Cooperative WLAN protocols for multimedia communication

Author(s):
Ettefagh, Azadeh

Publication Date:
2011

Permanent Link:
https://doi.org/10.3929/ethz-a-006363155

Rights / License:
In Copyright - Non-Commercial Use Permitted
Cooperative WLAN Protocols for Multimedia Communication

A dissertation submitted to
ETH ZURICH

for the degree of
Doctor of Sciences

presented by
AZADEH ETTEFAGH
M.Sc., University of Ulm
born July 20, 1975
citizen of Iran

accepted on the recommendation of
Prof. Dr. A. Wittneben, examiner
Prof. Dr. A. Czywicki, co-examiner

2010
To my dear family
Abstract

In the recent years, wireless communications and the WLAN technology, as one of the most prevalent wireless indoor technologies, have received a lot of attention. However, while the demand for multimedia traffic over WLANs increases rapidly, the existing WLANs approach their limits. It is known that multiuser cooperative communication can enhance performance of wireless networks. It can substantially increase the spectral efficiency of wireless networks by utilising interference rather than avoiding it. This paradigm shift has a great impact on the medium access control (MAC) protocol. Most of the existing MAC protocols are designed to reduce the interference. In this work, the focus is on the WLAN and its limitations on utilising cooperative communication. WLANs were originally designed for point-to-point links with data traffic. However, in order to be able to support different applications, each with its own special requirements, further enhancements are needed.

This thesis is divided into two main parts. In the first part, the fundamental bounds of the existing WLANs with respect to physical (PHY) and MAC layers are studied. Outage probability, throughput and delay have been chosen as the main performance criteria. The results in this part present some of the existing problems in WLANs including relatively small coverage range, low efficiency, and some fairness issues. Possible relaying methods as a way to improve the coverage range or the data rate of a WLAN are also discussed. In the second part of the thesis, the focus is on the possible enhancements to improve the WLAN performance. We start this part by looking into a multiuser single-hop scenario and then study three advanced PHY relaying methods namely: multiuser decode-and-forward, two-way, and multiuser zero-forcing relaying scenarios. It has been shown that these methods can improve the spectral efficiency. However, these methods as well as many other cooperative techniques require simultaneous multiuser transmissions and hence cannot be supported by the current IEEE 802.11 systems using carrier sense multiple access with collision avoidance (CSMA/CA). In this part, a novel cluster-based MAC as extension of CSMA/CA is proposed to facilitate these cooperative techniques by utilising the multiuser interference mitigation capability of PHY layer.

According to the proposed cluster-based CSMA/CA (CB-CSMA/CA) scheme, nodes in a
network, including stations and relays, are allocated to different clusters. The nodes, which belong to the same cluster, are allowed to transmit concurrently, provided that there are enough degrees of freedom available to efficiently decode the desired stream at destinations. We study the impact of signal processing on the MAC protocol design, and explain the basics of CB-CSMA/CA as well as the required modifications of the existing IEEE 802.11 MAC.

We have applied the proposed CB-CSMA/CA to the above-mentioned relaying scenarios. The typical uplink/downlink transmissions in a WLAN with a multiple-antenna access point can also be considered as a subset of the first scenario, i.e. multiuser decode-and-forward relaying scenario. The analysis of the network performance with respect to both MAC and PHY layer parameters leads to more comprehensive results compared to the cases where only PHY or MAC layer parameters are studied. The results show that the CB-CSMA/CA improves throughput and reduces delay significantly as compared to the current WLAN systems.

After studying the performance of CB-CSMA/CA with some ideal assumptions, we extend our study to more realistic scenarios and take into account some cross-layer enhancements of the CB-CSMA/CA. In this study we consider three possible backoff procedure models and compare their performances in the presence of PHY channel errors. We have also considered different clustering methods, where stations are grouped into different clusters based on their channel conditions or higher layer requirements. Up to this point, we have only considered scenarios where all cluster members set their contention window to the same value, i.e. they are all perfectly event-synchronised, and therefore transmit simultaneously whenever they access the channel. We have extended then our study by looking into scenarios where stations in a cluster are not anymore perfectly event-synchronised and hence they may transmit in different time slots. We have shown analytically that in the worst case scenario, i.e., when all stations in all clusters are asynchronous, then the CB-CSMA/CA throughput approaches that of the IEEE 802.11 system. In fact, with the same data rate, header durations and backoff model, throughput of the CB-CSMA/CA is always better than that of the IEEE 802.11.

So far, only one- or two-hop communication links have been taken into consideration. In the last chapter, an outlook for multihop communications and possible extensions of the CB-CSMA/CA for the multihop links are presented.

In this thesis we show that the proposed CB-CSMA/CA is a promising approach for a variety of network configurations including typical infrastructure WLANs as well as ad hoc networks with or without relays. The proposed scheme is simple but powerful and flexible. The performance results in Chapter 6 show that throughput and delay of an IEEE 802.11
system can be significantly improved when the CB-CSMA/CA is applied. The analysis and
discussion in Chapter 7 confirm that the performance can be further improved if we can take
some cross-layer parameters into account while choosing the backoff model and forming the
clusters.
Abstract
Kurzfassung


von Mehrbenutzer-Interferenzen auszunutzen, wird ein neuartiges, gruppiertes (cluster-based) MAC Protokoll als Erweiterung von CSMA/CA vorgeschlagen.

Knoten in einem drahtlosen Netzwerk, sowohl Endknoten als auch Relays, werden entsprechend dieser cluster-based CSMA/CA (CB-CSMA/CA) Methode verschiedenen Gruppen (Cluster) zugeordnet. Knoten in der gleichen Gruppe dürfen gleichzeitig übertragen, falls genügend Freiheitsgrade zur Verfügung stehen, um die gewünschten Daten am Zielknoten effizient zu decodieren. Der Einfluss der Signalverarbeitung auf das Design des MAC Protokolls wird untersucht, und sowohl die Grundlagen von CB-CSMA/CA als auch die notwendigen Modifikationen des existierenden IEEE 802.11 MAC werden besprochen.


Bis jetzt wurde einzig one- oder two-hop Kommunikation untersucht. Im letzten Kapitel wird ein Ausblick auf multihop Kommunikation gegeben und mögliche Erweiterungen von
CB-CSMA/CA vorgeschlagen.

Kurzfassung
## Contents

3.1.4.2 Throughput ................................................. 56  
3.1.5 Throughput Results ........................................ 57  
3.1.6 Delay ......................................................... 68  
3.2 Relaying Scenarios .............................................. 70  
3.2.1 Outage Probability ........................................... 71  
3.2.2 Coverage Angle ............................................... 74  
3.2.3 PHY Layer Data Rate ....................................... 78  
3.2.4 MAC Layer Throughput .................................... 79  
3.3 Summary ......................................................... 80

### II Cooperative WLAN Protocols

4 Representative PHY Layer Cooperative Protocols  
4.1 Multiuser Single-Hop Transmissions ............................ 87  
4.2 Multiuser DF Relaying ......................................... 88  
4.3 Two-Way Relaying ............................................... 92  
  4.3.1 Optimising Power Allocation at the Relay ............... 94  
  4.3.2 Asymmetric Traffic ........................................ 99  
  4.3.3 Semi-Two-Way Relaying .................................. 99  
4.4 Multiuser Zero-Forcing Relaying ............................... 100  
4.5 Summary ......................................................... 103

5 A Novel Cooperative MAC Protocol  
5.1 Cluster-Based CSMA/CA for Cooperative Wireless Networks  
  5.1.1 Basics of CB-CSMA/CA ..................................... 107  
  5.1.2 Collision Post-Detection .................................. 111  
  5.1.3 Impact of the PHY layer on the CB-CSMA/CA ........... 112  
  5.1.4 Impact of Control Frame Error ............................ 115  
  5.1.5 Cluster-Based CSMA/CA with RTS/CTS .................... 117  
  5.1.6 Cluster-Based Centralised Channel Access ............... 117  
  5.1.7 Time and Frequency Synchronisation ..................... 118  
5.2 Outline of Representative CB-CSMA/CA Applications .......... 121  
5.3 CB-CSMA/CA Throughput and Delay Investigation ............. 122  
  5.3.1 Application A .............................................. 122  
  5.3.2 Application B .............................................. 126
## Contents

5.3.3 Application C .................................................. 129
5.4 Summary .......................................................... 132

6 Network Performance of CB-CSMA/CA .................................. 133
6.1 Application-Specific Parameters ........................................ 134
6.2 Outage Probability .................................................. 136
6.3 Collision Probability ................................................ 139
6.4 Throughput .......................................................... 140
6.5 Delay ............................................................... 144
6.6 Summary ............................................................ 144

7 Cross-Layer Enhancements of CB-CSMA/CA ......................... 145
7.1 Impact of Backoff Procedure on CB-CSMA/CA Performance .... 145
7.1.1 Backoff Models ................................................. 146
7.1.2 CB-CSMA/CA Modifications .................................. 152
7.1.3 Performance Results .......................................... 154
7.2 Impact of Clustering on CB-CSMA/CA Performance ............. 160
7.2.1 Clustering According to the Higher Layer Parameters ....... 161
7.2.2 Clustering According to the PHY Layer Parameters .......... 164
7.3 Summary ............................................................ 171

8 Performance Bounds of CB-CSMA/CA ................................ 173
8.1 Impact of Event-Synchronisation Errors on CB-CSMA/CA Performance .... 173
8.1.1 Assumptions and Model ....................................... 174
8.1.2 Symmetric Scenario ............................................ 174
8.1.3 Asymmetric Scenario .......................................... 179
8.1.4 General Scenario .............................................. 183
8.2 Comparison with the Multipacket Reception Protocol .......... 184
8.2.1 Saturation Throughput ........................................ 186
8.2.2 Non-saturation Throughput .................................. 189
8.3 Summary ............................................................ 191

9 Outlook for Multihop Communications .................................. 193
9.1 Conceptual Discussion ............................................. 193
9.2 A Specific Proposal ............................................... 197
9.3 Summary ............................................................ 201
Part I

Fundamental Limits of the Existing WLANs
Chapter 1

Introduction

1.1 Motivation

During the last decade, wireless local area networks (WLANs) have received lots of attention. Fast growth of WLAN infrastructures on one side and significant increase in demand for wireless communications on the other side, have pushed both industry and scientific community to further develop the WLAN technology. Nowadays one can access WLANs almost everywhere. Hotels, airports, coffee shops and even many downtowns are covered by wireless hotspots. We are used to accessing WLANs everywhere and it is hard to imagine a world without WLANs anymore.

WLANs have been originally designed for point-to-point data communications. However during the last few years, applications have become very diverse, ranging from voice-over-WLAN to video streaming. Furthermore, originally IEEE 802.11 systems could only support data rates up to few Mb/s. Later on, the IEEE 802.11 standardisation bodies have developed several amendments succeeding the legacy standard to support higher data rates, quality-of-service (QoS) and enhanced compatibility, security etc.

The IEEE high-throughput standard, i.e. IEEE 802.11n, supports higher data rates up to 600 Mb/s for point-to-point links applying multiple-input multiple-output (MIMO) techniques [1]. However, to support this rate the following conditions should be satisfied:

1. Four transceivers at each transmit and receive side
2. Rich scattering environment
3. Strong channel on all spatial links in order to enable each link to support 64-QAM (quadrature amplitude modulation) with coding rate of 5/6
4. A guard interval (GI) of 400 ns, hence the longest path should experience a delay smaller than 400 ns

5. Bandwidth of 40 MHz

6. A medium access control (MAC) layer efficiency of 100%

Since in practice one (for example requirement 6) or more of these requirements cannot be satisfied, the achievable rate becomes lower than 600 Mb/s. Besides, higher data rates can only be achieved in a relatively small area. Furthermore, in many networks, nodes may still be equipped with only one transceiver for simplicity or cost reasons. Therefore to enhance performance of WLANs, it is worthwhile to consider other approaches.

The IEEE 802.11 "very high throughput" task group has also realised the need for further enhancements and is looking for techniques to improve WLAN systems [2, 3]. Currently, the task group (TG) ac considers operating at frequency band below 6 GHz while TGad is responsible for investigating operating at 60 GHz. Both groups are at the very preliminary phases and there is still no public draft available. However, some preliminary concepts exist which suggest the employment of multiuser MIMO techniques [4]. In addition to standardisation development, many research activities on different aspects of WLANs are in progress.

Multiuser MIMO techniques can also improve performance of wireless systems. In fact multiuser MIMO techniques can provide spatial multiplexing gain even in networks with single-antenna stations (STAs). However, to realise multiuser MIMO transmissions in WLANs we need to modify the existing WLAN systems. The fundamental channel access method of the IEEE 802.11 MAC is the distributed coordination function (DCF) which is based on carrier sense multiple access with collision avoidance (CSMA/CA) [5]. DCF has been originally designed for point-to-point links and prevents distributed spatial multiplexing by the collision avoidance mechanism. Therefore to support multiuser MIMO transmissions, modifications of MAC is necessary. Besides, the current WLAN MIMO physical layer (PHY) should also be adapted in a way that multiuser interference can be cancelled out.

In this thesis, we focus on WLAN as a widely-used and an evolving technology. We aim to enhance performance of the existing WLANs. In order to do so we focus on PHY and MAC layer cooperative protocols which can enable distributed spatial multiplexing. At first we investigate main limitations of the existing WLANs and then propose a novel protocol to overcome some of these limitations.
1.2 State-of-the-Art, Contributions and Outline

Part I: Fundamental Limits of Existing WLANs

In the first part of the thesis the existing WLANs are considered. At first motivation, contribution and outline of the thesis are presented. Then a comprehensive study of the existing WLANs is described.

To start with, a short review of WLAN standards is given in Chapter 2. Then a WLAN project is briefly presented which provides high quality voice over WLAN by integrating professional quality telephony into WLANs. Within this project framework we focused on the coverage investigation of the IEEE 802.11a system and the diversity and the relaying techniques to enhance the range. In addition to the project reports, some ideas and results have been published in:


Later in Chapter 3 performance bounds of the current single-input single-output (SISO) WLANs with respect to coverage range, outage probability, MAC layer throughput and delay are investigated. Although there are several publications which take different aspects of WLANs into account, such as [6–9], we intend to analyse all interesting aspects of WLANs using the chosen parameters and channel model. While studies in [6–8] focus on MAC throughput of IEEE 802.11 networks, [9–11] look into PHY layer aspects of OFDM-based WLANs as explained in the following paragraphs.
Chapter 1 Introduction

The basic channel access method of IEEE 802.11 is based on CSMA/CA where nodes try to avoid collision by setting their backoff window size randomly. In [6], a two-dimensional Markov Chain model for the backoff window size has been introduced. The model can be used to calculate the saturation throughput in scenarios with perfect channels, i.e. cases without packet error. On the other hand the authors of [7] have introduced an average cycle time approach to compute the throughput. They show that there exist a cycle time during which each node transmits on average one packet successfully. There are several works which are based on the model introduced in [6] while extending that in different ways. For example the authors of [8] have extended the model in [6] to account for both channel error and non-saturation conditions. As it will be explained in detail in Chapter 3, throughout this thesis the MAC throughput analysis is based on the model in [6]. Furthermore, we mainly focus on the scenarios where STAs transmit with the same data rate. Nevertheless the extension to multirate scenarios can be easily done as it has been explained in [12] [13].

In this thesis we consider outage probability and data rate as the main PHY layer metrics. The IEEE 802.11 defines different PHY layer techniques. The IEEE 802.11a, 802.11g and 802.11n are based on orthogonal frequency division multiplexing (OFDM). Different works have been performed on analysis the existing PHY or techniques to improve them [10, 11]. Some of them dealt with the packet error ratio estimation in WLANs. However, most works have considered measurement or simulations to compute the error probabilities. In [9] the authors have developed an analytical packet error rate expression for the OFDM-based WLAN with frequency-selective channel. However, while the analytical results are fairly accurate for the cases with convolutional coding rates of $1/2$, they need to be compensated for other cases.

In addition to SISO systems we also consider point-to-point WLAN MIMO systems and relaying networks in Chapter 3. It is known that MIMO techniques can enhance performance of a wireless system [14–17]. A comprehensive study on MIMO theory is given in [14] and [15]. By applying MIMO techniques one can achieve spatial multiplexing gain and improve sum rate and/or achieve diversity gain and enhance the link performance. However, there is a tradeoff between spatial multiplexing and diversity gains [17].

In addition to MIMO techniques, relaying methods can also enhance coverage range or data rate in a network. Relaying channels have been first considered in the seventies [18] and later have further been investigated [19–21]. In [19] probability of outage in different relaying networks, including decode-and-forward (DF) and amplify-and-forward (AF) relaying, is estimated. A study of outage probability in relaying networks for fading channels and the impact of power control on the outage behaviour have been reported in [20]. Coding
strategies for relay networks with cooperative nodes are developed in [21] where decode-and-forward and compress-and-forward relays are considered. Throughout the thesis we take DF and AF relays into consideration.

The discussion and analysis in the first part of the thesis show that there is still a lot of potential for improving WLAN performance. Our analysis in this part, provides us with the methodology to be applied for the analysis in the second part of the thesis. Besides, in this part we introduce main parameters and explain all performance criteria which are used in this thesis.

**Part II: Cooperative WLAN Protocols**

Multiuser MIMO techniques can exploit MIMO gains even in networks with single-antenna STAs [14] [22]. Furthermore, it has been shown that cooperative relaying protocols can extend the coverage range or the data rate of a given system [23]. However, current WLANs only support point-to-point MIMO and do not define any wireless relaying protocol except for the access point (AP). In an infrastructure network the AP can act as a relay and forward the packets from a source to its destination within the same basic service set [5]. In the second part of the thesis, cooperative protocols are taken into account to enhance performance of WLANs.

In Chapter 4 we focus on the PHY layer and investigate the integration of three representative cooperative techniques into WLAN. We have evaluated PHY data rate and outage probability in all three cases. In all of these cases there are enough degrees of freedom either at receivers or at some cooperative relays to enable multiple STAs to transmit simultaneously.

1. As the first application, we consider a network with a multiple-antenna DF relay, where multiple STAs in uplink transmit concurrently to a DF relay. The DF relay can also forward multiple independent streams during the downlink. In this application, the uplink constitutes a multiple access channel while the downlink establishes a broadcast phase. Signal processing and achievable rates in both multiple access and broadcast phases have been widely studied in the literature. During the multiple access phase two or more nodes transmit simultaneously to a single destination while in the broadcast phase, a single transmitter sends packets to several nodes concurrently. This scenario can be a typical application in a WLAN with a multiple-antenna AP. An analysis of the rate region for both multiple access and broadcast phase with single-antenna as well as multiple-antenna nodes are given in [14] [15] [24].

2. The second application is a two-way relaying application, where two STAs simultaneously transmit to a DF two-way relay and the relay combines the received messages
and forwards them to both STAs. The two-way relay channel has first been considered in [25] and later has been studied in other publications [26]. Contrary to [26] in which the authors optimised power allocation at the relay with respect to sum rate, here we find a simple power allocation to improve the link rate. The results have partially been published in:


Additionally we have analysed the two-way relaying network throughput in:


3. In the third application, an AF relaying network has been considered. In this application multiple STAs transmit to their destinations via several AF relays. The AF relays cancel the multiuser interference by using proper amplification gains [27].

All considered applications, as well as many other cooperative schemes, require concurrent transmissions from multiple users. However, current IEEE 802.11 systems using CSMA/CA allow only a single user to transmit in its neighbourhood. Hence, to solve this problem a new MAC protocol or modifications of the existing MAC are needed. This problem has already been discussed in some other publications.

A MAC protocol with antenna arrays is suggested for ad hoc networks in [28]. The antenna array is used to null the interference from neighbours. In this work as well as in [29], a control channel has been introduced such that nodes can learn about ongoing transmissions by monitoring this channel. Park et al. [30] have proposed a MAC protocol mitigating interference by using multiple antennas at the receivers. In [31], the authors propose a MAC protocol which considers the spatial correlation between the signal and interference to decide whether multiple links should transmit simultaneously or not. In their proposal, links are allowed to contend for the channel sequentially while transmitting data packets simultaneously. In [32], authors have proposed a slightly modified version of the existing MAC to support multiple streams. In their scheme whenever more than one user transmits in a given slot, if the receiver can decode multiple packets it simply acknowledges them.

To enable multiple concurrent transmissions in a WLAN, we propose a novel cluster-based MAC as the extension of CSMA/CA in Chapter 5. The proposed protocol facilitates
1.2 State-of-the-Art, Contributions and Outline

cooperative techniques by utilising the multiuser interference mitigation capability of the PHY.

The main idea of the proposed cluster-based CSMA/CA (CB-CSMA/CA) is to group STAs into clusters such that STAs belonging to the same cluster can transmit at the same time. As it is explained in Chapter 5 the CB-CSMA/CA scheme has some distinct advantages compared to the existing proposals, as it requires neither sequential contention per node for the data transmission as needed in [30] and [31], nor a control channel as in [28] and [29]. Besides, by grouping nodes into clusters, the collision probability of the CB-CSMA/CA depends on the number of contending clusters rather than the total number of contending stations.

In Chapter 5, the CB-CSMA/CA protocol description is presented and the impact of signal processing on the MAC protocol design has been discussed. To the best of our knowledge, this is the first work which takes into account the impact of AF relaying and its differences to DF relaying with respect to the IEEE 802.11 MAC protocol. Furthermore, throughput and delay equations are investigated for all three considered applications as well as their respective reference systems operating based on the IEEE 802.11 MAC.

