. 5658–5662, Oct 2017.

[50] V. Djaja-Josko and J. Kolakowski, "A new method for wireless synchronization and tdoa error reduction in uwb positioning system," in *2016 21st International Conference on Microwave, Radar and Wireless Communications (MIKON)*, pp. 1–4, May 2016.

[51] X. Li, S. Li, D. Zhang, J. Xiong, Y. Wang, and H. Mei, "Dynamic-music: accurate device-free indoor localization," pp. 196–207, 09 2016.

[52] P. Pivato, L. Palopoli, and D. Petri, "Accuracy of rss-based centroid localization algorithms in an indoor environment," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, pp. 3451–3460, Oct 2011.

[53] C. Xu, B. Firner, Y. Zhang, R. Howard, J. Li, and X. Lin, "Improving rf-based device-free passive localization in cluttered indoor environments through probabilistic classification methods," in *Proceedings of the 11th International Conference on Information Processing in Sensor Networks*, IPSN '12, (New York, NY, USA), pp. 209–220, ACM, 2012.

[54] E. Gülch, "Investigations on google tango development kit for personal indoor mapping," *studies*, vol. 1, p. 3.

[55] L. Li, P. Hu, C. Peng, G. Shen, and F. Zhao, "Epsilon: A visible light based positioning system," in *11th USENIX Symposium on Networked Systems Design and Implementation (NSDI 14)*, (Seattle, WA), pp. 331–343, USENIX Association, 2014.

[56] K. Chintalapudi, A. Padmanabha Iyer, and V. N. Padmanabhan, "Indoor localization without the pain," in *Proceedings of the Sixteenth Annual International Conference on Mobile Computing and Networking*, MobiCom '10, (New York, NY, USA), pp. 173–184, ACM, 2010.

[57] Y.-S. Kuo, P. Pannuto, K.-J. Hsiao, and P. Dutta, "Luxapose: Indoor positioning with mobile phones and visible light," in *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*, MobiCom '14, (New York, NY, USA), pp. 447–458, ACM, 2014.

[58] G. Oberholzer, P. Sommer, and R. Wattenhofer, "Spiderbat: Augmenting wireless sensor networks with distance and angle information," *Proceedings of the 10th ACM/IEEE International Conference on Information Processing in Sensor Networks*, pp. 211–222, 2011.

[59] M. Bouet and A. Santos, "Rfid tags: Positioning principles and localization techniques," pp. 1 – 5, 12 2008.

[60] M. Eichelberger, K. Luchsinger, S. Tanner, and R. Wattenhofer, "Indoor

localization with aircraft signals," in *Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems*, SenSys '17, (New York, NY, USA), pp. 12:1–12:14, ACM, 2017.

[61] M. Bargh and R. de Groote, "Indoor localization based on response rate of bluetooth inquiries," pp. 49–54, 01 2008.

[62] N. Mair and Q. H. Mahmoud, "A collaborative bluetooth-based approach to localization of mobile devices," *8th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*, pp. 363–371, 2012.

[63] A. N. Raghavan, H. Ananthapadmanaban, M. S. Sivamurugan, and B. Ravindran, "Accurate mobile robot localization in indoor environments using bluetooth," in *2010 IEEE International Conference on Robotics and Automation*, pp. 4391–4396, May 2010.

[64] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, pp. 1067–1080, Nov 2007.

[65] K. Kaemarungsi, R. Ranron, and P. Pongsoon, "Study of received signal strength indication in zigbee location cluster for indoor localization," *2013 10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, pp. 1–6, 2013.

[66] A. Varshavsky, E. de Lara, J. Hightower, A. LaMarca, and V. Otsason, "Gsm indoor localization," *Pervasive Mob. Comput.*, vol. 3, pp. 698–720, Dec. 2007.

[67] T. Kos, M. Grgic, and G. Sisul, "Mobile user positioning in gsm/umts cellular networks," in *Proceedings ELMAR 2006*, pp. 185–188, June 2006.

[68] J. Peral-Rosado, J. A. López-Salcedo, G. Seco-Granados, F. Zanier, and C. Massimo, "Achievable localization accuracy of the positioning reference signal of 3gpp lte," pp. 1–6, 06 2012.

[69] Y. Chen, D. Lymberopoulos, J. Liu, and B. Priyantha, "Fm-based indoor localization," in *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services*, MobiSys '12, (New York, NY, USA), pp. 169–182, ACM, 2012.

[70] S.-H. Fang and T.-N. Lin, "Cooperative multi-radio localization in heterogeneous wireless networks," *Trans. Wireless. Comm.*, vol. 9, pp. 1547–1551, May 2010.

[71] R. Carvalho, Shan-Ho Yang, Yao-Hua Ho, and Ling-Jyh Chen, "Indoor localization using fm and dvb-t signals," in *2016 13th IEEE Annual*

Consumer Communications Networking Conference (CCNC), pp. 862–867, Jan 2016.

[72] S. A. Shaukat, K. Munawar, M. Arif, A. I. Bhatti, U. I. Bhatti, and U. M. Al-Saggaf, "Robust vehicle localization with gps dropouts," in *2016 6th International Conference on Intelligent and Advanced Systems (ICIAS)*, pp. 1–6, Aug 2016.

[73] R. Calvo-Palomino, F. Ricciato, D. Giustiniano, and V. Lenders, "Ltess-track: A precise and fast frequency offset estimation for low-cost sdr platforms," pp. 51–58, 10 2017.

[74] S. Scholl, "Introduction and experiments on transmitter localization with tdoa," *Software Defined Radio Academy*, July 2017.

[75] S. W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing*. San Diego, CA, USA: California Technical Publishing, 1997.

[76] R. Calvo-Palomino, D. Giustiniano, V. Lenders, and A. Fakhreddine, "Crowdsourcing spectrum data decoding," in *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, pp. 1–9, May 2017.