In Chapter 6, we investigate network performance of the three considered applications, operating based on CB-CSMA/CA, by taking into account both MAC and PHY layer parameters. We compare their performance with that of a respective reference system operating based on the IEEE 802.11 MAC. The throughput and delay results show that the CB-CSMA/CA outperforms the IEEE 802.11 MAC significantly. In the considered infrastructure network with a four-antenna AP, the CB-CSMA/CA uplink has an aggregate throughput about 2.5 times greater than that in the reference system. This throughput gain has been achieved despite three times higher link rate in the reference system. Some of the results and the basic idea of CB-CSMA/CA protocol have been published in:


In Chapter 7, we extend the study to include different backoff models and clustering issues. The study of three backoff models, helps us to choose the appropriate model for each circumstance. Besides, we show that in many setups it is beneficial to allocate STAs with similar requirements/conditions (including average channel condition) into the same group. We consider a realistic scenario, where different sources and destinations are distributed in a given area such that they have different distances to the AP. Simulations show that by ap-
appropriate clustering we can gain a lot compared to a random clustering. In addition to the optimum clustering found by exhaustive search, a heuristic method has also been suggested.

The CB-CSMA/CA study up to this point has been based on the ideal performance of the protocol. In Chapter 8, we investigate the performance bounds of the protocol in imperfect situations. The analytical study in this chapter shows that the CB-CSMA/CA is robust to protocol synchronisation errors. In fact in a worst case scenario, where all STAs within a cluster transmit asynchronously, the CB-CSMA/CA performs similar to an IEEE 802.11n system. Simulation results validate our analytical model in the presence of synchronisation errors. Furthermore, we compare the throughput of the CB-CSMA/CA protocol with that of another multipacket protocol proposed in [32]. We show that under saturation condition, the CB-CSMA/CA outperforms the other protocol significantly. In this chapter we also take the non-saturation throughput into account and analyse the performance based on different clustering approaches. We show that in non-saturation scenarios we can benefit from adaptive clustering where stations are grouped into clusters according to the presence of traffic. Part of the results in this chapter is submitted to the EURASIP Journal on Wireless Communications and Networking:


In the last chapters of this thesis, outlook and conclusion are presented. In this thesis the emphasis has been on single- and two-hop communication links, which are the most typical patterns in infrastructure and many ad hoc networks. However, multihop links with more than two hops should also be supported in different types of networks. An outlook for multihop communications is presented in Chapter 9, where possible extension of the proposed CB-CSMA/CA to support multihop links are also discussed.

Our study shows that the CB-CSMA/CA is a promising protocol which utilises distributed spatial multiplexing in networks with multiuser interference cancellation capability. It is a simple but powerful method and is backward compatible with the existing IEEE 802.11 systems. The CB-CSMA/CA can be used in different types of networks including infrastructure, ad hoc and relaying networks.
Chapter 2

Basics of Existing WLANs

Currently, WLAN systems are one of the most frequently used wireless systems in indoor environments. The IEEE 802.11 [5] is a family of standards for wireless LAN medium access control and physical layer specifications.

In this chapter, first we discuss main features of the IEEE 802.11 systems with emphasis on those aspects which are more relevant to our work and referred to in the next chapters. Then as an example of a WLAN system, we briefly present the main features of a project, called WINDECT. The aim of the project has been integration of professional quality telephony based on DECT systems into WLANs [33]. Furthermore, to show the coverage limitations of the off-the-shelf WLAN devices, the results of a coverage measurement campaign carried out in a typical office building are also presented in this chapter.

2.1 IEEE 802.11

The IEEE 802.11 is the legacy standard for local and metropolitan area networks. The legacy 802.11, which deals with operation in 2.4 to 2.5 GHz band, defines several PHY layer signalling techniques and MAC protocol for wireless data communication.

The basic service set (BSS) is defined as the basic building block of an IEEE 802.11 network [5]. It is assumed that the stations within a BSS are in the communication range of each other and if one station moves out of its BSS it cannot anymore communicate with any other member of that BSS directly. Two types of networks are defined in this standard:

- The independent BSS (IBSS) or ad hoc network: this type of network includes only stations which are in the communication range of each other. The main characteristic of an ad hoc network is that it has a limited spatial as well as limited temporal extent.
In an IBSS, access to a distribution system (DS)\textsuperscript{1} is not available. STAs in an ad hoc network communicate directly with each other.

- Infrastructure network: an infrastructure network consists of one or more access point (AP) and zero or more portal entities as well as the DS medium. STAs in an infrastructure network can communicate with each other via an AP.

Two different coordination functions are defined in the legacy standard. The fundamental access method is a distributed coordination function (DCF) based on CSMA/CA. According to CSMA/CA, each station, which has a packet to transmit senses the channel. If the channel remains idle over a certain predefined duration, called DCF interframe space (DIFS), it starts the transmission otherwise it defers until the end of the current transmission. To avoid collisions, after a successful transmission or a deferral, the STA has to select a random backoff interval. The backoff interval counter is decremented while the medium is idle. The basic access mechanism is shown in Figure 2.1 where DIFS, SIFS, and PIFS indicate DCF interframe space, short interframe space and point coordination function interframe space. All of these values are defined for a specific PHY by the standard.

In addition to the DCF basic method, it is possible to exchange short control frames, i.e., request-to-send (RTS) and clear-to-send (CTS), between the transmitting and the receiving STA prior to the data transmission. By utilising these short control frames, the probability of collision is reduced. DCF can be applied to both ad hoc and infrastructure networks.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{basic_access_method.png}
\caption{Basic access method \cite{腳注5}.}
\end{figure}

The STA should set its backoff time according to the following equation:

\begin{equation}
\text{backoff time} = \text{Random} \times a\text{SlotTime}, \tag{2.1}
\end{equation}

\textsuperscript{1}The distribution system is the architectural component which is used to interconnect BSSs.
2.1 IEEE 802.11

where $aSlotTime$ is a value fixed per PHY and $Random$ is a pseudo-random integer which is taken from a uniform distribution over $[0,CW]$ with contention window (CW) being an integer within the range of $CW_{\min}$ and $CW_{\max}$. Both $CW_{\min}$ and $CW_{\max}$ values are determined for each specific PHY by the standard. The CW parameter takes an initial value of $CW_{\min}$. Whenever the STA transmission fails, it increases its CW to sequentially ascending integer powers of two, minus one up to the $CW_{\max}$ value. After every successful transmission the CW is reset to $CW_{\min}$. The performance of the DCF access method with respect to the throughput and delay under different conditions, like saturated or non-saturated traffic and non-ideal transmission channel, has been investigated in several research works.

Transmission and collision probabilities in the case of saturated traffic with unlimited retransmissions can be calculated using a Markov chain (MC) model, as in [6], or estimated based on mean value analysis similar to [34]. In [35] the authors also consider a stochastic analysis of the backoff window size based on a Markov chain but for a limited number of frame retransmissions.

Another coordination function defined in IEEE 802.11, is the point coordination function (PCF). PCF may be optionally used in infrastructure networks to support real-time applications. According to PCF, a point coordinator allocates slots to STAs for data transmission. Transmission under PCF is contention-free (CF). In this case contention-free period (CFP) should alternate with a contention period (CP). PCF has rarely been considered due to certain severe problems, like unpredictable beacon delays and unknown transmission durations.

The legacy IEEE 802.11 standard consists of several amendments which have been proposed to fulfill different requirements. Some of them need either extensions of the PHY layer, or modifications of the MAC layer, while some others may require extensions of both MAC and PHY layer features. For example, the IEEE 802.11a, 802.11b and 802.11g all specify the PHY.

The IEEE 802.11b [36] has been designed to support some higher rates than the IEEE 802.11, it provides data rate of 5.5 and 11 Mb/s in addition to the 1 and 2 Mb/s.

The IEEE 802.11a [37] and 802.11g [38] both specify a system based on OFDM supporting higher rates up to 54 Mb/s. The IEEE 802.11a supports application in the 5 GHz band while systems based on 802.11g operate in the 2.4 GHz band. To support quality-of-service (QoS) requirements, the IEEE standardization bodies have proposed amendment IEEE 802.11e [39] which enhances the legacy MAC. Recently IEEE 802.11n [1] task group has specified MAC and PHY layers, to enhance the data rate of a WLAN using multiple-input multiple-output techniques. As it has been mentioned in Chapter 1, the 802.11n supports a data rate up to 600 Mb/s (uncoded data rate of 720 Mb/s).
In practice the achievable data rate for any of the above-mentioned systems is less than the aforementioned values due to following reasons: The wireless channel is vulnerable to noise, fading and interference and it changes over time. Path loss, shadowing, and the characteristics of the propagation channel impact the received signal power. The signal power is attenuated by increasing the distance between source and destination. The value of the path loss exponent depends on the environment and the availability of the line-of-sight (LoS) component. Besides, operation at higher frequencies results in higher path loss as compared to the lower frequencies.

Another important factor which degrades the throughput of a system is the collision probability. In a WLAN several nodes may share the same channel and their transmission is orthogonalised in a time division multiple access (TDMA) manner, i.e., in the ideal case at any time instant only one node is allowed to transmit. Consequently, the throughput of each user decreases by increasing the number of users. Besides, the mandatory access method of the IEEE 802.11 systems is based on the CSMA/CA. The CSMA/CA has always a maximum normalised saturated throughput below 100% even with a perfect PHY [6]. This is due to the waiting times, signalling overheads as well as the time wasted in collisions. Although the CSMA/CA avoids collision by sensing the channel prior to transmission, the collision occurrence is not completely prohibited.

The IEEE standard documents usually present the bandwidth and the maximum allowed transmission power which are specified for applications in USA. For other countries the local standardisation bodies define the available channels at each frequency band as well as the maximum transmit power which is allowed to be used in a certain band. In Switzerland the Federal Office of Communications fulfills the regulatory tasks related to telecommunications and broadcasting [40].

In this section, the main specifications of the IEEE 802.11a, 802.11e and 802.11n which are referred to in the upcoming chapters are briefly presented.

### 2.1.1 IEEE 802.11a

The IEEE 802.11a has standardised transmission in the 5 GHz band based on the OFDM technique. According to the 802.11a amendment, the 20 MHz transmission bandwidth is divided into 64 subcarriers, 48 of them are allocated to data transmission while 4 of them are allocated to the pilots and the rest are the zero subcarriers. One OFDM symbol is 4 μs long from which 0.8 μs is allocated to the guard interval. The IEEE 802.11a provides communication data rates of 6, 9, 12, 24, 36, and 54 Mb/s by varying the modulation alphabet and the
2.1 IEEE 802.11

forward error correction (FEC) coding rate. Among these data rates, only support of transmission at 6, 12 and 24 Mb/s is mandatory. The modulation alphabet can be chosen from binary phase shift keying (BPSK), quadrature phase shift keying (QPSK), QAM or 64-QAM and the FEC coding rate can be set to 1/2, 2/3 or 3/4. For example by choosing the BPSK and a coding rate of 3/4, a data rate of 9 Mb/s is provided which corresponds to an uncoded data rate of 12 Mb/s\(^2\).

2.1.2 IEEE 802.11e

To overcome existing problems of the legacy MAC and to support QoS requirements, the IEEE task group has defined the amendment IEEE 802.11e [39]. The IEEE 802.11e offers a hybrid coordination function (HCF). The basic unit of allocation of the right to transmit is called the transmission opportunity (TXOP), which is defined by a starting time and a maximum duration. The STAs can either obtain the right to transmit by winning enhanced distributed channel access (EDCA) contention during the CP, or by receiving a CF-Poll during the CP or CFP.

In CFP, the hybrid coordinator (HC), which is collocated with the AP, allocates time slots to the terminals by a polling mechanism according to hybrid coordination function controlled channel access (HCCA). During CFP, the HC receives add traffic stream requests from non-AP QoS stations and the HC may accept or deny the requests according to an admission control policy.

The HCF differs from the legacy MAC in several aspects, including the possibility of prioritisations among STAs with different traffic classes and performing polling during both CP and CFP. Priorities are differentiated by varying CW limits, the arbitration interframe space (AIFS), and the TXOP limit. Each station starts a backoff, after detecting the channel being idle for an AIFS. AIFS has a length equal to or larger than the DIFS. According to the IEEE 802.11e there are altogether four access categories (AC), i.e., in ascending order of priority, access category background, best effort, video, and voice.

In addition to the above features, the IEEE 802.11e also supports direct link setup (DLS) between two stations in the infrastructure mode [39]. To initiate a direct link, the transmitting STA, sends a DLS request to the AP. The AP forwards the request to the recipient. If the receiver accepts the DLS, it sends a DLS response to the AP and the AP forwards it to the transmitter and the direct link becomes active. Figure 2.2 shows the four steps of DLS handshake.

---

\(^2\)For more details on the rate dependent parameters of the IEEE 802.11a see Table 78 in [37].
Furthermore the IEEE 802.11e has introduced the block acknowledgement (ACK) mechanism to improve efficiency. According to this mechanism a block of data frames can be acknowledged with a single ACK frame.

It has been shown that in a heterogenous system where nodes with different priorities exist, the performance of high-priority AC nodes is much better as compared to those in a legacy system while performance of low-priority nodes is significantly worse [41].

### 2.1.3 IEEE 802.11n

The IEEE 802.11n task group specifies MAC and PHY layers to enhance the data rate of a WLAN using MIMO techniques, like transmit beamforming and space-time block codes (STBC) or hybrid STBC/spatial multiplexing [1].

The IEEE 802.11n PHY is also based on OFDM technology operating at 2.4 or 5 GHz frequency band and is backward compatible with 802.11a and 802.11g. Some of the OFDM parameters of the high throughput (HT) mode [1] of the IEEE 802.11n are different from those of 802.11a. In the HT mode there is a possibility to choose between 20 and 40 MHz bandwidth. Besides, one can set the guard interval to 0.4 or 0.8 $\mu$s. According to the IEEE 802.11n, support of 20 MHz with one and two spatial streams is mandatory at APs while only support of 20 MHz with one spatial stream is mandatory at non-AP STAs, here called STAs. However, in order to achieve higher data rates, spatial multiplexing is necessary. The maximum IEEE 802.11n data rate with one and two spatial streams are 150 Mb/s and 300 Mb/s, respectively.

An OFDM symbol in the 20 MHz HT mode consists of 52 data subcarriers and 4 pilots.
### Table 2.1: IEEE 802.11n MCS for 20 MHz, GI of 800ns and one spatial stream.

<table>
<thead>
<tr>
<th>MCS Index</th>
<th>Modulation</th>
<th>FEC Rate</th>
<th>$N_{\text{CBPS}}$</th>
<th>$N_{\text{DBPS}}$</th>
<th>Data rate (Mb/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>BPSK</td>
<td>1/2</td>
<td>52</td>
<td>26</td>
<td>6.5</td>
</tr>
<tr>
<td>1</td>
<td>QPSK</td>
<td>1/2</td>
<td>104</td>
<td>52</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>QPSK</td>
<td>3/4</td>
<td>104</td>
<td>78</td>
<td>19.5</td>
</tr>
<tr>
<td>3</td>
<td>16-QAM</td>
<td>1/2</td>
<td>208</td>
<td>104</td>
<td>26</td>
</tr>
<tr>
<td>4</td>
<td>16-QAM</td>
<td>3/4</td>
<td>208</td>
<td>156</td>
<td>39</td>
</tr>
<tr>
<td>5</td>
<td>64-QAM</td>
<td>2/3</td>
<td>312</td>
<td>208</td>
<td>52</td>
</tr>
<tr>
<td>6</td>
<td>64-QAM</td>
<td>3/4</td>
<td>312</td>
<td>234</td>
<td>58.5</td>
</tr>
<tr>
<td>7</td>
<td>64-QAM</td>
<td>5/6</td>
<td>312</td>
<td>260</td>
<td>65</td>
</tr>
</tbody>
</table>

[Figure 2.3: The IEEE 802.11n HT-mixed format PLCP protocol data unit (PPDU). PSDU is the (PLCP) service data unit [1].]

... out of totally 64 subcarriers, while in the 40 MHz operation mode there are 108 data subcarriers and 6 pilot subcarriers out of total 128 subcarriers. Table 2.1 shows all possible IEEE 802.11n modulation and coding schemes (MCS) for a single spatial stream, where $N_{\text{CBPS}}$ and $N_{\text{DBPS}}$ are respectively, the number of coded bits and data bits per OFDM symbol.

In both IEEE 802.11a and 802.11n systems, the pilots are used to make the coherent detection robust against frequency offsets and phase noise. The long training sequence at the beginning of the frame is used for channel estimation. The IEEE 802.11n introduces three different preamble formats: Non-HT format, HT-mixed format and HT-greenfield format, where HT stands for high throughput. The non-HT format is defined according to the SISO standards. The HT-mixed format, which is used throughout this thesis, contains a preamble which is compatible with the non-HT standards as well as a part which is particular for HT devices and cannot be decoded by non-HT devices. The HT-greenfield format does not contain a non-HT compatible part and its support is optional. It should be noted that, although non-HT preamble and MAC headers are slightly shorter than their corresponding
Chapter 2 Basics of Existing WLANs

<table>
<thead>
<tr>
<th>Octets:</th>
<th>2</th>
<th>2</th>
<th>6</th>
<th>6</th>
<th>6</th>
<th>4</th>
<th>0-7955</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frame Control</td>
<td>Duration/ID</td>
<td>Address 1</td>
<td>Address 2</td>
<td>Address 3</td>
<td>Sequence Control</td>
<td>Address 4</td>
<td>QoS Control</td>
<td>HT Control</td>
</tr>
</tbody>
</table>

MAC Header

Figure 2.4: The IEEE 802.11n MAC frame format [1].

![Figure 2.4: The IEEE 802.11n MAC frame format](image)

Figure 2.5: A-MSDU encapsulation [42].

![Figure 2.5: A-MSDU encapsulation](image)

values defined by HT IEEE 802.11n specifications, in order to be consistent, throughout this thesis we will use headers according to the HT-mixed format for all IEEE 802.11 numerical results, cf. Figure 2.3, and Figure 2.4.

According to the standard 802.11n, a high throughput STA has MAC features which include block acknowledgement, in a similar way as defined in the IEEE 802.11e, frame aggregation, power-save multi-poll operation, and reverse direction as explained in the following paragraph [1].

According to the IEEE 802.11n, aggregation can be applied to the MAC service data unit (MSDU) or to the MAC protocol data unit (MPDU). In the first case, i.e., aggregate MSDU (A-MSDU), several MSDUs which are destined for the same receiver and have the same traffic identifier can be aggregated in a single MPDU. In the second case, i.e., aggregate MPDU (A-MPDU), the fully formed MPDUs are aggregated in a single MPDU, such that each MPDU component has a short MPDU delimiter [42]. Figures 2.5 and 2.6 illustrate these two aggregation methods. In these figures DA, SA, FCS and CRC denote destination address, source address, frame check sequence, and cyclic redundancy check, respectively.

The power-save multi-poll (PSMP) mode provides the schedule at the beginning of the

---

3 According to the IEEE 802.11e a QoS facility uses a traffic identifier to determine differentiated services per each MSDU.

4 The MPDU delimiter is used such that the structure of the A-MPDU can be recovered when one or more MPDU delimiters are received with errors.
2.2 WINDECT

During the last few years, WLAN applications have become very diverse, ranging from voice to video streaming. WLANs were originally designed for data transmissions. Hence, to support such a wide range of applications, we should deal with several challenges. In this section, we have referred to a project, called WINDECT, which has aimed to integrate professional quality telephony into WLAN. This project has been one of the first projects, which used IEEE 802.11e to provide high QoS voice-over-WLAN. In this section, we describe main specifications of this project.

One of the main challenges for real-time applications over WLAN is to provide full coverage within an area. Here we briefly present results of our coverage measurement campaign carried out in a typical office building. The results show that the coverage range of an IEEE 802.11a AP is relatively small. Within the framework of this project we have investigated diversity and different relaying techniques for point-to-point links to extend this

PSMP phase, so STAs can shut down until they receive or transmit packets according to the schedule. Consequently in this mode the power consumption is reduced.

Finally, the reverse direction protocol allows the owner of a TXOP to allocate the unused part of the TXOP to its communication peer. This can be done when uplink and downlink are not symmetric. Often the uplink is underutilised, for example when only ACK frames are transmitted during the uplink. Hence, the reverse direction protocol can lead to a fully utilised TXOP.

**Figure 2.6:** A-MPDU encapsulation [42].
Chapter 2 Basics of Existing WLANs

range. We have included the techniques which we used for our coverage analysis and coverage extension, in Chapter 3, where a comprehensive study of the existing WLANs is given. In the second part of the thesis we proceed by taking cooperative relaying techniques into consideration.

2.2.1 WINDECT Specifications

Wireless LAN with integration of professional-quality digital enhanced cordless telecommunications (DECT) [43] telephony (WINDECT), was a European project under the sixth framework program. WINDECT has been based on a novel approach to merge wireless networks by integrating professional quality telephony into WLANs to obtain high quality voice over WLAN. DECT is well-known for its high quality voice transmissions over a low-bandwidth while WLANs provide a high data rate but best-effort bursty traffic.

The WINDECT PHY layer is mainly based on the IEEE 802.11a standard and its MAC sublayer is designed according to the IEEE 802.11e. Main features of these standards are described in Section 2.1. It should be noted that at the time of this project, the IEEE 802.11n standard did not exist.

DECT is a low-power digital wireless communication system and is the standard for digital cordless telephony. It is designed for wireless use and applied worldwide for professional-quality telephony. It uses 10 ms adaptive differential pulse code modulation (ADPCM) frames. ADPCM has a moderate complexity and causes almost no extra delay to the system. In DECT systems the standard speech packets with a 64 kbps are compressed into 32 kbps ADPCM [43, page 33].

DECT has a strict TDMA structure and consequently provides a high QoS. In order to achieve such high QoS and to ensure that each STA has a guaranteed transmission time, the polling access method of IEEE 802.11e is used in WINDECT. The AP sends a voice packet to each STA with a piggybacked CF-poll. It ends the CFP once all STAs with active voice connections are polled.

Data packets are transferred during the CP. In WINDECT a maximum CFP duration of 5.9 ms is considered. In the considered WLAN MAC the beacon intervals are multiples of 1024 $\mu$s. This makes the direct mapping to 10 ms DECT frame intervals impossible. In WINDECT, the beacon interval is set to 10.24 ms but as explained on the next page every STA receives voice packet only every other beacon interval.

An in-house network with DECT over WLAN looks like the network depicted in Figure 2.7. As it is shown in this figure, the old solution, where in-house wireless networks
The WINDECT approach is to merge the upper layers of the DECT protocol stack with an existing WLAN physical layer and MAC layer technology, using a protocol adaptation layer (PAL) as shown in Figure 2.8. As the higher layers of DECT are optimised for the voice transmission, a high voice quality is expected.

In addition to that, DECT is easily implemented on lightweight devices and its component costs are very low while their battery life is quite long. Although DECT and WLAN each can bring certain advantages to the system, the integration of both in one merged network leads to some challenges, including, PAL design, handover, QoS issues, power consumption, and voice optimisation. These aspects are studied in the framework of the project by several partners. In the following sections, we briefly consider some of these challenges with focus on our main research work in this project, i.e., coverage investigation and enhancement.

**PAL Design:** The PAL needs to map functionality of DECT into that of the 802.11 WLAN and vice versa. Some features like authentication are provided by both systems. Therefore the PAL should disable one of them. Some other concepts are similar but with slightly different functionality in the two systems. In these cases the PAL should map between the concepts. In cases of totally different concepts, like handover, the PAL needs to hide these...
Figure 2.8: The WINDECT protocol Stack [44].

differences while still providing the necessary services [45].

**Seamless Handover:** DECT provides seamless handover, by allowing a STA to be connected to two different base stations at the same time. However, in a WLAN each STA can only be associated with a single AP. Therefore, to support seamless handover a modification in the WLAN protocol is required.

In WINDECT a novel approach was proposed to solve this association problem, as explained in the following paragraph. In WINDECT a beacon interval is set to 10.24 ms but every STA receives packets only every other beacon interval [46]. In this way, the STA has enough time to scan for other AP’s beacons and a second possible association. The desired maximum of 20 supported STAs per AP is reached by assigning half the connections to the even and the other half to the odd beacons.

To avoid problem with double association, the DS is not informed about the second association as long as the handover is not completed. After completion of the handover process the STA is disconnected from the first AP and the DS is informed about the actual AP connection [44,45].

### 2.2.2 Coverage Measurement

In order to compare the coverage range of an IEEE 802.11a network with that of a DECT system, a measurement campaign has been carried out in a typical office environment at our university. Two devices were used, a stand-alone access point and a CardBus card in a laptop
2.2 WINDECT

No or poor reception (FER > 5%)

Full coverage (FER ≤ 5%)

Figure 2.9: Map of arena showing FER for data rate of 6 Mb/s.

Figure 2.10: Map of arena showing RSSI.
computer. An Atheros test program was used to perform the measurements. This test program is primarily intended for testing the hardware during development and manufacturing. It allows fine control over the hardware in a way that is not possible with the normal software intended for production use in real networks.

The software was used in its link test mode where a unidirectional link is used. The access point was configured as the receiver and the laptop as the sender. The sender sent 500 packets at each data rate. The transmit power was set to its maximum allowed value in Switzerland which at the time of the measurement was 50 mW. The automatic repeat request (ARQ) was disabled in these measurements. The receiver then showed statistics for each of the different data rates, such as the number of received packets and the average received signal strength indication (RSSI). Since 500 packets was not sufficient to produce an accurate statistical results, a number of such runs (a minimum of 8) were performed. The laptop was placed on a trolley and moved around the building. Parts of the results obtained from these experiments can be seen in Figure 2.9-2.10.

Figure 2.9 shows the measurement results according to the frame error rate (FER) value, for a certain location of the AP. The results are depicted for data rate of 6 Mb/s. The measurements of FER show that a threshold has been reached, after which the FER rapidly increases. This results in a quick transition from working to not working. It has been assumed that an area is fully covered if the FER in that area is less than or equal to 5%. Figure 2.10 depicts the RSSI for the same scenario. The RSSI value is a measure of the received field strength. It is not calibrated and does not represent any well-known units. However, they can be used as measurements relative to other RSSI measurements reported by the same test equipment. In our case RSSI is a number between 0 and 60.

As expected, full coverage has been obtained in the area close to the AP, but propagation down the corridor shows some unusual fluctuations. This can be due to the reflections from neighboring buildings as well as a copper layer which is installed on a part of the corridor’s roof. Inside the rooms we obtained higher signal strength close to doors and windows. One reason for that can be the metal trunking under the windows in the offices which runs the entire length of the building. More results and details of measurements can be found in [45].

The same measurement campaign has been performed using DECT handsets, locating the DECT base station (BS) in the same place as the AP in the previous measurement. This time we had full coverage for the entire floor.

The results in Figure 2.9 and 2.10 are obtained for the lowest IEEE 802.11a data rate, i.e., 6 Mb/s. The higher link data rates in IEEE 802.11a systems are achieved by applying
higher-order modulation schemes which are not as robust as the lower-order modulations. Consequently smaller coverage areas are attained for larger data rates as reported in [45].

The coverage measurement results show that the coverage range of an IEEE 802.11a WLAN is much smaller than that of a DECT system. Often diversity and relaying methods have been considered as techniques to improve the coverage range. Within the framework of the project we investigated possible diversity methods as well as relaying protocols to enhance the coverage range of the WINDECT. As WINDECT has been designed for optimal transmission of voice packets, we studied simple diversity techniques like antenna selection which does not add additional delay. Outage probability, outage capacity and coverage angle are some of the performance criteria which have been considered in our analysis. Outage probability is defined by the probability that the instantaneous rate drops below a certain value called outage rate. More explanation about all considered parameters can be found in the next chapter.

In this section we only present one outage probability plot as an example of our coverage investigations. The outage probability versus distance for IEEE 802.11a with different degrees of frequency diversity ranging from one (single-tap channel) to infinity i.e. additive white Gaussian noise (AWGN) channel is plotted in Figure 2.11. For comparison the outage probability of the DECT system is also depicted. DECT has much lower bandwidth as compared to the IEEE 802.11a but it profits from higher transmit power and lower outage rate.

Figure 2.11: Outage probability vs. distance between source and destination [47].
As it is seen applying diversity methods can improve the coverage range of an IEEE 802.11a AP considerably\textsuperscript{5}.

Relaying is another way which can be used to improve the coverage range of an AP. However, as current IEEE 802.11 standard does not define any relaying function, except that for the AP, we need to define a relaying protocol which fulfills WINDECT requirements. In this chapter we only study the relaying for WINDECT and we postpone the general concepts of relaying to Section 3.2. More details on coverage investigation for point-to-point communication links are included in Chapter 3.

### 2.2.3 Relaying Protocol for WINDECT

One way to extend the coverage range of an AP is to install relays in a BSS, for example close to the border of the AP coverage area. Due to high QoS requirements, in WINDECT we consider only dedicated relays which are fixed. These relays only forward packets without private benefit.

In general, relay signalling can be done in both CFP and CP, but during CP there is no guarantee for the relay to access the channel at a certain time. Smart use of EDCA can increase the probability of accessing the channel for a certain user but still cannot guarantee it. Hence, for high quality voice systems like WINDECT, transmission of voice packets during CFP is recommended. We distinguish two different scenarios where AP and the relay operate at different frequencies and when they share the same frequency band.

**Relay and AP operate at different frequencies:**

In this case we assume that the relay has all functionalities of an AP but has no backbone configuration. To guarantee the QoS, voice packets are delivered during CFPs while all data packets are transferred during CPs. The relay is served by the first AP or potentially another relay. It is assumed that AP and relay communicate at frequency $f_1$ while relay and STA(s) communicate at frequency $f_2$. Hence, the relay should switch to a different frequency in order to connect to the AP. Since this tuning in WINDECT is done during the CP, the amount of time which can be used for data transmission is reduced. This method can be beneficial when the AP is heavily loaded. The packet transfer in this case is shown in Figure 2.12.

As explained in Section 2.2.1, in WINDECT for each associated STA, data is only transmitted every other beacon. This structure has been defined in this way to allow seamless

\textsuperscript{5}It should be noted that some of the simulation parameters which are used here were chosen according to the IEEE 802.11a and according to local regulations at the time of simulations in 2006, these parameters are partly different from those used within the rest of the thesis.
2.2 WINDECT

Figure 2.12: The relaying method proposed for WINDECT, relay and AP operate at different frequencies. STA$_1$ communicates directly with the AP at $f_1$, while STA$_2$ communicates with the AP via the relay at $f_2$.

handover in WINDECT. We can benefit from this structure for relay signalling. If the AP uses odd beacons for transmitting the data to the relay, then the relay has to use the even beacons for transmissions to the STAs.

As it is seen in Figure 2.12 in the first beacon frame, the AP communicates with the relay. It allocates time slots to the relay to transmit packets to the STAs which are connected via the relay. In the second beacon frame the relay, which has already tuned to another frequency ($f_2$), communicates with the STAs which are connected to it, while the AP is transmitting packets to other STAs. In this way, there is no need for the AP to be silent during the second beacon interval. It is assumed that the relay similar to the AP can poll STAs and transmit voice packets during CFP.

**Relay and AP operate at the same frequency:**

Again, similar to the previous case, we just consider transmission during the CFP. However, here the AP and the relay use the same frequency band for their transmission. Figure 2.13 shows an example of packet transfer in this setup.

Like the previous case, during the first beacon interval, AP is communicating with several STAs and the relay. During the second beacon interval AP assigns time slots to the relay and has to be silent while relay is transmitting to the STAs. This is due to the fact that both AP and relay are working at the same frequency and hence fewer degrees of freedom are available than in the previous case.

Assuming that the AP can communicate with totally 10 STAs within one beacon frame and a scenario where all STAs communicate with the AP via a relay, then at most 10 users can be served during the CFP within two beacon frames. In this case, in the first beacon interval the AP can only communicate to the relay and allocate TXOPs for these 10 users.
During the second beacon interval, the AP has to be silent while the relay is transmitting to these STAs. So data rate is reduced by a factor 1/2 due to the relaying.

With respect to data rate this scenario is less efficient compared to the first scenario. However, as both AP and relay share the same bandwidth, smaller frequency spectrum is used. Consequently, other APs/systems, which operate in the same region, benefit from higher degree of freedom in their frequency planning.

2.3 Summary

In this chapter we have briefly presented some aspects of the IEEE 802.11 standard. The focus has been on features and parameters which are used throughout this thesis. As an example for WLAN applications, the WINDECT project is presented. Supporting high quality voice transmission on WLAN has been the main goal of this project. We have briefly presented some of the challenges, such as the required modifications as compared to the existing systems, and part of the results of our coverage measurement campaign. The coverage measurements showed that there is still considerable room for improving the coverage range of IEEE 802.11a systems. In the next chapter we will consider the fundamental bounds of the existing WLAN systems and analyse their features numerically.
Chapter 3

Multimedia Traffic over WLANs

WLANs were originally designed for data communication. However, with the fast growth of wireless telecommunications and increasing demand for multimedia applications, most wireless systems are required to support multimedia traffic. Multimedia traffic includes a combination of different traffic types like text, image, audio, video, interactive content etc. Each one of these traffic types has its own requirements. Some of them, like video and data traffic, need high data rates while others like voice require lower data rates. On the other hand while usually delay is not critical for data traffic, high quality real-time applications including voice and live video streaming are very sensitive to the delay. Furthermore, some traffic types are bidirectional, e.g., voice and interactive applications while others, e.g., video broadcasting, are unidirectional.

In this chapter, we investigate both PHY and MAC layer specifications. The performance of the existing WLAN systems with respect to coverage, data rate, and throughput is analysed. In Section 3.1 we only take single hop topologies into account while in Section 3.2 the benefits of relaying in the considered network are discussed. In the next chapters, we also present some of the differences between unicast and multicast or broadcast transmissions, where the same packet is transmitted to several stations. Furthermore, we study some of the differences between point-to-point and multiaccess communications, also referred as multipoint-to-point [48].

3.1 Fundamental Bounds and QoS Limitations

To realise high-QoS multimedia over WLANs, one needs to overcome fundamental bounds and QoS limitations of the existing WLANs. These bounds limit or sometimes even prevent
supporting certain applications over WLANs. In this section, we focus on the main limitations of the existing WLANs with respect to PHY and MAC layers. At first we discuss the problems and then in the next chapters we refer to some of these problems and suggest appropriate solutions.

3.1.1 Channel Model

In wireless networks the received signal strength is impacted by different effects including path loss, shadowing, interference, doppler, noise, and multipath. There are different path loss models proposed for different environments. In this thesis we consider a simplified path-loss model defined by the following equation [14, page 46]:

\[ P_r = P_t K_0 \left( \frac{d_0}{d} \right)^\gamma \]  \hspace{1cm} (3.1)

where \( P_r \) and \( P_t \) are the received and the transmitted power, respectively. \( d \) is the distance between the transmit and the receive antennas, \( d_0 \) is the reference distance for the antenna far field, \( \gamma \) is the path loss exponent, and \( K_0 \) is a unitless constant which for omnidirectional antenna can be calculated from the following equation:

\[ K_0 dB = 20 \log_{10} \left( \frac{\lambda}{4\pi d_0} \right), \]  \hspace{1cm} (3.2)

with \( \lambda \) being the wavelength. A path loss exponent of 2 is assumed for free-space propagation environment. In this case the above path loss equation is the same as Friis formula [49]. Different values of path loss exponent between 1.6 to 6.5 have been reported for different environments [14, page 47]. In addition to the path loss, the wireless channel shows a fading behaviour due to multipath components and a doppler shift due to movement of any object in the area. As we focus on indoor environments, where all objects are either static or move with the pedestrian speed, the doppler effect is neglected. In this thesis we consider the non-line-of-sight (NLoS) condition where the direct link between the transmitter and receiver is obstructed. A widely accepted statistical model for the wireless NLoS channel is the Rayleigh model. According to this model each complex channel coefficient consists of a real and an imaginary random variable, i.e., \( h_x \) and \( h_y \) with mean zero and equal variance of \( \sigma^2 \). In this case the amplitude of the channel tap \( A = \sqrt{h_x^2 + h_y^2} \) is Rayleigh distributed. The model is based on the fact that a large number of statistically independent reflected or scattered paths with random amplitudes contribute to a single tap.

Throughout the thesis, the ETSI non-line-of-sight indoor channel model A with root-
3.1 Fundamental Bounds and QoS Limitations

Table 3.1: ETSI Channel Model A, corresponding to a typical office environment for NLoS conditions and a rms delay spread of 50 ns in the 5 GHz band.

<table>
<thead>
<tr>
<th>Tap Number</th>
<th>Delay (ns)</th>
<th>Average Relative Power (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>-0.9</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>-1.7</td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td>-2.6</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>-3.5</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>-4.3</td>
</tr>
<tr>
<td>7</td>
<td>60</td>
<td>-5.2</td>
</tr>
<tr>
<td>8</td>
<td>70</td>
<td>-6.1</td>
</tr>
<tr>
<td>9</td>
<td>80</td>
<td>-6.9</td>
</tr>
<tr>
<td>10</td>
<td>90</td>
<td>-7.8</td>
</tr>
<tr>
<td>11</td>
<td>110</td>
<td>-4.7</td>
</tr>
<tr>
<td>12</td>
<td>140</td>
<td>-7.3</td>
</tr>
<tr>
<td>13</td>
<td>170</td>
<td>-9.9</td>
</tr>
<tr>
<td>14</td>
<td>200</td>
<td>-12.5</td>
</tr>
<tr>
<td>15</td>
<td>240</td>
<td>-13.7</td>
</tr>
<tr>
<td>16</td>
<td>290</td>
<td>-18.0</td>
</tr>
<tr>
<td>17</td>
<td>340</td>
<td>-22.4</td>
</tr>
<tr>
<td>18</td>
<td>390</td>
<td>-26.7</td>
</tr>
</tbody>
</table>

mean-square delay spread of 50ns is used. This model and other ETSI channel model for different environments are given in [50]. The channel model is frequency selective and consists of 18 taps. The channel taps are given in Table 3.1. A time resolution of 10 ns (a bandwidth of 100 MHz) is assumed. All elements of the channel coefficients in the time-domain are zero-mean complex normal random variables with unit variance. They are scaled by the taps and normalised such that on average a power of one is preserved.

In this thesis we consider only thermal noise and implementation loss. The thermal noise is modeled by a white Gaussian distribution with zero mean and power spectral density of $N_0/2$. The total noise power within $2BW$ bandwidth can be calculated by:

$$ N_0/2 \cdot 2BW = N_0 \cdot BW = k \cdot T_0 \cdot NF \cdot BW, $$

(3.3)

where $k$ is the Boltzmann’s constant ($1.3806503 \times 10^{-23}$ Joule/°K), $T_0$ is the receiver temperature in kelvin and NF is the noise factor. Throughout the thesis, an additional 5 dB implementation loss is considered which is simply included in the noise variance. Interference from other BSSs or other non-WLAN devices is neglected.
Table 3.2: PHY layer simulation parameters.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmit Power</td>
<td>2.5mW per subcarrier(^1)</td>
</tr>
<tr>
<td>Path Loss Exponent</td>
<td>3.5</td>
</tr>
<tr>
<td>Reference Distance (d_0)</td>
<td>1 m</td>
</tr>
<tr>
<td>Temperature</td>
<td>25°C</td>
</tr>
<tr>
<td>Noise Figure</td>
<td>10dB</td>
</tr>
<tr>
<td>Implementation Margin</td>
<td>5dB</td>
</tr>
<tr>
<td>Carrier Frequency</td>
<td>5.2GHz</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>20MHz</td>
</tr>
<tr>
<td>Subcarrier frequency spacing</td>
<td>20 MHz/64=312.5 KHz</td>
</tr>
<tr>
<td>Number of data subcarriers</td>
<td>52</td>
</tr>
<tr>
<td>Number of total subcarriers</td>
<td>64</td>
</tr>
</tbody>
</table>

3.1.2 PHY Layer Data Rate

With the fast development of wireless communications and the change of application from pure data to multimedia, there is a considerable demand for high data rate services.

Conventional wireless systems support a maximum rate of 54 Mbit/s. Recently the WLAN standard for MIMO systems, i.e., IEEE 802.11n, has been finalised \(^1\). The IEEE 802.11n PHY is based on MIMO OFDM in the 2.4GHz and 5GHz band, operating in the 20 or 40 MHz bandwidth. The highest link data rate in IEEE 802.11n system is 600 Mb/s which is achieved by increasing the number of spatial streams to four, applying 64-QAM with a coding rate of 5/6, a shorter guard interval, i.e., 400 ns, and 40 MHz bandwidth. Since the guard interval should be longer than the longest path delay, using a short guard interval may not be possible in all environments. In this thesis, we consider a guard interval duration of 0.8 \(\mu\)s and a bandwidth of 20 MHz.

The ergodic capacity of a MIMO system without channel state information at the trans-\(\text{mitter (CSIT) over a frequency selective channel with } N \text{ frequency flat sub-channels can be calculated by the following formula} \(^5\):\(^2\):

\[
C = \mathbb{E}\left[\frac{1}{N} \sum_{i=1}^{N} \log_2 \det \left[ I_{M_t} + \frac{\rho}{M_t} H_i H_i^H \right] \right] \text{[bps/Hz]},
\]

\(^1\)According to the Swiss Federal Office of Communications in the considered band a maximum radiated power of 200 mW EIRP is allowed and the maximum power mean EIRP density shall be limited to 0.25 mW/25 KHz in any 25 KHz bandwidth \(^5\). Throughout this work we simply assume that the maximum transmit power is uniformly split over all sub-channels. Taking into account the power loss due to the guard interval, this leads to a transmit power of 160/64 mW per subcarrier.
where \( \rho = P/\sigma^2_{n_i} \) with \( P \) being the transmit power per sub-channel and \( \sigma^2_{n_i} \) the noise variance per sub-channel, \( H_i \) is the \( i \)th sub-channel gain matrix and the number of antennas at the transmitter and the receiver is respectively \( M_t \) and \( M_r \).

Throughout this thesis, we consider an OFDM system with a total bandwidth of 20 MHz [1]. In multiple-antenna setups, it is assumed that the signals transmitted from different antennas are uncorrelated. Transmit power is uniformly distributed over all subcarriers. Furthermore, there is a signal to noise ratio (SNR) loss at the receiver due to the power which is allocated to the cyclic prefix. The cyclic prefix is utilised as a guard interval and it leads to SNR and spectral efficiency loss [53,54]. The SNR loss at the receiver due to the guard interval is determined by [54, page 38]:

\[
\text{SNR loss} = -10\log_{10} \left( 1 - \frac{T_{GI}}{T_{\text{sym}}} \right),
\]

in which \( T_{GI} \) is the duration of the guard interval and \( T_{\text{sym}} \) is the duration of one OFDM symbol. Besides, due to the guard interval the spectral efficiency is also reduced by a factor of \((1 - \frac{T_{GI}}{T_{\text{sym}}})\). For a system based on the IEEE 802.11a and 802.11n [37], [1] with an OFDM symbol duration of 4 \(\mu\)s and a guard interval of 0.8 \(\mu\)s the loss is 20\%. This leads to a SNR loss of approximately 1 dB. The simulation parameters, which are used throughout this thesis, are shown in Table 3.2. Taking parameters of IEEE 802.11n into account for a GI of 0.8 \(\mu\)s the spectral efficiency can be obtained from:

\[
R = 0.8E \left[ \frac{1}{64} \sum_{i=1}^{52} \log_2 \det \left( I_{M_r} + \frac{P}{M_t \sigma^2_n} H_i H_i^H \right) \right] \text{[bps/Hz]},
\]

where an optimal codebook is assumed. The pre-factor 0.8 is due to the fact that 20\% of each symbol is allocated to the GI. As it is mentioned in Section 3.1.3, there is also a power loss due to the GI which has been reduced from the total transmit power. The transmit power \( P \) indicates the final transmit power per subcarrier, cf. Table 3.2. This equation provides an upper bound for the spectral efficiency. Although for practical modulations the achievable rate is below the values calculated from (3.6), in this work this upper bound is used for comparison-based results. In this way, our PHY analysis is kept as simple as possible compared to other PHY analyses, like estimation of the packet error ratio, which requires a complete description of the whole PHY layer. In our analysis we also take a noise figure and an implementation margin into account, cf. Table 3.2. This leaves an additional margin for some degradation of the received SNR caused by different components for example due to the implementation errors. As a result, applying the considered parameters, fair comparisons
of PHY layer performance of different systems can be made.

The data rate for SISO, 1x4 single-input multiple-output (SIMO), 4x1 multiple-input single-output (MISO) and 4x4 MIMO are depicted in Figure 3.1. As it is seen the MIMO system shows a considerable spatial multiplexing gain. However to realise this gain some requirements should be fulfilled. For example to utilise the spatial multiplexing gain a rich scattering environment is needed. Besides, correlation among antennas degrade the capacity performance\(^2\). In many networks complexity of the STAs should be as small as possible. In such networks, STAs may have only a single transceiver, therefore with conventional point-to-point links we cannot gain from spatial multiplexing gain anymore. According to the IEEE 802.11n the support of only one spatial stream is mandatory at STAs. Furthermore, the IEEE 802.11n supports only point-to-point links, hence as it is explained in the next chapters, in order to realise cooperative networks which require simultaneous multiuser transmissions one needs to modify the existing protocols.

\(^2\)We assume that the spaces between the antennas are large enough and there are enough scatters such that the spatial correlation can be neglected.
3.1.3 Coverage

It is expected that future wireless communication systems will operate even beyond 5 GHz. Moving towards higher frequencies leads to larger path loss and as a result to smaller coverage area. Besides, for real-time applications\(^3\) full coverage would be necessary, as for example in cellular telephony networks. However, providing full coverage is problematic mainly due to the random nature of the wireless channels and shadowing. Furthermore, the real-time applications require seamless handover which implies having enough overlaps between ranges of neighbouring access points.

In Chapter 2, we presented the results of our coverage measurements in a typical office environment using off-the-shelf IEEE 802.11a devices. It was shown that the coverage area served by a WLAN AP operating in the 5 GHz band in NLoS condition at the lowest possible data rate is in general small. The range was even much smaller when the WLAN AP was operating at higher data rates. For example at 54 Mbit/s we could achieve a coverage range of only few meters.

There are several ways to increase the coverage area of a WLAN system:

- Increasing the number of APs in a certain region: as the APs are connected to the backbone, by increasing the number of APs infrastructure costs increase and system flexibility is reduced. In addition to this, due to small number of non-overlapping channels, frequency planning problems arise when the number of APs in a given area is increased.

- Increasing the transmit power: the transmit power is usually limited by the local regulations and standards.

- Applying antenna diversity techniques: these techniques increase the diversity gain and as a result enhance the coverage range. Diversity gain increases the slope of the symbol error ratio (or outage probability) curves versus SNR while array gain increases the average received SNR \[14\] \[52\], page 205] [52, page 88].

- Relaying: in relaying networks the relay node assists the communication between the source and the destination node. This results in either a larger coverage area for the same data rate or a higher data rate within the same coverage area.

In this section we analyse the coverage range of a WLAN with single-hop communication links. The point-to-point relaying methods are discussed in Section 3.2.

\(^3\)For data traffic dead spots within the coverage area are acceptable since the damaged or lost data packets can be retransmitted.
### Outage Performance

Throughout this work, we consider fading channels. As a result the instantaneous rate, which is supported by the channel, is a random variable. We use the instantaneous channel capacity as upper bound on the instantaneous achievable rate. By definition an outage occurs, when the instantaneous rate $R$ drops below the target outage rate $R_{\text{out}}$. Thus, the outage probability of a link is defined as [15, page 187]:

$$P_{\text{out}} = \Pr(R < R_{\text{out}}), \quad (3.7)$$

We assume an application which requires at least a data rate of $R_{\text{out}}$ on the channel. Therefore we cannot avoid outage events even by knowing CSIT and link adaptation possibility. We employ outage probability as a parameter which indicates the coverage range. Besides, the slope of the outage probability curves in high SNR-regime shows the diversity order of the respective system.

We define coverage range as the largest radius of a circular area where outage probability is below a desired value, e.g. 1%. To determine coverage range, we calculate the outage probability for a given outage rate via Monte-Carlo simulations and plot it versus the distance between source and destination. We fix outage probability to the desired value. The respective value on the X-axis indicates the maximum possible range between the source and its destination. Even though, in practice the coverage range highly depends on channel conditions and characteristics of the indoor environment, the outage probability curves can be used as the basis for comparisons.

First we consider a SISO system and assume that an uncoded data rate of $R_{\text{out}} = 13, 26, 52, \text{ and } 78 \text{ Mb/s}$ is respectively required. These rates are the possible uncoded data rates per link which can be supported by an IEEE 802.11n device [1]. These rates are achieved by varying the modulation scheme from BPSK to QPSK, 16/64-QAM (cf. Table 2.1). Since we use capacity values in (3.7), the calculated outage probabilities are obtained even in the presence of the most sophisticated code. In practice, for a given SNR, we have higher outage probabilities than the ones obtained from (3.7). However we consider the theoretical value as a simple parameter which shows behaviour of a system and we apply it for coverage investigation and comparisons of different PHY layer schemes.

In a similar way, one can achieve outage probability according to a required SNR, i.e. $\text{SNR}_{\text{thr}}$. In this way, outage probability can be defined by the $\Pr(\text{SNR}_r < \text{SNR}_{\text{thr}})$ where $\text{SNR}_r$ is the received SNR. The IEEE 802.11n standard specifies the minimum sensitivity to achieve for a PSDU of 4096 octets, a packet error rate less than 10% for different modulations. The minimum input sensitivity is measured as the average power per receive
3.1 Fundamental Bounds and QoS Limitations

![Figure 3.2: Outage probability vs. distance between transmitter and receiver, simulation parameters are shown in Table 3.2.](image)

antenna [1]. However, the SNR values are not averaged over fading. Therefore, some additional loss should be considered in the calculated SNR\text{thr} values depending on the channel model. Throughout this thesis we consider the outage probabilities defined according to the first method, i.e., rate-dependent method, for our comparison-based analysis.

The outage probability of the SISO system versus distance between the transmitter and the receiver is plotted in Figure 3.2. As shown in this figure, for a fixed $P_{\text{out}}$ the largest coverage range is achieved for the lowest outage rate. Similar behaviour has been also observed in our measurement results presented in Section 2.2.2. However, some of the parameters which are used here are different from those in the measurements.

As shown in Figure 3.2, the NLoS coverage range of the SISO system is relatively small. For example at the highest possible uncoded data rate (78 Mb/s) the coverage range is limited to about 8 m for outage probability value of 1%.

In order to enhance data rate without sacrificing coverage range we need to take into account other techniques than changing the modulation alphabet. While spatial multiplexing enhances the data rate, diversity techniques mainly improve robustness and outage behaviour of a system. There are several possible approaches to attain diversity gains in time, frequency, and/or space domain [14]. However, time diversity methods like automatic repeat request introduce additional delay and hence usually cannot be applied for real-time traffic.
In an OFDM system we can gain from frequency diversity by for example coding across subcarriers. In order to benefit from space diversity we take multiple-antenna nodes into account.

Among different cases in Figure 3.2, we consider the case with the smallest coverage range, i.e., the scenario with \( R_{\text{out}} = 78 \) Mb/s and investigate how much the coverage range improves using multiple antenna at the transmit or receive side. We simulate the outage probabilities for a \( 1 \times 4 \) SIMO, a \( 4 \times 1 \) MISO, and a \( 4 \times 4 \) MIMO system. The outage probability of the SISO system is also plotted as the reference. Figure 3.3 shows the results. It should be noted that the difference between SIMO and MISO curves is due to the lack of CSI at the transmitter. The \( 4 \times 4 \) MIMO significantly improves the outage performance, e.g., at \( P_{\text{out}} = 1\% \) it supports almost 4.2 times larger distance between a transmitter and its receiver compared to the SISO case. However, since many nodes in a system may still be equipped with only a single antenna, we may achieve in many practical scenarios at most the SIMO or MISO gains (depending on whether CSI is available at the transmitter or not).

Since we are interested in indoor WLAN systems, supporting constant data rates over slow fading channels, throughout this thesis outage probability is considered as the main criteria for the PHY layer performance.
3.1.4 MAC Layer Throughput Analysis

In practice the achievable throughput in the MAC layer is much less than the PHY data rates discussed in the previous subsection. This is mainly due to the redundancies and overheads attached to the sequence of data bits, the time wasted during erasure transmissions and the waiting time due to the backoff procedure. The latter is needed to avoid collisions on a shared medium. For an IEEE 802.11a network with a single source, which transmits data packets of 1000 bytes over an ideal channel without error and with data rate of 54 Mb/s, a maximum achievable throughput of around 46% has been reported in [55]. Throughput of a single-hop link based on DCF has been widely studied in several papers such as [6], [8] and [56].

Throughout the thesis, we use Bianchi’s model [6] to calculate collision and transmission probabilities. In [6] an ideal channel with no channel errors, hidden nodes and capture is assumed. All nodes operate in saturation condition. Furthermore, no retransmission limit is considered. This model has further been extended to include packet retransmission limit and transmissions error [35] [57], non-saturated condition and capture effect [8], backoff freezing [58] [59] and presence of anomalous slots, i.e. the slots in which probability of transmission and collision are different from the other ones [60] [61].

In the following paragraphs, first the original model is explained and then different extensions are discussed.

3.1.4.1 Markov Chain Model

**Original Model:** Bianchi has modeled the backoff procedure with a two-dimensional Markov chain where backoff time counter, \( b(t) \), and backoff stage, \( s(t) \), are represented by two stochastic processes. The key approximation in this model is that each packet at each transmission collides with a constant and independent probability, \( p = P_{\text{col}} \), regardless of the actual backoff stage. This assumption is also known as the decoupling assumption [62] [63]. Furthermore, it is assumed that backoff counter decremented at the beginning of time slot regardless of the channel status\(^4\). Besides in this model an infinite number of retransmissions has been assumed.

The respective Markov chain, shown in Figure 3.4, models the bidimensional process \( \{ s(t), b(t) \} \). Each stage \( b_{i,k} \) represents the stationary distribution of the chain where \( b_{i,k} = \lim_{t \to \infty} \Pr\{ s(t) = i, b(t) = k \}, i \in (0, m), k \in (0, W_i - 1) \), where \( W_i = 2^i W \) and \( i \in (0, m) \) is the backoff stage. We can calculate \( W \) and \( m \) from the minimum and

\(^4\)According to the IEEE 802.11 standard each STA decrements its backoff only if the channel is sensed idle.
Figure 3.4: Markov chain model for the backoff procedure.
maximum contention window sizes, denoted by \( CW_{\text{min}} \) and \( CW_{\text{max}} \) respectively, as follows: 
\[
W = CW_{\text{min}} + 1 \quad \text{and} \quad CW_{\text{max}} = 2^m W - 1.
\]
Both \( CW_{\text{min}} \) and \( CW_{\text{max}} \) are defined by the standard [1].

The non-zero transition probabilities in this Markov Chain are\(^5\) [6]:

\[
\begin{align*}
\Pr\{i, k|i, k + 1\} &= 1, \quad k \in (0, W_i - 2) \quad i \in (0, m) \\
\Pr\{0, k|i, 0\} &= (1 - p)/W_0, \quad k \in (0, W_0 - 1) \quad i \in (0, m) \\
\Pr\{i, k|i - 1, 0\} &= p/W_i, \quad k \in (0, W_i - 1) \quad i \in (1, m) \\
\Pr\{m, k|m, 0\} &= p/W_m, \quad k \in (0, W_m - 1).
\end{align*}
\]

(3.8)

We obtain the first equation by assuming that the backoff time is decremented at the beginning of each time slot with probability one. The second equation models the fact that after a successful transmission the backoff stage resets to 0 and the backoff is uniformly chosen in the range \((0, W_0 - 1)\). The third equation shows that after an unsuccessful transmission, the backoff stage increases and the new backoff value is uniformly chosen in the range \((0, W_i)\). However, as the fourth equation shows once the backoff stage reaches its maximum value, here denoted by \( m \), then it is not increased anymore in the next transmissions.

Using the transition probabilities and taking into account the normalisation condition, one can easily obtain a closed-form solution [6]. In this way, assuming a fixed number of contending nodes \( n \), the probability that a node transmits in a randomly chosen time slot, \( \tau \), can be obtained from the following equation [6]:

\[
\tau = \frac{2(1 - 2p)}{(1 - 2p)(W + 1) + pW(1 - (2p)^m)}. \tag{3.9}
\]

Under ideal channel conditions, the transitional probability from one backoff stage to the next one, \( p \), is the same as the conditional collision probability, i.e., \( p = P_{\text{col}} \), where \( P_{\text{col}} \) is the probability of collision for a packet which is being transmitted on the channel. As calculated in [6], each transmitted packet collides if at least one of the other nodes transmits in the same time slot:

\[
P_{\text{col}} = 1 - (1 - \tau)^{n-1}. \tag{3.10}
\]

By solving the equations (3.9)-(3.10) transmission and collision probabilities are obtained.

\(^5\)It should be noted that we have used the same notation as in [6] in which:
\[
\Pr\{i_1, k_1|i_0, k_0\} = \Pr\{s(t + 1) = i_1, b(t + 1) = k_1|s(t) = i_0, b(t) = k_0\}
\]
**Channel Error:** In the presence of channel errors, the IEEE 802.11 MAC cannot distinguish them from the collision events. Therefore, the CW increases after each unsuccessful transmission to prevent further collisions. Accordingly in (3.9), $p$ should be replaced by \[ p \] = 1 - (1 - P_{\text{col}})(1 - P_e) = P_e + P_{\text{col}} - P_eP_{\text{col}}, \] where $P_e$ is the packet error probability. To obtain this expression, it is assumed that all STAs on average have the same channel error probability and that the channel errors and collision events are independent.

**Limited Number of Retransmissions:** According to the IEEE 802.11 standard, a data packet, which does not use RTS/CTS handshake, can be transmitted up to a $\text{ShortRetryMax}$ equals 7 times. Accordingly the original model has been extended such that it includes the packet retransmission limit [35] [57] [64]. Based on this extended model a packet can be retransmitted only $m'$ times. After the $m'$th attempt, if it has not been received correctly, it will be dropped and the contention window will be reset to its initial value. However, as it has been shown in [57] for a retry limit larger than six, the throughput does not improve significantly with increase in the retry limit.

**Backoff Freezing Probability:** The backoff model in [6] assumes that the backoff counter is decremented at the beginning of a time slot regardless of the fact whether that slot is idle or busy. However according to the standard [5] the backoff counter is decremented at the end of a time slot and if the channel is sensed busy during a time slot the backoff procedure is suspended.

In [58] and [59] the authors have taken the freezing probability into account. In [58] the backoff decrement probability, i.e. in the considered Markov Chain (MC) the transition probability from stage $b_{i,k+1}$ to stage $b_{i,k}$, is set to the probability that no other STA transmits in that slot. For a given number of contenting nodes, this leads to smaller transmission probability than that in the original model. In fact, as it is discussed in [60] [65] and [66], these approaches suffer from the assumption of statistical independence of consecutive time slots. Simulations in [60] have shown that in this approach throughput values are less accurate than those obtained by the original model.

**Anomalous Slots:** A more accurate model for the IEEE 802.11 backoff procedure has been developed in [60] and [66]. As explained in [60], in saturation mode the time slot following a successful transmission can only be accessed by the station which has successfully transmitted in the previous channel access. Hence in these time slots (denoted by anomalous
slots) the channel access probability is much lower than in other slots. On the other hand, the time slot following a collision cannot be accessed by any node.

In [66], the authors consider the backoff freezing probability but correct the model in [59] by taking the dependency of channel access probability and the collision probability on the channel status. Their new model is a tri-dimensional MC, which compared to the original model, includes a new state, indicating the status of the previous state (busy or idle).

In [60], the authors consider the presence of anomalous slots while retaining the basic assumption about statistical independence of consecutive slots as in the original model. Similar to [66] they take the anomalous slots (the slots immediately after a successful transmission) into account. Additionally, they consider the fact that the time slot immediately following a collision cannot be accessed by any station. Their model leads again to a bi-dimensional MC. However, they show that using this model same results can be achieved as the ones applying the alternative tri-dimensional MC. They have validated their model by performing simulations and shown that the model is very accurate.

Summary of the MC Model Discussion

In this thesis, we apply the original backoff model introduced in [6] which is simple but accurate and has been widely used in the research community.

The model has been validated in [6] for IEEE 802.11b parameters. In [60] the authors have considered the IEEE 802.11a parameters and showed that the results using the original model are very close to the results obtained from a more accurate model introduced in [60]. Furthermore, as it has been shown in [63] for saturated networks Bianchi’s decoupling assumption is valid even for small networks.

We focus on the saturation throughput. The saturation throughput shows the limit of the system throughput as the offered load increases. We consider the extension of the original model which includes channel error probability. Additionally, we take the idle time slot following reception of an error frame into account and assume an infinite retransmission limit. We have plotted the throughput taking into consideration the parameters which are used in this thesis (cf. Table 3.3) for the case with no retransmission limit and the retransmission limit of 7. As shown in Figure 3.5, the results are close especially for a low number of STAs. For example at data rate of 58.5 Mb/s and for 5 STAs there is only about 0.2% throughput.

However, it has been shown that in the cases where the frame duration is comparable to the backoff slot size the original model leads to less accurate results. In the considered IEEE 802.11n MIMO setup in [60], where there are five stations and a packet of 1500 byte is transmitted at 600 Mb/s, the original model leads to 3.5% higher throughput compared to the simulation results. It should be noted that in our study much lower rates (at most 58.5 Mb/s) are used.
Figure 3.5: Aggregate throughput vs. number of STAs with and without retransmission limit for two different data rates.

difference between these two models. As it will be explained in the next chapters, our proposed MAC protocol reduces the number of contending units in the network and hence the regime with low number of contending units is more interesting for us.

3.1.4.2 Throughput

Throughput is defined by the average payload bits which are transmitted successfully in a time slot divided by the duration of the time slot. After calculating collision and transmission probabilities (as explained in the previous section), we need to obtain the following parameters to analyse throughput.

The probability that there is at least one node that transmits in a slot, $P_{tr}$, and the probability that only one node transmits in a slot or in other words a collision-free transmission occurs in a time slot, $P_s$, are expressed as [6]:

\[
P_{tr} = 1 - (1 - \tau)^n, \quad (3.12)
\]
\[
P_s = n\tau(1 - \tau)^{n-1}. \quad (3.13)
\]

Assuming the same average channel conditions, data rates, and average payload size for all
3.1 Fundamental Bounds and QoS Limitations

STAs, the throughput can be defined as [8]:

\[ S = \frac{P_s(1 - P_e)L}{T_{\text{slot}}} = \frac{n\tau(1 - \tau)^{n-1}(1 - P_e)L}{T_{\text{slot}}}, \] (3.14)

\[ T_{\text{slot}} = (1 - P_{tr})\sigma + P_s(1 - P_e)T_s + (P_{tr} - P_s)T_c + P_sP_eT_e; \] (3.15)

where \( L \) is the payload size, \( \sigma \) is the duration of an empty time slot, \( T_s, T_c, \) and \( T_e \) are the average time the channel is sensed busy due to successful transmission, collision, and channel error, respectively. \( T_s, T_c, \) and \( T_e \) are defined as follows:

\[ T_s = T_{\text{Data}} + \delta + \text{SIFS} + T_{\text{ACK}} + \delta + \text{DIFS}, \] (3.16)

\[ T_c = T_{\text{Data}} + \delta + T_{\text{ACKtimeout}} + \text{DIFS}, \] (3.17)

\[ T_e = T_{\text{Data}} + \delta + T_{\text{ACKtimeout}} + \text{DIFS}, \] (3.18)

where SIFS and DIFS are the short interframe space (IFS) and the DCF IFS, respectively [5]. \( \delta \) is the propagation delay and \( T_{\text{ACKtimeout}} = \text{SIFS} + T_{\text{ACK}} + \delta + \sigma \) is the ACK timeout. In the case of ideal channel conditions or operation at high SNR, \( P_e = 0 \) and hence the throughput expression in (3.14) can be simplified. The duration of the data and the ACK frames in an OFDM-based WLAN can be obtained as follows [67]:

\[ T_{\text{Data}} = T_{\text{PLCP}} + T_{\text{PLCPSIG}} + \left\lceil \frac{40 + (16 + 6)/8 + L}{BpS(M)} \right\rceil \cdot T_{\text{sym}}, \] (3.19)

\[ T_{\text{ACK}} = T_{\text{PLCP}} + T_{\text{PLCPSIG}} + \left\lceil \frac{14 + (16 + 6)/8}{BpS(M')} \right\rceil \cdot T_{\text{sym}}, \] (3.20)

where \( T_{\text{PLCP}} \), \( T_{\text{PLCPSIG}} \), and \( T_{\text{sym}} \) are the durations of the PHY layer convergence protocol preamble (PLCP), PLCP SIGNAL and one OFDM symbol, respectively. The number of bytes per OFDM symbol for a given modulation \( M \) and the payload size are denoted by \( BpS(M) \) and \( L \), respectively. The \( BpS \) values for each IEEE 802.11n MCS can be obtained from Table 2.1 by dividing the number of data bits per symbol, i.e. \( N_{\text{DBPS}} \), by 8.

3.1.5 Throughput Results

Throughout this thesis we use the IEEE 802.11n values, i.e. \( W = 16 \) and \( m = 6 \), cf. Table 3.3. In Figure 3.6 the throughput of a WLAN system is plotted for different number of
Table 3.3: The analysis parameters which are used throughout the thesis.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIFS</td>
<td>16 (\mu s)</td>
<td>SIFS time</td>
</tr>
<tr>
<td>DIFS</td>
<td>34 (\mu s)</td>
<td>DIFS time</td>
</tr>
<tr>
<td>(\delta)</td>
<td>1 (\mu s)</td>
<td>Propagation delay</td>
</tr>
<tr>
<td>(CW_{\text{min}})</td>
<td>15</td>
<td>Minimum CW</td>
</tr>
<tr>
<td>(CW_{\text{max}})</td>
<td>1023</td>
<td>Maximum CW</td>
</tr>
<tr>
<td>(\sigma)</td>
<td>9 (\mu s)</td>
<td>Slot time</td>
</tr>
<tr>
<td>(L)</td>
<td>1024 Byte</td>
<td>Payload size</td>
</tr>
<tr>
<td>Data MAC(H)+FCS</td>
<td>40 Byte</td>
<td>Data MAC Header and FCS field length</td>
</tr>
<tr>
<td>ACK, CTS MAC(H)+FCS</td>
<td>14 Byte</td>
<td>ACK and CTS MAC Header and FCS field length</td>
</tr>
<tr>
<td>RTS MAC(H)+FCS</td>
<td>20 Byte</td>
<td>RTS MAC Header and FCS field length</td>
</tr>
<tr>
<td>(T_{\text{PLCP}} + T_{\text{PLCP}})</td>
<td>36 (\mu s)</td>
<td>PLCP Preamble and header duration</td>
</tr>
</tbody>
</table>

Figure 3.6: Throughput vs. number of STAs for basic and RTS/CTS access mechanisms and a data rate of 6.5 and 58.5 Mb/s.
3.1 Fundamental Bounds and QoS Limitations

Figure 3.7: Comparison of throughput results between the results obtained from the simulation and those achieved using the model.

nodes. The parameters which are used for the analysis are shown in Table 3.3. It is assumed that the control frames\(^7\) are transmitted at the lowest basic rate, i.e., 6.5 Mb/s. As it is seen in this figure, a WLAN system operating at the lowest data rate has a greater normalised throughput\(^8\) than one with a high data rate. This is due to the fact that the PHY headers and ACK frames are transmitted with 6.5 Mb/s regardless of the payload rate. The results show that the throughput decreases with increase in the number of STAs. This happens since collision occurs more often as the number of nodes in the network increases.

Simulation Results: In order to validate the analytical results we have simulated the same scenario in MATLAB. We have considered a network with \(n\) STAs and simulated the random backoff procedure. A perfect channel has been assumed (no channel error and no path loss). Figure 3.7 shows the throughput values obtained using the analytical model and the ones from the simulation when number of STAs varies from 2 to 8. As it can be observed the model provides good estimation for throughput results. However, as it has been expected, the model becomes more accurate as number of STAs in a network grows. Besides, it is more accurate for scenarios with lower data rates.

RTS/CTS Access Method: We have included the RTS/CTS throughput in Figure 3.6.

\(^7\)Control frames include signalling messages like ACK, RTS and CTS frames.

\(^8\)Normalised throughput is defined by the time used for data transmission divided by \(T_{\text{slot}}\), cf. (3.15).
The average time the channel is sensed busy due to successful transmission and collision in the case of RTS/CTS can be respectively calculated from:

\[
T_s = \text{RTS} + \delta + \text{SIFS} + \text{CTS} + \delta + \text{SIFS} + T_{\text{Data}} \\
+ \delta + \text{SIFS} + T_{\text{ACK}} + \delta + \text{DIFS},
\]

\[
T_c = \text{RTS} + \delta + T_{\text{CTStimeout}} + \text{DIFS},
\]

where the durations of the RTS and the CTS frames can be obtained as in (3.20) by taking into account the respective MAC headers. \(T_{\text{CTStimeout}}\) is the CTS timeout which is equal to \(\text{CTS} + \delta + \text{SIFS} + \sigma\). Using RTS/CTS access method reduces the time channel sensed busy due to collision but introduces more overhead. The RTS/CTS throughput performance strictly depends on the data rate. The RTS/CTS access mechanism outperforms the basic access mechanism in the low data rate case. However, the basic access method outperforms the RTS/CTS in cases where packets are transmitted with a high data rate. Similar results have been reported in [68]. However, in our analysis the packets have longer PHY headers compared to that in [68]. Furthermore, here the data rate is slightly higher and there is no retransmission limit. In our analysis the RTS/CTS frames, similar to ACK frames, are transmitted at the lowest basic rate.

Despite lower throughput of the RTS/CTS scheme, for large data packets and to cope with hidden nodes it is still beneficial to use it. In this thesis, we mainly consider the basic access method, however the extension to RTS/CTS method is straightforward.

**Multi-rate Network:** In the above analysis it is assumed that all STAs transmit with the same PHY rate. The throughput of a WLAN system with asymmetric data rates has also been studied. It has been shown that in that case the STA with slowest rate may significantly limit the throughput of STAs with higher rates [69, 70]. This phenomenon which is known as performance anomaly happens when for example all STAs can transmit the same amount of data with the same channel access probability but some of them are further away from the AP and support lower data rates. Hence, those STAs with lower rates occupy the channel for a longer time as compared to the STAs with higher data rates. Consequently, the throughput of STAs does not increase linearly proportional to their PHY data rates anymore. The IEEE 802.11e can cope with this problem by limiting the maximum TXOP and/or assigning users different priority categories.

**Infrastructure Network:** The throughput of uplink and downlink in an infrastructure
3.1 Fundamental Bounds and QoS Limitations

Figure 3.8: Aggregate throughput vs. number of STAs with a data rate of 58.5 Mb/s.

WLAN based on DCF is not symmetric and is limited by the downlink rate. This is due to the fact that the AP has the same channel access probability as that of the other STAs. However, assuming $n - 1$ STAs are communicating via the AP, during the uplink all $n - 1$ STAs transmit to the AP and in the downlink the AP forwards the received packets to all STAs. The uplink and downlink throughput are as follows [71]:

$$U_{\text{up}} = \frac{n - 1}{n} S,$$  \hspace{1cm} (3.23)

$$U_{\text{down}} = \frac{1}{n} S,$$  \hspace{1cm} (3.24)

where $S$ can be calculated from (3.14).

Figure 3.8 shows how dramatically aggregate throughput degrades with increase in number of STAs when all of them transmit over a two-hop link through the AP. This problem can be solved for example by applying EDCA and estimating backoff parameters for STAs in a way that weighted fairness can be realised [72] or similar to [73] by letting the AP forward the downlink packets after a SIFS right after the uplink reception. The throughput of the latter case, which is here called "enhanced infrastructure" is also depicted in Figure 3.8. In this way uplink and downlink utilisations become equal. Besides, in the absence of hidden nodes, collision may occur only during the first hop.

Impact of PHY Link Errors: So far MAC layer throughput of a system with ideal PHY
Figure 3.9: Throughput per STA vs. number of STAs in symmetric and asymmetric scenarios with different values of PER. Data rate is 58.5 Mb/s and ACK rate is 6.5 Mb/s.

layer has been considered. However, throughput significantly drops with increase in channel errors.

One drawback of the IEEE 802.11 is that it is not able to distinguish between unsuccessful transmissions due to link errors or due to collisions, and, in both cases, the backoff interval increases exponentially up to a certain value [5]. As explained in [74] and [75], the binary exponential backoff may reduce the efficiency and cause unfairness in the presence of channel errors. This happens since in the case of channel error backoff window size increases unnecessarily. Hence, the channel is unused more often and the efficiency is reduced. Additionally, as it is shown in the following, when STAs suffer from different packet error rates (PER) fairness cannot be achieved. Therefore, it is important to modify the IEEE 802.11 system such that it can distinguish between channel errors and collisions.

In Figure 3.9 the throughput of an asymmetric case is plotted where half of the STAs in the BSS have a PER of zero and the other half have a PER of $P_e = 50\%$. A data rate of 58.5 Mb/s and basic rate of 6.5 Mb/s are assumed. In order to obtain these results we need to distinguish between these two groups of STAs and adapt the respective equations in previous sections as follows.

Assume that there are totally $n$ STAs and $n_1$ STAs have no channel error, while $n_2$ of them have a non-zero $P_e$. The transmission probability of STAs in the first group is $\tau_1$ and
the transmission probability of the STAs in the second group is $\tau_2$. In this case collision probability of each STA in one of these groups and the probability that in a time slot only one STA transmits can be obtained from:

$$
\begin{align*}
P_{\text{col}1} &= 1 - (1 - \tau_1)^{n_1-1}(1 - \tau_2)^{n_2}, \\
P_{\text{col}2} &= 1 - (1 - \tau_2)^{n_2-1}(1 - \tau_1)^{n_1}.
\end{align*}
$$

(3.25)

$$
\begin{align*}
P_{s_1} &= n_1 \tau_1 (1 - P_{\text{col}1}), \\
P_{s_2} &= n_2 \tau_2 (1 - P_{\text{col}2}).
\end{align*}
$$

(3.26)

The throughput of each group of STAs can be calculated as:

$$
\begin{align*}
S_1 &= \frac{P_{s_1}L}{T_{\text{slot}}}, \\
S_2 &= \frac{P_{s_2}(1 - P_e)L}{T_{\text{slot}}},
\end{align*}
$$

(3.27)

where $T_{\text{slot}}$ is:

$$
T_{\text{slot}} = P_{\text{id}}T_s + \left(P_{s_1} + P_{s_2}(1 - P_e)\right)T_s + P_{s_2}P_eT_e + P_cT_e,
$$

(3.28)

in which $P_{\text{id}}$ is the probability that the channel is idle and $P_c$ is the probability of collision. These values can be calculated from:

$$
\begin{align*}
P_{\text{id}} &= (1 - \tau_1)^{n_1}(1 - \tau_2)^{n_2}, \\
P_c &= 1 - P_{\text{id}} - P_{s_1} - P_{s_2}.
\end{align*}
$$

(3.29)

(3.30)

$T_s$, $T_c$ and $T_e$, can be calculated from (3.16)-(3.18), respectively.

The user throughput of the symmetric scenarios, i.e., where for all STAs once $P_e$ is set to zero and another time to 50%, are also plotted in Figure 3.9 for comparison. It should be noted that this is an extreme case and in practice situations with much lower channel error probabilities are desirable. However, analysis of such extreme cases, shows the existing problems and helps us to understand the performance bounds. As it is shown in Figure 3.9,

\footnote{According to the IEEE 802.11a standard the packet error rate should be less than 10% at a PSDU length of 1000 bytes for the considered minimum receiver sensitivities [37].}
in the asymmetric case STAs with a poor wireless link suffer significantly from their backoff
growth due to the link error. The throughput of these STAs is even smaller than that of the
symmetric scenario where all STAs have $P_e = 50\%$. However, in the asymmetric scenario
the STAs with $P_e$ of zero benefit from weak performance of the other group. Consequently,
the throughput of these STAs becomes significantly larger than that in the symmetric scenario
where all STAs have no channel error.

In [74] the authors suggest a loss-distinguishable MAC based on the reception of CTS
and ACK frames in RTS/CTS access method. According to their proposed scheme if the
transmitter receives the CTS frame in response to its RTS packet but it does not receive the
ACK frame it assumes that a link error has occurred. For the basic access method, they
suggest to send a NAK control frame whenever the MAC header is received correctly but the
MAC body is received with error. In this case the transmitter can presume that a link error
has occurred.

**Real-Time Applications:** Throughput analysis for real-time applications can differ from
data applications in different ways. For real-time applications, it is desirable to schedule
STAs in a deterministic way. Hence, contention-free access methods are preferable. The
contention-free access methods, defined by the IEEE 802.11 standard, are based on a polling
scheme, cf. Section 2.1.

However, assuming that real-time applications should also be supported during the conten-
tion period (for example in case contention-free method has not been implemented in a sys-
tem), then following points should be considered.

- ARQ methods introduce additional delay to the system. Consequently, in real-time
applications a packet may be sent only once, regardless of the status of the transmis-
sion, i.e. successful or failure. In this case the number of retry limit, $m'$ is equal to
zero and we have $m = m' = 0$. Accordingly in the analysis of the throughput, one
needs to replace (3.9) by $\tau = \frac{2}{(W+1)}$ [6]. In this case, contention window is fixed to its
minimum value and a packet is dropped with the probability $p^{m+1} = p$.

- In case of voice transmissions, size of packets are usually much smaller than that in
data applications. Since there are always some constant overheads attached to a frame,
having a short payload makes the contention access methods inefficient.

- In a system where only short voice packets are transmitted, analysis of the throughput
in saturation mode may not be useful anymore. In this case, as each user generates
a short voice packet with interval of a few milliseconds, one needs to take the input
arrival rate into account.
In this section, we consider a real-time application, where packets of 1024 Byte are transmitted at a data rate of 6.5 Mb/s. We assume that the system is saturated but there is no retransmission attempt. The throughput of such a system is plotted in Figure 3.10. As it is seen in this figure, the throughput of the real-time application degrades substantially with increase of number of STAs. This is mainly due to the constant channel access probability which leads to a high collision probability and consequently a high packet drop probability for large number of STAs. As shown in [75], to improve the throughput in the absence of the exponential backoff, one needs to optimise the contention window parameters. In [75] the authors find an optimum value for \( \tau \) (the transmission probability) such that \( T_{\text{slot}} \) is minimised and therefore throughput is maximised.

**QoS Issues:** In order to support real-time applications, the IEEE 802.11e amendment has been proposed. The IEEE 802.11e provides QoS by defining different access categories (ACs). Throughput of IEEE 802.11e WLANs has been studied in different papers [41] [76]. In [76] a simulation-based study of IEEE 802.11e is presented while in [41] both analytical model and simulations are studied. The IEEE 802.11e throughput analysis shows that it can improve the throughput of high-priority streams while the low-priority streams significantly suffer with increase in traffic with higher priorities.

**MIMO Systems:** So far all throughput results in this section are calculated for a single
spatial stream. In IEEE 802.11n MIMO systems up to four spatial streams can be supported. However, in point-to-point systems, to support \( k \) spatial streams, there should be at least \( k \) antennas at both ends of a communication link. Furthermore, as it has been shown in [60] the channel utilisation efficiency of the IEEE 802.11n system can be very poor (even around 10\%). This happens when at each transmission attempt a single packet is transmitted at a very high rate over multiple antennas. In this case the overheads occupy a significant fraction of each transmission duration. Therefore, in practice the use of packet aggregation in point-to-point MIMO systems should be considered. In addition to frame aggregation, block ACK and reverse direction protocol can further improve the IEEE 802.11n throughput [1] [77].

**Fairness Issues:** For a perfect PHY channel, in the long term, the legacy DCF is fair, i.e., it provides an equal channel access opportunity to all STAs [69]. However, as it has been discussed, as soon as some STAs have better PHY links than the others or they support different data rates, the system is not fair anymore. Besides, recent enhancements of the IEEE 802.11 standard, like priorities in IEEE 802.11e, make the IEEE 802.11 systems even more diverse. Although utilising higher throughput at some STAs could improve the aggregate throughput and improve these STAs' satisfactions, other STAs may dramatically suffer from low throughput.

**Example:** Here, we focus on the performance enhancements with respect to individual user. As an example we consider a network of heterogenous STAs equipped with different number of antennas. Let us assume that half of the sources and their respective destinations have a single antenna while others have \( N_a > 1 \) antennas. If we intend to improve the aggregate throughput, one way is to let the multiple-antenna STAs transmit at each transmission attempt aggregated packets or to increase the channel access opportunity of these STAs. The latter can be achieved for example by scheduling them more often or increasing their TXOP. On the other hand we can improve the performance of the single-antenna nodes by using the same techniques, e.g., longer TXOP.

Let us take two different methods into consideration. In both methods, the multiple-antenna STAs transmit \( N_a \) parallel data packets at total data rate of \( N_a R \), whenever they access the channel. However, the single-antenna STAs behave differently according to the applied method.

First Method: In the first method each single-antenna STA transmits only a single packet at data rate \( R \) at each transmission attempt. In this case, assuming the same packet size, the data duration is the same for both multiple-antenna and single-antenna STAs, i.e., \( T_{Ls} = T_{Lm} \) where indexes \( s \) and \( m \) indicate single-antenna and multiple-antenna, respectively. However, while the multiple-antenna STAs transmit \( N_a L \) data bytes, the single-antenna STAs transmit
only $L$ bytes.

Second Method: According to the second method, the single-antenna STAs aggregate $N_a$ data packets into one frame successively. Hence, taking only the payload durations without headers into account, in this case $T_{L_s} = N_a T_{L_m}$. Since at each transmission attempt, multiple-antenna STAs transmit $N_a$ independent spatial streams while each single-antenna STAs sends a single stream with $N_a$ aggregated packets, we have $L_s = L_m = N_a L$ bytes. The receiver confirms reception of packets by sending a block ACK back to the source.

In this example to calculate transmission probability, we can apply (3.9)-(3.11). However, to calculate throughput in the presence of heterogeneous nodes with different packet lengths, we need to consider the following points. Since the frame length of the two groups of STAs is different we should differentiate between the data packet and ACK of the two groups. Consequently, the time, during which the channel is sensed busy due to either successful or failed transmissions, is also different for each group.

In the considered scenario, the collision probability and the probability that only one STA transmits can be calculated from (3.25)-(3.26). Assuming $P_c = 0$ we have $\tau_s = \tau_m$. Therefore, aggregate throughput of each group of STAs is defined by:

$$S_i = \frac{P_s L_i}{T_{\text{slot}}}, \quad (3.31)$$

where $T_{\text{slot}}$ and $T_{s_i}$ are:

$$T_{\text{slot}} = P_{id} \sigma + \sum_i P_{s_i} T_{s_i} + P_c T_c,$$

$$T_{s_i} = T_{\text{Data}_i} + \text{SIFS} + \delta + T_{\text{ACK}_i} + \text{DIFS} + \delta, \quad (3.32)$$

where $P_{id}$ can be calculated directly from (3.29).

To calculate $P_c$ we should differentiate between collision of packets with the same duration, $P_{ci}^{(\text{intra})}$, and that of packets with different lengths, $P_{ci}^{(\text{inter})}$. In the considered scenario there are only two different packet durations. Whenever short packets collide with long packets, the channel remains busy according to the duration of the long packets\(^{11}\). This leads

\(^{11}\)We assumed same PHY rate per link for both groups of STAs.
Chapter 3  Multimedia Traffic over WLANs

\[
P_{ci}^{(\text{intra})} = \left( 1 - \left( (1 - \tau)^{n_i} + n_j \tau (1 - \tau)^{n_i-1} \right) \right) \cdot (1 - \tau)^n, \quad j \neq i
\]
\[
P_{c1}^{(\text{inter})} = 0,
\]
\[
P_{c2}^{(\text{inter})} = (1 - (1 - \tau)^{n_2}) \cdot (1 - (1 - \tau)^{n_1}),
\]
\[
P_c = P_{ci}^{(\text{intra})} + P_{c1}^{(\text{inter})},
\]

(3.33)

where the first term in \( P_{ci}^{(\text{intra})} \) indicates the probability that more than one STA from the same group transmit and the second term indicates that no other STAs from another group transmits. It should be noted that for inter-group collisions we just need to take the collision of longer packets with shorter ones into consideration. As we have already taken the collisions of both groups in \( P_{c2}^{(\text{inter})} \) into account, we have \( P_{c1}^{(\text{inter})} = 0 \). The total time during which channel is busy due to collision is:

\[
T_c = P_{c1} T_{c1} + P_{c2} T_{c2},
\]

(3.34)

where \( T_{c1} = T_{\text{Data1}} + \delta + T_{\text{ ACKtimeout1}} + \text{DIFS} \).

Figure 3.11 shows the user throughput of both methods for both multiple-antenna and single-antenna nodes. For this analysis the parameters from Table 3.3 are used. Except that for the multiple-antenna transmissions, the PLCP header includes orthogonal training sequences which lasts 4\( \mu \)s longer for each additional spatial stream. For simplicity, all transmissions in both methods use the same data and ACK MAC headers. Moreover, both types of STAs transmit using the same modulation alphabet and hence the same PHY rate per link is considered. A data rate of 58.5 Mb/s and an ACK rate of 6.5 Mb/s are used. Number of antennas at multiple-antenna STAs is once set to two and another time to four.

As shown in Figure 3.11, using the second method, the throughput of the STAs with the single-antenna is improved substantially at the cost of throughput loss of the multiple-antenna nodes. Nevertheless, this can be a useful method in the scenarios where the weakest link determines the overall performance and/or where fairness with respect to rate is desired (max-min fairness).

3.1.6 Delay

An important measure of QoS for real-time services is delay. Usually, in high-quality real-time services the maximum acceptable delay in the system is predefined and it should not
3.1 Fundamental Bounds and QoS Limitations

![Graph](attachment:graph.png)

(a) \( N_a = 2 \)

(b) \( N_a = 4 \)

**Figure 3.11:** Throughput per STA in a network with diverse antenna configurations for method 1 and 2.

be exceeded. For example, voice applications over DECT experience a delay of around 25ms [78]. In this thesis, we consider only the packet delay introduced by the MAC protocol, i.e., head-of-line (HoL) delay. The HoL is defined by the time interval between the moment a packet reaches the head of the queue until it is successfully received by its destination. As investigated in [61] [60] for a network with \( n \) nodes, the packet delay \( D \) of a single-hop link can be calculated via Little’s Result as:

\[
D = \frac{n}{S/L}.
\]

The Little’s result states that in a queuing system the average number of users in the system is equal to the average arrival rate multiplied by the average time spent in the system [79]. The HoL delay of the considered ad hoc network in Section 3.1.4, cf. Figure 3.6, is depicted in Figure 3.12. Again both basic and RTS/CTS access methods are considered. As shown in this figure the delay of the lowest rate rapidly increases with increase in number of STAs. In our considered model, an infinite number of retransmissions have been assumed. The average HoL delay is reduced when the number of retransmissions is limited but at the expense of some dropped packets. However for small network sizes, the delay difference between the case with retry limit of 6 and the case with infinite retry is relatively small [64].

One should keep in mind that the shown values only present the delay introduced by the MAC layer. However, in each system there are other sources of delay. For example
the queueing delay, which is the waiting time a packet experiences in a queue before it is transmitted. The queueing delay depends on the traffic arrival rate and the MAC protocol. In practice, one may need to measure the end-to-end delay. However, since we are focused on the MAC and PHY layers we only take the HoL delay into account.

As it is shown in Figure 3.12, the HoL delay rapidly increases with the number of STAs. For example in a network with 20 STAs, using the basic access method and the data rate of 6.5 Mb/s, the HoL delay is about 41 ms. Hence, taking other delay sources into account, the total delay may be too large for using such setups in high-QoS voice applications.

### 3.2 Relaying Scenarios

Relaying has become a major topic in the research literature as it has a great potential for improving system performance. In this section we focus on point-to-point networks and discuss the main advantage of relaying in such networks and explain how it can be used to solve some of the problems mentioned in the previous section.

Several wireless relaying protocols are presented in the literature [19–21]. Throughout this thesis, we focus on two of them, DF and AF relaying. A DF relay decodes the data packet before forwarding it. Contrary to DF relays, AF relays only amplify the received packets and forward them without decoding. One of the advantages of AF relays is that they...
are transparent to data modulation and coding. As a result, they are less complex compared to DF relays. However, while DF relays can easily be considered in the current WLANs, the implementation of AF relays in a WLAN system needs further considerations.

In this section, at first we evaluate the outage probability performance of a system with a DF and an AF relay. Then applying the coverage angle concept, the coverage improvement of a DF relaying network is investigated and compared with that of a direct link, i.e., where source and destination communicate directly. Furthermore, we give an estimation of the number of required DF relays or APs to fully cover a circular area. Afterwards, we focus on the data rate performance and study the impact of DF relaying on PHY data rate and MAC throughput of a WLAN.

### 3.2.1 Outage Probability

In this section, we investigate the outage probability of a wireless network in a similar way to that in Section 3.1.3 but this time one or more relays assist the communication between a source and its corresponding destination. It is assumed that there is no direct link between a source and its destination. Furthermore, it is assumed that relays are not connected to the backbone and they forward the packets which they have received from different sources. Throughout this thesis, uplink is considered to be the link from the source to the relay/AP while downlink is the link from the relay/AP to the destination. In the first time slot, STA₁ transmits to the relay and in the second time slot the relay forwards the received signal to its destination, i.e., STA₂. Similarly STA₂ can transmit to the STA₁ via the relay using the third and fourth time slots, cf. Figure 3.13.

For communication links with more than one hop, depending on the requirements, the target outage rate can be defined by either end-to-end data rate or data rate per hop. As an example assume a scenario where a data rate of $R_t$ is required to support a certain application.
In this case, the outage rate can be defined as $R_{\text{out}} = R_t$ and the outage probability can be calculated by comparing the end-to-end data rate with $R_t$. However, in scenarios where a minimum data rate per hop, $R_i$, has to be supported the outage rate can be set to $R_i$. In this case, an outage occurs whenever any of the hops are in outage.

**DF Relay:** Assuming that a source communicates to its destination via a DF relay, then outage occurs if any of the source-relay or relay-destination links are in outage. The data rate of a two-hop link over a DF relay is:

$$R_{\text{DF}} = \min\{\frac{k_1}{k_1 + k_2} R_{sr}, \frac{k_2}{k_1 + k_2} R_{rd}\},$$  \hspace{1cm} (3.36)

where $R_{sr}$ and $R_{rd}$ are the data rate of the source-relay and relay-destination links, respectively. $k_1$ and $k_2$ are fractions of time that uplink and downlink transmissions occupy, respectively. Assuming that for a transmission from a source to a destination via the DF relay two channel uses of equal duration are needed then $k_1/(k_1 + k_2) = k_2/(k_1 + k_2) = 1/2$. Since in this thesis we consider the same constant data rate per each hop we have $k_1/(k_1 + k_2) = k_2/(k_1 + k_2) = 1/2$.

For DF relaying link we can extend the outage definition in (3.7) as:

$$P_{\text{DF}} = \Pr \left( \min\{R_{sr}, R_{rd}\} < R_{\text{out}} \right),$$  \hspace{1cm} (3.37)

where the outage rate, $R_{\text{out}} = R_{\text{out single-hop}}$, is the required PHY rate per hop. For the two-hop DF links, we can equally define an end-to-end outage rate as: $R_{\text{out two-hop}} = \frac{1}{2} R_{\text{out single-hop}}$.

**AF Relay:** The spectral efficiency of an AF relaying scenario is given by:

$$P_{\text{AF}} = \left( \frac{1}{2} \log_2 \left( 1 + \frac{P_s |h_{sr}|^2 |h_{rd}|^2}{\sigma_n^2 + \sigma_d^2} \beta^2 \right) \right),$$  \hspace{1cm} (3.38)

where $h_{sr}$, and $h_{rd}$ are the channel coefficients between source and relay and relay and destination, respectively. $P_s$ is the transmit power at source, $\sigma_n^2$ and $\sigma_d^2$ are the noise variance at the relay and destination. In an OFDM system, instantaneous rate per OFDM subcarrier can be obtained from (3.38).

In order to remain within a relay transmit power constraint of $P_r$, the relay amplification
3.2 Relaying Scenarios

Figure 3.14: Outage Probability of DF and AF Relaying, \(d_{sr} = 20\, \text{m}\).

The outage probability in an AF relaying scenario can be obtained by applying (3.7), (3.38) and (3.39) and replacing the \(R_{out}\) with the two-hop outage rate.

We have simulated the outage probability for both AF, and DF relaying by assuming that the source and the relay are fixed while relay-destination distance, \(d_{rd}\), changes. For the analysis a source-relay distance of \(d_{sr} = 20\, \text{m}\) and an end-to-end outage rate of 13 Mb/s (or in relaying scenarios a single-hop outage rate of 26 Mb/s) are assumed. Furthermore, the same transmit power and same noise variance at all nodes are considered. The outage probability versus source-destination distance is plotted in Figure 3.14.

In the DF relaying scenario for \(d_{rd} < 20\, \text{m}\), the uplink limits the performance and as \(d_{sr}\) is fixed, \(P_{out}\) reaches a constant value. For small values of \(d_{rd}\) the AF rate is also determined by its uplink channel. However, with increase in \(d_{rd}\), the downlink channel impacts the overall rate.
3.2.2 Coverage Angle

Sofar we have considered coverage extension using a one-dimensional model. In order to have a more realistic model we extend our scenario to a two-dimensional scenario, where an AP is located in the centre of a circular area with the radius of $d_r$. It is assumed that the AP has a coverage range of $d_{AP} \leq d_r$. Therefore, to fully cover the area we need to extend the coverage range of the AP. In order to do so, we distribute some relays uniformly around the AP such that an area with a radius of $d_\alpha \geq d_r$ is fully covered. We refer to the $d_\alpha$ as the coverage range. We consider a performance metric called coverage angle which has been introduced in [80]. In [81] we have applied the same concept to determine the coverage of an AP in an IEEE 802.11a system.

The coverage angle, $\alpha$, determines the size of the circle sector supported by one specific relay. In Figure 3.15 an example of such a network is shown where six relays extend the coverage range of an AP. As it is seen in this figure, $d_\alpha$ is the largest radius which gives a fully covered circular area. In this specific example $\alpha = 60^\circ$ hence $N_r = \lceil \frac{360}{\alpha} \rceil = 6$ relays are needed to fully cover the circular area with the radius of $d_\alpha$.

There is a tradeoff between the number of relays and the corresponding coverage angle. A smaller coverage angle leads to a larger coverage range but requires a higher number of relays. In order to investigate how many DF relays we need to fully cover a certain area, we simulate the coverage angle. Again a single-antenna AP is assumed at the centre of the area with (0,0) coordination and a DF single-antenna relay is located on the virtual horizontal line which passes through the AP. Due to symmetry, it is sufficient to estimate the coverage angle of a single relay. We change the location of the destination within the whole
3.2 Relaying Scenarios

Figure 3.16: Coverage Angle for the DF relaying network for different relay-AP distances.

Figure 3.17: 1%-outage capacity of the DF relaying network in [bps/Hz].
area. The destination connects either directly to the AP or via the relay. It chooses the link which leads to higher data rate. At each location, we calculate the 1%-outage capacity of the destination. The 1%-outage capacity is defined by the largest transmission rate by which outage probability remains smaller than 1% [15, page 188]. In this simulation an outage rate of 13 Mb/s per hop is assumed.

The simulation is repeated for different distances between the relay and the AP. Figure 3.16 shows the coverage angle versus coverage range $d_{c}$ for different relay-AP distances, i.e. different $d_{sr}$. It is seen that a maximum coverage range of about 41 m at coverage angle of 60° is reached when the relay is located at 22 m distance from the AP. If the relay gets closer or farther from the AP the range is decreased. In this scenario, to fully cover a disk-shaped area, i.e. coverage angle of 360° with radius of 41 m, one AP and six relays are required in such a way that relays are placed uniformly around the AP and all have a distance of 22 m from AP, as it is shown in Figure 3.15. We have also plotted the 1%-outage capacity when relay is 22m away from the AP (cf. Figure 3.17).

**APs versus Relays:** In Section 3.1.3 different methods for the coverage enhancement are pointed out. An intuitive way to do so is to increase the number of APs in a given area. However, this leads to higher infrastructure costs and requires careful frequency planning. In this section, we assume that we can add additional APs to the considered network. Using the coverage angle metric, we estimate the number of APs needed to fully cover a circular area.

The graph of the coverage angle versus coverage distance is shown in Figure 3.18. In this scenario, a second AP is used to extend the coverage range. As, it is shown to have a coverage range of 41 m, as in the DF relaying network, we can locate the second AP 52 m away from the first one. However, this time a coverage angle of about 91° is achieved. This means that we need only four AP which are uniformly located on a circle with radius of 52 m while the first AP is at the centre of the circle. The 1%-outage capacity for two neighbouring APs with a distance of 52 m is depicted in Figure 3.19.

It is important to note that in indoor scenarios shape of rooms, location and type of walls and furniture can change the desired coverage area significantly. For example to fully cover a long narrow corridor, it is not necessary to consider a coverage angle of 360°. Nevertheless our simulation is realistic for general scenarios and fair comparisons between adding more relays or APs can be performed based on that. In the above scenario it is more reasonable to employ relays as long as six relays are less expensive than four APs and where flexibility in the system is favoured.

By using relays instead of increasing the number of APs in such networks, flexibility is
3.2 Relaying Scenarios

Figure 3.18: Coverage Angle for the network with an additional AP for different distances between the two APs.

Figure 3.19: 1%-outage capacity of the network with an additional AP in [bps/Hz].
increased and infrastructure cost is reduced. At the same time the delay is increased and the capacity is reduced. The flexibility can be further improved in mobile scenarios. In such cases, each STA can relay packets for a neighbouring user when this user is not in the communication range of an AP. This is realistic for data but not for real-time traffic services, where to limit delay and ensure system resiliency further considerations are required.

### 3.2.3 PHY Layer Data Rate

The IEEE 802.11n supports several rates by using different modulation alphabets and/or different FEC coding rates. The IEEE 802.11n document determines the minimum receiver sensitivity for each rate [1]. Taking into account these parameters we can obtain the corresponding minimum receiver SNR values ($SNR_{min}$) using the following formula [82]:

\[
SNR_{\text{min}} = S_{\text{min}}/(kT_0 \cdot NF \cdot BW),
\]

In our analysis $T_0 = 298.15^\circ K$. Furthermore, an implementation loss of 5 dB is used, cf. Table 3.2. The minimum receiver sensitivity and SNR for all IEEE 802.11n link data rates are shown in Table 3.4.

We simulate the data rate for both direct link and DF relaying network. This time a link adaptation algorithm is considered. According to this algorithm whenever the average receiver SNR is above the calculated minimum SNR, the data rate is set to the respective link rate. In this simulation the source and the relay are fixed, $d_{sr} = 20$ m, while the destination is placed at different positions. In Figure 3.20, the data rate based on this algorithm for direct

<table>
<thead>
<tr>
<th>Modulation</th>
<th>Coding Rate</th>
<th>Minimum Sensitivity (dBm)</th>
<th>SNR range(dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPSK</td>
<td>1/2</td>
<td>-82</td>
<td>3.84 - 6.84</td>
</tr>
<tr>
<td>QPSK</td>
<td>1/2</td>
<td>-79</td>
<td>6.84 - 8.84</td>
</tr>
<tr>
<td>QPSK</td>
<td>3/4</td>
<td>-77</td>
<td>8.84 - 11.84</td>
</tr>
<tr>
<td>16-QAM</td>
<td>1/2</td>
<td>-74</td>
<td>11.84 - 15.84</td>
</tr>
<tr>
<td>16-QAM</td>
<td>3/4</td>
<td>-70</td>
<td>15.84 - 19.84</td>
</tr>
<tr>
<td>64-QAM</td>
<td>2/3</td>
<td>-66</td>
<td>19.84 - 20.84</td>
</tr>
<tr>
<td>64-QAM</td>
<td>3/4</td>
<td>-65</td>
<td>20.84 - 21.84</td>
</tr>
<tr>
<td>64-QAM</td>
<td>5/6</td>
<td>-64</td>
<td>&gt; 21.84</td>
</tr>
</tbody>
</table>
3.2 Relaying Scenarios

Figure 3.20: Link adaptation for direct link and a DF relaying link vs. $d_{rd}$.

The downlink and uplink benefit from a shorter communication path, resulting in larger data rates.

For the two-hop link, we consider an overall rate which is defined according to the minimum SNR of the first and the second hop for each channel realisation. Since for transmission from source to destination via the relay, two channel uses of equal duration are assumed, we scale this rate by a factor $1/2$. The overall data rate of the relayed link is also depicted in Figure 3.20. As it is shown in this figure, despite using two hops for transmission of a single packet, the relaying link provides higher data rate as compared to the direct link for large $d_{sd}$. However, where source and destination are close, transmitting via relay degrades the rate performance.

3.2.4 MAC Layer Throughput

The IEEE 802.11 standard does not define any wireless relaying method except for the AP. In Section 3.1.4 we have studied the throughput of single communication links in an ad hoc network. We have also considered uplink and downlink throughput of an infrastructure network. Here, we focus on the throughput of a network with two-hop communication links.

In an infrastructure network, communications between two STAs in the same BSS, i.e., intra-BSS communications, form wireless two-hop links. The throughput of intra-BSS links
Chapter 3 Multimedia Traffic over WLANs

in an infrastructure network obtained directly from (3.23),(3.24):

\[
S_D = \frac{1}{n-1} \min\{U_{\text{up}}, U_{\text{down}}\} \\
= \frac{1}{n-1} U_{\text{down}}, \text{ for } n \geq 2. \quad (3.41)
\]

Now we focus on throughput of an ad hoc scenario with two-hop links. Throughput of ad hoc networks with multihop links are analysed in different works such as [83] and [84]. We consider a BSS where there are \( n \) nodes all communicating with each other over two hops. The throughput in such a system can be calculated using [83]:

\[
S_{\text{Multihop}} = S \frac{S_{\text{RF}}}{N_{\text{hops}}}, \quad (3.42)
\]

where \( S_{\text{RF}} \) is the spatial reuse factor, \( N_{\text{hops}} \) is the average number of hops between the source and its destination (here it is equal to two) and \( S \) is the throughput of a network with single-hop links. The spatial reuse factor indicates the number of simultaneous transmissions which can be carried on the same channel in a given region. Relays should be located such that they can communicate with the source and the destination through the wireless channel. In this case, assuming operation at the same frequency carrier leads to \( \frac{S_{\text{RF}}}{N_{\text{hops}}} < 1 \). Consequently throughput of the multihop network is smaller than that of a single-hop network. Assuming that all nodes are in the radio range of each other then \( S_{\text{RF}} = 1 \). Hence, for the network with the two-hop traffic pattern we get \( S_{\text{Two-hop}} = \frac{1}{2} S \). As presented in [83] HoL delay also increases by factor of \( \frac{N_{\text{hops}}}{S_{\text{RF}}} \). In this work, in Chapters 2 to 8 we limit the number of hops to a maximum of two and consider cases where all STAs are in the radio range of each other. In Chapter 9 we consider multihop setups with more than two hops.

3.3 Summary

Presented results in this chapter show the limitations of the existing WLAN to support different types of applications with respect to coverage, MAC throughput, fairness issues and delay. However, there are other limitations like hidden node problems [85], handover, frequency planning, inter-BSS interference, etc which are not included in this work.

In this chapter only point-to-point links are taken into account. We have taken into consideration both PHY and MAC performance bounds and studied the possibility of using relays
in a WLAN and its impact on coverage, data rate, and throughput. By installing relays in the border of the AP coverage area, we can extend the coverage range of the AP. However, the relaying may reduce spectral efficiency and throughput due to additional hop (resource) which is used for transmission of each single packet.

In the next part, we study cooperative networks utilising point-to-multipoint and multipoint-to-point links.
Part II

Cooperative WLAN Protocols
Chapter 4

Representative PHY Layer Cooperative Protocols

In the previous part, the performance of the existing WLANs with respect to PHY and MAC layers were studied and the fundamental bounds of the current systems were discussed. In this part, we study cooperative protocols in WLANs and provide solutions to the coverage and data rate limitations which are discussed in Sections 3.1 and 3.2.

In this chapter, we focus on the PHY layer and for the moment we neglect the MAC scheme. Although in practice the MAC layer is necessary and its performance impacts the overall system performance, this assumption allows us to investigate possible signal processing techniques in a WLAN independent of the MAC protocol. We take into consideration distributed MIMO schemes, which are based on multiuser interference cancellation (MUIC). In order to efficiently decode an intended data packet in the presence of multiuser transmissions, the interference from other nodes has to be cancelled out. We refer to this as MUIC. In this work, we consider the MUIC techniques in the following networks:

**Point-to-Multipoint/Multipoint-to-Point Networks:** In these networks, a multiple-antenna node at one side of the transmission is able to mitigate the multiuser interference. This can be efficiently done in the following situations:

- The receiver has $N_r$ antennas while there are $K$ transmitters, each with $N_t$ antennas. In this case in order to have $N_t$ independent streams from each transmitter the receiver needs $N_r \geq KN_t$ antennas [86].

- CSIT is available and the transmitter has more antennas than the sum of the antennas at individual receivers. In this situation, the transmitter can for example perform transmit beamforming as explained in Section 4.2.
Distributed MIMO Networks: In distributed MIMO networks, individual nodes share their antennas and form a virtual antenna array [22]. Hence, the MIMO gain can be leveraged in a system with single-antenna nodes. Contrary to the point-to-multipoint/multipoint-to-point networks, here cooperation among individual nodes is needed to carry out MUIC.

Relay-based Networks: In this work we consider three different relaying scenarios as representative examples of the above-mentioned cooperative networking:

- MIMO DF relaying: here we consider the extension of the point-to-point MIMO networks to the multiuser MIMO networks with two-hop communication links, where multiple users transmit simultaneously.

- Two-way relaying: the two-way relaying [26] is a relaying scheme that avoids the pre-log factor $1/2$ in the rate expression which is introduced by the relay in two-hop links. Here we investigate the integration of two-way DF relaying in IEEE 802.11n and analyse its outage probability. Contrary to [26] in which the authors optimised power allocation at the relay with respect to sum rate, we find a simple power allocation to improve link rate. Besides, we briefly discuss an asymmetric setup, where one user has higher data rate than the other one, and an adapted version of the two-way relaying which can easily be supported by the existing standard.

- Multiuser zero-forcing relaying (MUZFR): in a MUZFR scenario each source communicates with its destination via several AF relays [27]. As shown in [27] by appropriate allocation of the gain factors at the AF relays, multiuser interference can be cancelled out.

Figure 4.1 depicts a general relaying scenario where several single-antenna STAs communicate via relay/s. In the following sections we consider each relaying scenario and study its PHY layer performance in details.

Simulation Setup: In all scenarios we have assumed an OFDM PHY with a uniform power allocation across all subcarriers. For each scenario, both outage probability and data rate are analysed. In relaying networks, there are several single-antenna sources and destinations and one or more single- or multiple-antenna relays (depending on the scenario). To estimate coverage range, we fix the location of the relay/s with respect to the sources and change the distance between destination(s) and relay(s) and simulate the outage probability versus distance. Furthermore, we assume a constant target rate per user. If a user can support an instantaneous rate greater than or equal to the target rate, it transmits with the target rate otherwise its rate drops to zero. We will plot the data rate versus distance. It is assumed that the AP has full CSI while all other nodes (including two-way relay) only have CSI at the
4.1 Multiuser Single-Hop Transmissions

The first scenario is a typical single-hop scenario in WLANs which includes uplink and downlink transmissions in an infrastructure network. The network includes a multiple-antenna AP and several single-antenna STAs. The standard IEEE 802.11n supports point-to-point MIMO systems [1]. Accordingly, in such an infrastructure network, STAs access the channel one after the other. On the other hand the distributed spatial multiplexing can increase capacity significantly if multiple STAs are able to transmit simultaneously. Therefore, to allow multiple STAs to transmit concurrently, an enhancement is made to the standard system. Since the AP is equipped with several antennas it can decode multiple streams in the uplink (for example by successive interference cancellation\(^1\)), or separate them in the downlink transmissions (for example by performing zero-forcing). In this way, the uplink is the multiple access phase while the downlink is the broadcast phase. It is assumed that the AP has full CSI of the uplink and downlink channels, while the STAs only have the CSI at the receiver side. The uplink channel is a MIMO multiple access channel without CSIT.

---

\(^1\)One way to perform multiuser detection is to apply parallel or successive interference cancellation. In parallel interference cancellation, the interference is cancelled out after users are detected simultaneously. In successive interference cancellation users are detected one at a time and then they are subtracted out from users to be detected [14].
The broadcast channel capacity region for Gaussian parallel channels has been studied in [87] and in a similar way for OFDM systems in [88]. In [87] in addition to the capacity region for parallel broadcast channels, power allocations to achieve all points on the boundary of the capacity region and the region characterization for a given power constraint are given. The broadcast capacity region of OFDM multiuser networks under a sum power constraint is described in [88].

The multiple access channel capacity region for OFDM systems has been considered in [89] and [90]. Let $R_{up}^i$ be the instantaneous rate of STA $i$. We define the uplink rate tuple $(R_{up}^1, R_{up}^2, ..., R_{up}^{N_{mac}})$. The OFDM symbols consist of $N_{sub}$ data subcarriers out of $N_{tot}$ subcarriers. Thus MIMO multiple access channel capacity region is defined by the following system of inequalities:

$$\sum_{j \in A} R_{up}^j \leq \frac{1}{N_{tot}} \sum_{l=1}^{N_{sub}} \log_2 \left( \det (I_{N_a} + \sum_{j \in A} \frac{P}{\sigma_n^2} h_{j,l} h_{j,l}^H) \right),$$

(4.1)

for all $A \subseteq N$ where $N$ includes index of all STAs. As we consider the rates in bps/Hz the sum is normalised to the total number of subcarriers. The channel vector $h_{j,l}$ consists of all channel coefficients between STA $j$ and the AP at the $l$th subcarrier. The transmit power per data subcarrier at each node, i.e., STA and AP, is set to $P$ and the noise is additive white Gaussian with zero mean and variance $\sigma_n^2$. In our case we require that all STAs transmit with a rate $R_{up}^i = R_{out}$. The uplink is in outage if any of the inequalities in (4.1) are not satisfied. This scenario is a subset of the multiuser DF relaying scenario, which is considered in the next section.

### 4.2 Multiuser DF Relaying

As the first relaying scenario we consider a typical infrastructure network, where the AP acts as a wireless DF relay for STAs within its BSS. Therefore, for each pair of nodes, there is a source-destination link via the AP which establishes a wireless two-hop link. Figure 4.2 shows this scenario where there are an AP with $N_a = 4$ antennas and four sources and destinations, each equipped with a single antenna. In the following, first we consider the uplink and then the downlink phase.
4.2 Multiuser DF Relaying

Figure 4.2: The MIMO DF relaying scenario for four source-destination pairs.

**Uplink Phase: Multiple Access Phase**

During the first hop, i.e., uplink, all sources transmit simultaneously to the AP. The uplink outage probability can be obtained from (4.1), as it is explained in the previous section.

**Downlink Phase: Broadcast Phase**

The capacity of the broadcast channel with CSIT can be achieved by dirty paper coding [91]. However in practice dirty paper coding is difficult to implement. We assume that the AP performs transmit beamforming to orthogonalise the channels among different STAs. In this work we consider zero-forcing beamforming, i.e., the multiuser interference is nulled by the beamforming algorithm [92].

In our setup as all STAs are equipped with a single antenna, a perfect orthogonalisation is only possible if the number of antennas at the AP, $N_a$, is equal to or greater than number of STAs which are going to be served at the same time. In order to do so the AP should use a precoding matrix to be equal to the pseudo-inverse of the downlink channel. Accordingly, the data packets from the AP to the destinations can be transmitted simultaneously without causing interference at destinations.

The IEEE 802.11n also supports transmit beamforming, as one of its optional features, but only for point-to-point communication links [1].

To calculate outage, we apply the following algorithm. In order to support a rate of $R_{\text{out}}$ on each link, the AP needs a minimum transmit power for each link. Since the subchannels are orthogonalised and uniform power loading per subcarrier is assumed, we just need to allocate a minimum required power of $P_i$ per subcarrier to each link in order to support
In the DF relaying scenarios $d_{sr}$ is 20 m.

a data rate of $R_{\text{out}}$ over the whole bandwidth. In this case, outage occurs if the required transmit power (summed up over all antennas at the AP) exceeds the maximum allowed transmit power per subcarrier. Hence, the downlink outage probability can be obtained by the following formula:

$$P_{\text{out}}^d = \Pr \left( \sum_{i=1}^{N_a} P_i > P \right).$$

(4.2)

It is assumed that a source-destination (SD) link is in outage if its uplink or downlink channel is in outage.

**Simulation Results:** Let us assume an AP with $N_a = 4$ antennas and $N_{nc} \leq N_a$ sources which transmit concurrently. Furthermore, let us assume that all sources have the same distance to the AP. Similarly, all destinations have the same distance to the AP. The distance between the AP and the sources is set to 20 m while the distance between destinations and the AP is increased. We simulate the outage probability for $N_{nc} = 4$ and the reference case with $N_{nc} = 1$. All other simulations parameters are taken from Table 3.2.

The case $N_{nc} = 1$ corresponds to the point-to-point system and it is considered as the reference system. In the reference system, a two-hop scenario is considered where sources transmit to the AP, one after the other. The AP performs transmit beamforming and forwards
4.2 Multiuser DF Relaying

Figure 4.4: Sum rate in bps/Hz vs. distance between the source and the destination, i.e., $d_{sd}$. In the DF relaying scenarios $d_{sr}$ is 20 m.

The data packet to a single destination at a time. Therefore, in the reference system there is no multiuser interference and the multiple-antennas at the AP are used for enhancing the link quality. Uplink and downlink channels are SIMO and MISO channels, respectively.

We depict the outage probability versus distance between the source and the destination(s) in Figure 4.3. The figure shows the outage probability of the two-hop transmission links via the AP in both multiuser MIMO DF and Reference scenarios. For comparison purpose, the outage probability of the direct link between a source and its destination is also plotted. It should be noted that x-axis denotes the distance between a source and its destination which is the sum of the source-relay and relay-destination distances in two-hop scenarios.

A target end-to-end outage rate of 0.65 bps/Hz, which is equivalent to an outage rate of 1.3 bps/Hz per hop for two-hop links, is assumed. For the considered bandwidth, this leads to an end-to-end outage rate of 13 Mb/s which is one of the IEEE 802.11n supported uncoded data rates on the PHY layer. For $N_{nc} > 1$ it is assumed that the network is in outage whenever any of the concurrent transmissions is in outage

As it is seen, the SD link has a lower outage probability in the reference system than in the case of $N_{nc} = 4$. The reason is that in both cases we have the same number of antennas at the AP. When multiple STAs transmit concurrently, the antennas at the AP are used to

\[ \text{Reference, } N_{nc}=1 \]
\[ \text{MIMO DF, } N_{nc}=4 \]
\[ \text{Direct link, } \text{SISO} \]
cancel the multiuser interference. In the reference system, since at each time instant only one STA transmits to the AP, all antennas at the AP are used for one link. Hence, at a certain distance, reference system has lower outage probability than the multiuser MIMO DF case. However, the outage probability does not represent the overall performance. While the reference system provides the PHY rate of 0.65 bps/Hz for a single STA, the multiuser MIMO DF scenario supports four STAs at each transmission attempt. As it is depicted in Figure 4.4 the multiuser MIMO DF at high SNR regime achieves a sum rate which is four times larger than that in the reference scenario.

4.3 Two-Way Relaying

As it has already been discussed in Chapter 3, the conventional half-duplex DF relays suffer from a reduced spectral efficiency because of the factor 1/2 (cf. (3.36), (3.38)). To compensate for this loss, a signalling scheme called two-way relaying has been proposed in [26]. The two-way relaying increases the spectral efficiency of an IEEE 802.11n as we have shown in [93] and explained in more details in the following paragraphs.

Let us assume two single-antenna STAs, e.g., STA₁ and STA₂, which have packets for each other and a multiple-antenna DF relay. In a two-way relaying scenario at first both STAs transmit their symbols to the relay in the same time slot (time slot 1) and the same bandwidth (cf. Figure 4.5). The relay decodes both streams and forwards a linear combination of the two symbols of STA₁ and STA₂, e.g. the sum of them, in the second time slot. Since STA₁ and STA₂ know their own transmitted symbols they can subtract the back-propagating self-interference prior to decoding and decode the unknown symbol.

We assume full CSI only at the receivers. We consider \( h_{11} \) and \( h_{12} \) as the vectors of the channel coefficients between STA₁ and the relay in uplink and downlink, respectively. For
the sake of simplicity, we focus on one OFDM subcarrier and neglect the subcarrier index. In order to obtain the capacity expressions, we first need to consider the conventional relaying rate (one-way relaying). The instantaneous rate per OFDM subcarrier in the conventional relaying can be calculated by the following relationships:

\[
R_{DF}^1 = \frac{1}{2} \min \left\{ f_R \left( \frac{P}{\sigma_n^2}, h_{11} \right), f_R \left( \frac{P}{\sigma_n^2}, h_{22} \right) \right\}, \]

\[
R_{DF}^2 = \frac{1}{2} \min \left\{ f_R \left( \frac{P}{\sigma_n^2}, h_{12} \right), f_R \left( \frac{P}{\sigma_n^2}, h_{21} \right) \right\}, \tag{4.3}
\]

where we use the auxiliary function \( f_R \) to denote the instantaneous rate of a SIMO/MISO system per OFDM subcarrier with a channel vector \( h \):

\[
f_R \left( \frac{P}{\sigma_n^2}, h \right) = \log_2 \left( 1 + \frac{P}{M_n \sigma_n^2 \|h\|_F^2} \right), \tag{4.4}
\]

where number of transmit antennas, \( M_t \), is equal to one for the uplink channels and \( \|h\|_F^2 \) denotes the squared Frobenius norm of vector \( h \). In the two-way relaying, there is a multiple access channel during the first time slot (uplink) and a broadcast channel during the second time slot (downlink). Therefore, the data rate for STA\(_1\) and STA\(_2\) in the two-way relaying scenario can be calculated by the following formulae, respectively:

\[
R_{2way}^1 = \min \left\{ \frac{1}{2} R_{MA}, \frac{1}{2} \min \left\{ f_R \left( \frac{P}{\sigma_n^2}, h_{11} \right), f_R \left( \frac{\beta P}{\sigma_n^2}, h_{22} \right) \right\} \right\},
\]

\[
R_{2way}^2 = \min \left\{ \frac{1}{2} R_{MA}, \frac{1}{2} \min \left\{ f_R \left( \frac{P}{\sigma_n^2}, h_{21} \right), f_R \left( (1 - \beta) \frac{P}{\sigma_n^2}, h_{12} \right) \right\} \right\}, \tag{4.5}
\]

where \( \beta \) is a factor which controls the power allocation of the relay in the second time slot and \( R_{MA} \) is the multiple access rate and it is given by:

\[
R_{MA} = \frac{1}{2} \log_2 \left( 1 + \frac{P}{\sigma_n^2 \|h_{11}\|_F^2} + \frac{P}{\sigma_n^2 \|h_{21}\|_F^2} \right). \tag{4.6}
\]

The factor 1/2 in front of \( R_{MA} \) in (4.6) assures that both \( R_{2way}^1 \) and \( R_{2way}^2 \) are always in the rate region of the two-user multiple access channel. It should be noted that this is not optimal with respect to the sum rate, but it provides a fair situation for both users in scenarios where the data traffic is symmetric in both directions.

If \( \beta \) is equal to 1/2, the relay allocates the power equally to both STA\(_1\) and STA\(_2\). However, assuming a sum power constraint at the relay, in asymmetric situations the factor \( \beta \) can be
used to strengthen the weaker channel by taking power away from the stronger channel. A power allocation for SISO two-way relaying that is optimal with respect to the sum rate is derived in [94]. In the next section we calculate the optimum $\beta$ with respect to the outage probability. To enable a successful cancellation of the self-interference at the two STAs, the power distribution factor $\beta$ has to be known to both STAs.

In [95] the two-way relaying is extended to the multiuser case where more than two STAs can be served in each transmission attempt. However as the principle of the two-way relaying remains the same, for simplicity we restrict our study to two users.

**Simulation Results:** We assume that communications between STA$_1$ and STA$_2$ are performed only via the relay. Distance between STA$_1$ and the relay ($d_1$) is set to 20 m while distance between STA$_2$ and the relay ($d_2$) changes. For the simulations again an OFDM system with parameters given in Table 3.2 are assumed. We have considered a WLAN system which operates over 20 MHz bandwidth, we have applied (3.6) and adapted (4.3)-(4.6) to calculate the overall OFDM data rate.

It is assumed that the relay does not have CSIT and it allocates power equally to both STAs, i.e., $\beta = 1/2$. Outage probability and data rate of both two-way relaying scenario and the conventional DF relaying case are plotted in figures 4.6 and 4.7, respectively. As it is shown in these plots, although sum rate has been improved by using the two-way relaying method, outage probability of STA$_1$ is higher compared to that of the conventional relaying. This is due to the fact that the relay forwards packets of STA$_1$ in the second hop, i.e. the longer hop, with $P/2$ while it uses the full power $P$ in the conventional case. Besides, in the absence of CSIT at the relay, the 1x4 SIMO channels, i.e. channels from STAs to the relay, on average support higher data rates than the 4x1 MISO channels, i.e., channels from the relay to each STA. STA$_2$ has the same outage probability values in both two-way and conventional DF relaying scenarios. For STA$_2$ the bottleneck is in its first hop where it transmits, in both scenarios, with full power $P$ over a SIMO channel.

Assuming that the relay knows the second-order statistics of the downlink channels then, as shown in the next section, it is able to allocate power in a better way to each STA, such that the outage probability of the weaker STA is improved.

### 4.3.1 Optimising Power Allocation at the Relay

So far in the downlink, the relay has allocated equal power to both links. Equal power allocation at the relay is a simple method which is specially interesting in cases where the
4.3 Two-Way Relaying

![Graph](image)

**Figure 4.6:** Outage probability of the two-way relaying vs. distance between relay and STA$_2$ with uniform power allocation at the relay.

![Graph](image)

**Figure 4.7:** Data rate of the two-way relaying vs. distance between relay and STA$_2$ with uniform power allocation at the relay.
Chapter 4  Representative PHY Layer Cooperative Protocols

relay has neither CSIT nor information about path loss. However, to improve the outage behaviour in the two-way relaying scenario, the relay should allocate the power optimally to each link. Since we are interested in improving the outage probability we should allocate more power to the STA with the weaker channel (here STA$_1$) while reducing the power allocated to the stronger link, i.e. the link between the relay and STA$_2$. In other words we need to find the $\beta$ which fulfils:

$$R_{1\text{way}}^{2way} = R_{2\text{way}}^{2way}. \quad (4.7)$$

Considering this equation and with the assumption of equal transmit power over all frequency subcarriers we calculate the parameter $\beta$. In order to reduce the calculation overhead at the relay we assume that the relay calculates $\beta$ only based on the distribution of the channels and not based on instantaneous channel realisations. However, similar analysis can be performed to find the solution for the latter case.

In order to find optimum $\beta$ we consider two different regions: (i) region I where $d_1 < d_2$ and (ii) region II where $d_1 \geq d_2$. We define $C(P/\sigma_n^2, h) = \mathbb{E}[f_R(P/\sigma_n^2, h)]$.

**Region I**

Having $d_1$ smaller than $d_2$ results in $C(P/\sigma_n^2, h_{22}) < C(P/\sigma_n^2, h_{11})$. Since for all $\beta \in [0, 1], C(\beta P/\sigma_n^2, h)$ is a strictly increasing function of $\beta$ we have:

$$C(\beta P/\sigma_n^2, h_{22}) < C(P/\sigma_n^2, h_{11})$$

and therefore:

$$\mathbb{E}[R_{1\text{way}}^{2way}] = \frac{1}{2} C(\beta P/\sigma_n^2, h_{22}), \quad (4.8)$$

$$\mathbb{E}[R_{2\text{way}}^{2way}] = \begin{cases} 
\frac{1}{2} C((1 - \beta) P/\sigma_n^2, h_{12}) & \text{for } \beta \geq 1 - M_t(d_1/d_2)^\gamma \\
\frac{1}{2} C(P/\sigma_n^2, h_{21}) & \text{for } \beta < 1 - M_t(d_1/d_2)^\gamma 
\end{cases} \quad (4.9)$$

For $\beta < 1 - M_t(d_1/d_2)^\gamma$ there is not any $\beta$ which fulfills (4.7). But for $\beta \geq 1 - M_t(d_1/d_2)^\gamma$, by setting $\mathbb{E}[R_{1\text{way}}^{2way}]$ equal to $\mathbb{E}[R_{2\text{way}}^{2way}]$ we have:

$$C(\beta P/\sigma_n^2, h_{22}) = C((1 - \beta) P/\sigma_n^2, h_{12}). \quad (4.10)$$

96
or

\[
\log_2 \left(1 + \frac{\beta P}{M_i \sigma_n^2} \mathbb{E}[\|h_{22}\|^2_F]\right) = \log_2 \left(1 + \frac{(1 - \beta)P}{M_i \sigma_n^2} \mathbb{E}[\|h_{12}\|^2_F]\right). 
\]  
(4.11)

in which the order of expectation and logarithm operator is changed by using the Jensen’s inequality [25]. Since the loss is the same in both sides, equality is still satisfied.

Assume \(a_i\) and \(b_i\) to be the elements of channel vectors \(h_{22}\) and \(h_{12}\), respectively. Since \(\mathbb{E}[|a_i|^2] = 1/d_2^\gamma\) and \(\mathbb{E}[|b_i|^2] = 1/d_1^\gamma\) from (4.10) we obtain:

\[
\beta^* = \frac{\mathbb{E}[\|h_{12}\|^2_F]}{\mathbb{E}[\|h_{22}\|^2_F] + \mathbb{E}[\|h_{12}\|^2_F]} = \frac{d_2^\gamma}{d_2^\gamma + d_1^\gamma},
\]  
(4.12)

where \(\gamma\) is the path loss exponent.

**Region II**

Like the previous case, since \(d_1 \geq d_2\) we have:

\[
\mathbb{E}[R_{2\text{Way}}^2] = \frac{1}{2} f_R((1 - \beta)P/\sigma_n^2, h_{12}),
\]  
(4.13)

\[
\mathbb{E}[R_{1\text{Way}}^2] = \begin{cases} 
\frac{1}{2} f_R(\beta P/\sigma_n^2, h_{22}) & \text{for } \beta \leq M_i \left(\frac{d_2}{d_1}\right)^\gamma \\
\frac{1}{2} f_R(P/\sigma_n^2, h_{11}) & \text{for } \beta > M_i \left(\frac{d_2}{d_1}\right)^\gamma
\end{cases}
\]  
(4.14)

By solving (4.7) in a similar way as the analysis for region I, we get the same value for the optimum \(\beta\) as in (4.12).

We assume that the relay knows the second-order statistics of the downlink channels or the path loss/distance from each station and hence it can easily calculate the \(\beta^*\). Each STA may achieve path loss knowledge in different ways:

- If the relay process time is known at the STAs, then in the two-way-relaying scenario each STA can calculate the round trip time of its own transmitted signal and from that it can estimate its distance to the relay. The STAs feedback these values to the relay and the relay uses the received information to calculate \(\beta^*\). If one of the nodes moves to another location it should estimate its new distance and inform the relay accordingly.
- Assuming that the path loss exponent of the environment is known, then the relay can estimate \(\beta^*\) using (4.12). Since, the exact path loss exponent may be unknown, it is
Figure 4.8: Outage probability of the two-way relaying vs. distance between relay and STA₂ with $\beta = \beta^*$. 

better to calculate the path loss directly rather than obtaining that from the distance, as it is explained in the following paragraph.

- According to IEEE 802.11k amendment [96] a STA can request another STA to respond with a Link Measurement Report frame. The link measurement report frame contains a transmit power control element which indicates the power used to transmit this frame. In this way link measurement report enables measurement of link path loss.

Simulation Results: We keep the same setup as before except that now we choose $\beta = \beta^*$. The simulation results are shown in Figure 4.8 and Figure 4.9. As shown, in two-way relaying setup, both STA₁ and STA₂ have the same outage probability and their performance is very close to that of the conventional relaying method. Performance of the conventional relaying method is limited by the performance of the weaker link (here the link from STA₁ to STA₂). In this way, by using the optimum power allocation at the relay, the same coverage range for both two-way and conventional relaying methods are achieved. However, the two-way relaying still provides higher spectral efficiency compared to the conventional DF relaying, cf. Figure 4.9.
4.3 Two-Way Relaying

Figure 4.9: Data rate of the two-way relaying vs. distance between relay and STA2 with \( \beta = \beta^* \).

4.3.2 Asymmetric Traffic

In the previous setups, STA1 and STA2 have transmitted with the same data rate to each other. However, due to different channel properties or different application types, the data traffic and its rate may not be the same for both STAs. For example consider the case where \( R_1 > R_2 \), in which \( R_i \) is the uplink data rate of STA\(_i\). In this case as we proposed in [97] the AP can transmit multicast packets to both STAs with a rate proportional to the \( \min\{R_1, R_2\} = R_2 \). However, as during a certain time more packets arrive at the AP from STA1 than from STA2, the AP can transmit the additional packets as unicast packets to the STA2 with a data rate proportional to the difference of \( R_1 \) and \( R_2 \).

4.3.3 Semi-Two-Way Relaying

The two-way relaying can not be implemented according to the actual IEEE 802.11n. One reason is that in a BSS the two STAs ideally avoid collision and hence cannot transmit at the same time. Besides the relay should be able to either combine the two received packets (symbolwise or bitwise) or process the packets in a way that each user’s packet is forwarded over one of its antennas. In the latter case, the interleaving and coding should be adapted such that each antenna transmits only data bits destined to one of the users.
In order to overcome the first problem, namely simultaneous transmissions by two users, we investigate a "semi-two-way" relaying method in which the uplink is handled similar to the standard, i.e., each STA transmits to the relay in an individual time slot. However, the relay decodes the received packets, performs superposition coding and multicasts the new packet in the third time slot to both STAs. Figure 4.10 shows this scenario which is here called *semi-two-way* relaying.

The achievable rates by STA 1 and STA 2 in the semi-two-way relaying scenario are respectively:

\[
R_{\text{semi-2way}}^1 = \frac{1}{2} \min \left\{ f_R(P/\sigma_n^2, h_{11}), f_R(\beta P/\sigma_n^2, h_{22}) \right\},
\]

\[
R_{\text{semi-2way}}^2 = \frac{1}{2} \min \left\{ f_R(P/\sigma_n^2, h_{21}), f_R((1 - \beta)P/\sigma_n^2, h_{12}) \right\}.
\]

(4.15)

### 4.4 Multiuser Zero-Forcing Relaying

In this section we consider a scenario in which there are no any AP or DF relay which assist the communication. Instead, there are some idle nodes which can act as AF relays. As one example of such a cooperative relaying network we choose a multiuser zero-forcing relaying (MUZFR) network [27], where each source communicates with its destination via several AF relays. Let \( N_r \) be the number of single-antenna AF relays. In [27] it has been shown that by appropriate allocation of the gain factors at the AF relays, multiuser interference can be cancelled out. In order to perform interference cancellation for \( N \) SD pairs, at least

\[
N_r = N(N - 1) + 1.
\]

(4.16)

relays are required [27]. During the first hop, \( N \) sources transmit to all relays concurrently through the same physical channel. The relays scale and rotate the received signals, and
forward them in the second hop to the $N$ destinations.

Let $\mathbf{h}_{SR}^i$ be the vector of channel coefficients from the $i$th source to all relays, and $\mathbf{h}_{RD}^j$ the vector of channel coefficients from all relays to the $j$th destination at a certain subcarrier. The equivalent channel matrix $\mathbf{H}_{SD}$ for this subcarrier is the concatenation of the source-relay channel matrix $\mathbf{H}_{SR} \in \mathbb{C}^{N_r \times N}$, the relay gain factors, and the relay-destination channel matrix $\mathbf{H}_{RD} \in \mathbb{C}^{N \times N_r}$. The elements of the equivalent channel matrix at this subcarrier can be written as [27]:

$$
\mathbf{H}_{SD}[i,j] = g^T \cdot (\mathbf{h}_{RD}^j \odot \mathbf{h}_{SR}^i),
$$

(4.17)

where $g \in \mathbb{C}^{N_r}$, is the relay gain vector. The operator $\odot$ indicates Hadamard (element-wise) product. We define the matrix $\mathbf{H}_I$ as:

$$
\mathbf{H}_I[i,j] = (\mathbf{h}_{RD}^j \odot \mathbf{h}_{SR}^i), \quad \forall i \neq j.
$$

(4.18)

If the relay gain vector $g$ is chosen in a way that $g^T \cdot \mathbf{H}_I = 0$, the interference between different source-destination links is nulled. Any vector $g$ that lies in the nullspace of $\mathbf{H}_I$ fulfils the above equality. The relay gain factors depend on the source and destination channels. They can be distributed in an initialisation phase and have to be updated occasionally depending on how fast the channels are changing.

In a MUZFR network, all relays have to be synchronous and need to know the channel state information of all other relays for both uplink and downlink.

The instantaneous sum rate of the $i$th SD link over two-hop at a given subcarrier is obtained
Figure 4.12: Outage probability vs. $d_{sd}$ for the MUZFR relaying scenario.

from:

$$R_{i}^{2\text{hops}} = \frac{1}{2} R_i = \frac{1}{2} \sum_{i=1}^{N} \log_2 (1 + \text{SNR}_i),$$  \hspace{1cm} (4.19)

where the instantaneous SNR per subcarrier for the $i$th SD link is given by:

$$\text{SNR}_i = \frac{P |H_{SD}[i,j]|^2}{\sigma_n^2 (1 + (g \odot h_{RD}^i)) h (g \odot h_{RD}^i)}.$$  \hspace{1cm} (4.20)

Similar to other relaying scenarios, the factor $\frac{1}{2}$ in (4.19) is introduced since for transmitting a packet from a source to a destination two channel uses of equal durations are assumed.

**Simulation Results:** We consider a MUZFR setup where four SD pairs of STAs communicate via 13 AF relays. Again 20 m distance between sources and relays is assumed, while distances between relays and destinations take different values. The maximum transmit power per subcarrier at each node and relay is set to $P$. However, depending on the gain coefficients, each relay may transmit with $P_r \leq P$.

For the numerical analysis, it is assumed that outage occurs when at least one of the concurrent transmissions faces outage. The outage probability and sum rate of MUZFR are depicted in Figure 4.12 and Figure 4.13, respectively. For comparisons, direct link and
AF relaying scenario with a single-antenna relay (SISO) are also considered. As it is seen in Figure 4.13, the spectral efficiency of MUZFR is significantly improved in comparison with the SISO AF relaying and direct link.

4.5 Summary

In this section three different multiuser relaying schemes are studied. The presented schemes are some examples of recently proposed relaying methods. Taking parameters of the IEEE 802.11n into account, it has been shown that these schemes have better spectral efficiency compared to the conventional relaying scenarios. However, to facilitate these schemes in WLANs, the actual CSMA/CA protocol has to be modified. In all of these schemes, more than one node should be able to transmit in a particular time slot. This is contrary to the existing CSMA/CA protocol which is originally designed for point-to-point communications. In the next chapter, we briefly outline the limitations of the existing MAC and then we propose a novel CSMA/CA-based MAC protocol which can be applied in the considered relaying scenarios as well as many other cooperative networks.
Chapter 5

A Novel Cooperative MAC Protocol

The IEEE 802.11n specifies MIMO techniques to enhance data rates in WLANs. However, using CSMA/CA it can support only point-to-point links. On the other hand, it is known that multiuser MIMO techniques significantly increase the spectral efficiency of networks. As it has been explained in Chapter 4, there are already enhanced MIMO signal processing techniques available which enable concurrent multiuser transmissions by utilising multiuser interference cancellation techniques. Multiuser interference can also be cancelled in a distributed manner. In Chapter 4, we have explained three different representative applications which exploit multiuser interference cancellation techniques.

In order to realise distributed spatial multiplexing gain in WLANs, the interference avoidance scheme should be modified to facilitate controlled interference. To enable multiple users to transmit simultaneously in IEEE 802.11n systems, we propose a novel cluster-based CSMA/CA (CB-CSMA/CA) protocol. According to this protocol, nodes in a network are grouped into multiple clusters in such a way that the nodes which belong to the same cluster can transmit or receive simultaneously.

The IEEE 802.11 standard does not define any wireless relaying method except for the AP. As shown in Section 3.2, conventional relaying methods may degrade the spectral efficiency and the delay performance of a given network.

In the following sections, we first outline the limitations of CSMA/CA-based MAC protocols. Then we introduce a novel CB-CSMA/CA protocol to solve the above-mentioned problems. In this chapter, in addition to describing the main aspects of the MAC layer, we elaborate the impact of interference mitigation and signal processing on the MAC protocol design. Basic concepts and the analysis of the MUZFR application have been partially published in [98].
Limitations of the Conventional CSMA/CA: In order to utilise the MUIC techniques in a CSMA/CA-based system, we need to modify the existing MAC protocol. The CSMA/CA has originally been designed for point-to-point links and avoids multiuser transmissions by a random exponential backoff procedure. In the current standard control frames including ACK are designed for point-to-point communication. Therefore, the control signalling should also be modified to support concurrent multiuser transmissions. Furthermore, as explained in the following sections, the frame format should also be adapted.

Again we consider a wireless network where STAs belong to the same BSS. Furthermore, for practical reasons, we consider only half-duplex nodes which are not able to transmit and receive simultaneously.

5.1 Cluster-Based CSMA/CA for Cooperative Wireless Networks

Assume a network where the destination nodes can efficiently decode the desired signal out of multiple streams. This is possible if there are enough degrees of freedom available at the destinations, e.g., each destination has enough antennas, or appropriate cooperation exists among nodes. In such a network, to enhance the spectral efficiency, multiple nodes should be able to transmit simultaneously. In order to do so, we classify the nodes in a network into clusters. The nodes belonging to the same cluster can access the channel at the same time. Similarly, nodes belonging to the same cluster may receive simultaneous streams provided that the intended stream at each destination can be decoded. Hence, from the MAC layer point of view, the clusters replace the individual nodes, and the nodes belonging to the same cluster look like a single node, i.e., a "power-station". Since multiple transmissions from a single cluster can be resolved, a collision happens only if more than one cluster transmits in the same time slot. Consequently, the spectral efficiency for the given system improves and at the same time the probability of collision in the network can be substantially reduced.

We consider three types of clusters: source-, destination- and relay-clusters. In a system with totally $V$ clusters, we denote the $v^{th}$ cluster by $C_v$, where $v \in \{1, 2, \ldots, V\}$. A cluster, which generates the data packets, is called the source-cluster and denoted by $C_{sv}$. The size of the source-cluster is limited by the maximum number of concurrent streams which can be efficiently decoded. A cluster which is receiving data is a destination-cluster and denoted by

\[^{1}\] As it is defined in [15, page 28], degrees of freedom of the channel is the dimension of the received signal space.
5.1 Cluster-Based CSMA/CA for Cooperative Wireless Networks

![Diagram of a wireless network with antennas and clusters]

**Figure 5.1:** An infrastructure network with four antennas at the AP and 8 source-destination pairs.

$C_{d_v}$. Even though forming a destination-cluster is not always necessary, for simplicity we assume that the destinations of a given source-cluster belong to a single destination-cluster. A relay-cluster, denoted by $C_{r_v}$, is a cluster of relays, which receives and forwards data to other clusters without having their own data packets.

Clusters can be formed based on application or network structure. For example in an infrastructure network with a multiple-antenna AP, the AP forms clusters of STAs such that it can mitigate multiuser interference in uplink and downlink transmissions. Figure 5.1 depicts a multiuser MIMO scenario where the AP has four antennas and each source- and destination-cluster consists of four single-antenna STAs.

In this chapter, we focus on the data transmission phase and assume that clusters are predefined. Clustering issues are studied in Section 5.1.6 and later in Chapter 7.

### 5.1.1 Basics of CB-CSMA/CA

In this section first we explain the general modifications of the DCF access method to support the CB-CSMA/CA. Afterwards advantages of the CB-CSMA/CA protocol in comparison with other proposals are discussed. In the next sections we study the impact of the PHY layer structure and signal processing on the CB-CSMA/CA and distinguish between infrastructure and ad hoc networks.
Chapter 5  A Novel Cooperative MAC Protocol

**General Modifications:** According to the CB-CSMA/CA, nodes in a cluster behave similar to a single power-station node, i.e., they access the channel at the same time and transmit simultaneously. Therefore, assuming the DCF access method is used, in addition to forming the clusters, we need to consider two major modifications to the current backoff procedure: (i) setting the same initial backoff duration for all cluster members and (ii) updating this value at the same time. The first requirement can be achieved for example by having the same random generator seed in each cluster, so that the same pseudo-random numbers are generated. To satisfy the second requirement, the stations within a cluster should be made *event-synchronous*. This is necessary since according to the IEEE 802.11, the contention window size doubles after each unsuccessful transmission up to a maximum value \[5\]. Thus, after a transmission, stations in a cluster should all know whether they need to increase the backoff interval or not, as explained in the following pages. In this way, each node in a cluster contends only with other clusters and not with other nodes in the same cluster. The assumption is that the clusters are constructed in such a way that destinations are able to decode the intended message. Consequently, simultaneous transmissions of the STAs within a cluster are resolved and do not lead to collisions.

Clusters access the channel one after the other, in a similar way as individual nodes in the current CSMA/CA systems. In this chapter we assume that STAs in each cluster are perfectly event-synchronised. However, as it will be shown in Chapter 8, the CB-CSMA/CA is robust to event synchronisation errors. In the worst case scenario, when all cluster members are not synchronous anymore, the CB-CSMA/CA is reduced to standard CSMA/CA.

In a SISO WLAN system, the header of each transmitted frame can be ideally decoded by any node which is in the communication range of the transmitter. Accordingly nodes within a WLAN learn about an ongoing transmission by decoding the headers which include the duration field. However, the situation changes in a CB-CSMA/CA system, where several frames might be transmitted in parallel. The system is structured in such a way that each destination is able to decode its intended packet’s header as well as its payload. However, other nodes in the network may not be able to do so. In order to solve this problem, a cluster common preamble should be sent prior to the data transmission. The common preamble, which is sent at the lowest rate, could have a similar format as that of the legacy preamble of the PLCP header used in the IEEE 802.11n mixed format [1], see Figure 2.3. It additionally includes the address of the source-cluster and its LENGTH field should be set to the maximum length of all concurrent packets. In this way, the common preamble can be decoded by

---

1. It consists of the short and the long training sequences and a SIGNAL field. The SIGNAL field includes a RATE and a LENGTH field, one reserved bit, a parity bit and 6 tail bits and is one OFDM symbol long.

108
5.1 Cluster-Based CSMA/CA for Cooperative Wireless Networks

all STAs in the network.

Two-Hop Links: Since, the IEEE 802.11 standard does not define any wireless relaying method, for two-hop relay links, our CB-CSMA/CA protocol adopts an approach which is similar to [23]: relay clusters forward the received data packets after a SIFS, without participating in another contention procedure prior to the second-hop transmission. In this way, delay is reduced and the uplink and downlink utilisation is balanced.

CB-CSMA/CA Advantages: There are different research papers which deal with MAC layer limitations of supporting multiuser transmissions. A MAC protocol with antenna arrays is suggested for ad hoc networks in [28]. The antenna array is used to null the interference from neighbours. In this work as well as in [29], a control channel has been introduced in a way that nodes can learn about ongoing transmissions by monitoring this channel. In [29] once a transmission request is accepted, a busy tone on a control channel is used to acknowledge the request and to prevent other nodes to transmit at the same time. To let more than one node to transmit, the number of control channels should be equal to the maximum number of nodes which can transmit concurrently in the network.

Park et al. [30] have proposed a MAC protocol mitigating interference by using multiple antennas at the receivers. The proposed MAC has been designed for a multihop network with multiple-antenna nodes. Each transmission interval is divided into a negotiation period and a contention-free period. During the negotiation period STAs contend to access the channel based on the RTS/CTS access method. However, in their considered scenario there are two contention slots available prior to the contention-free period, and hence two RTS/CTS frames can be exchanged. The contention-free period includes the required training sequences and a data slot for simultaneous transmissions of data packets.

In [31], the authors propose a MAC protocol which considers the spatial correlation between the signal and interference to decide whether multiple links should transmit simultaneously or not. In their proposal, links are allowed to contend for the channel sequentially while transmitting data packets simultaneously. In [32], authors have proposed a slightly modified version of the existing MAC to support multiple streams. In their scheme whenever more than one user transmits in a given slot, if the receiver can decode multiple packets it simply acknowledges them.

In addition to works which have considered multiuser transmissions there are some works which deal with resolving collision. In [99] the authors came up with a MAC protocol with the ability to resolve collisions. According to this protocol, when a collision occurs, the collided packets have to be buffered. In the following time slots, a set of nodes behave as non-regenerative relays and forward the signal they have received during the collision slot. This
method has been extended and applied to a large-scale wireless sensor network by dividing the network into several clusters \([100, 101]\). Each cluster has a clusterhead, and the nodes in a cluster only communicate with the clusterhead. Collisions within a cluster are resolved as in \([99]\), while the clusterhead acts like a base station. The initial work in \([100]\) which was developed for a flat fading channel has been extended in \([101]\) to frequency selective channels.

Contrary to \([99]\) and \([100]\), in which the authors focus on the diversity techniques, we focus on utilising the spatial multiplexing gain in CSMA/CA-based systems.

The CB-CSMA/CA scheme has distinct advantages compared to other existing proposals: it requires neither sequential contention per node for the data transmission as in \([30]\) and \([31]\), nor a control channel as in \([29]\), \([28]\) and \([99]\).

Additionally, by grouping nodes into clusters, the collision probability of the CB-CSMA/CA is reduced as collision probability depends on the number of contending clusters rather than total number of contending stations. The conditional collision probability of CB-CSMA/CA versus the number of STAs in the network, for different numbers of STAs per cluster is shown in Figure 5.2. For comparison the conditional collision probability of conventional CSMA/CA is also depicted. As it is shown in Figure 5.2, the CB-CSMA/CA conditional collision probability is considerably reduced as compared to the stan-

**Figure 5.2:** Conditional collision probability vs. number of STAs in the network for CB-CSMA/CA with different number of STAs per cluster \((N_{nc})\).
5.1 Cluster-Based CSMA/CA for Cooperative Wireless Networks

dard IEEE 802.11 MAC, even for a cluster size of only two STAs. For a certain number of active STAs, larger source-cluster size results in a lower collision probability.

As it is explained in the next section, one advantage of the CB-CSMA/CA is that it is able to post-detect collisions in high SNR regime. Furthermore, in contrast to [32] where there are multipacket transmissions, only if accidentally more than one STA transmits in a time slot, the CB-CSMA/CA is designed such that each transmission attempt contains as many data packets as possible.

5.1.2 Collision Post-Detection

In WLANs packets may fail due to collision and channel error. The cause of failure however, cannot be distinguished in current WLAN systems [74, 75]. Hence, in both cases, the backoff interval increases exponentially up to a maximum value which is defined by the standard [5]. As it has been briefly mentioned in Section 3.1, in the presence of channel error the binary exponential backoff may reduce efficiency and cause unfairness.

An inherent advantage of our proposed CB-CSMA/CA is, its ability to differentiate between these losses in practical SNR regimes. We assume that all nodes in a BSS are in radio range of each other. At each transmission attempt a common preamble is transmitted by all members of the source-cluster. The common preamble is short and transmitted at the lowest data rate. Therefore, it is assumed that it is error-free. However, as different clusters send different common preambles, the common preambles collide if more than one cluster transmits in a time slot. Accordingly, upon reception of the common preamble the collision can be post-detected as explained in the following paragraph.

If the common preamble cannot be decoded, it can be assumed that a collision has occurred. However, if the common preambles collide, the information about involved clusters may not be obtained. Hence, a short packet, called contention window update request (CWUR) packet, should be broadcasted to all clusters, e.g. by the AP. Upon reception of CWUR all source-clusters enter the next backoff process. However, only the source-clusters which have been involved in the pending transmission increase their CW unless the maximum CW value has already been reached. In this case they keep the maximum CW for the upcoming transmission.
5.1.3 Impact of the PHY layer on the CB-CSMA/CA

The CB-CSMA/CA can be applied to networks with different signal processing methods. However, some of the PHY layer characteristics affect the MAC operation and should be considered in the course of the MAC protocol design. In this subsection, we study the PHY characteristics which impact the frame structure, backoff and ACK procedures and provide details on how the CB-CSMA/CA scheme operates.

**Point-to-Multipoint and Multipoint-to-Point:** Assume a network where a multiple-antenna node performs the multiuser interference cancellation. In this case, the multiple-antenna node can ensure the collision post-detection and CW update.

If the multiple-antenna node is the receiver, it can detect the collision upon reception of the common preambles. In case of collision, the multiple-antenna receiver informs individual transmitters to update their backoff interval. This can easily be done by sending the CWUR packet to all clusters after a SIFS period, following the packet reception. Upon reception of a CWUR, only STAs which are expecting an ACK but instead receive the CWUR packet, update their backoff interval to the next larger value. However, all nodes which have a packet ready to transmit, enter the next backoff process. On the other hand, if the multiple-antenna node is the transmitter, then it can detect collisions, by checking the number of received ACK frames. If collisions among clusters happen, most of the STAs will be affected. Therefore, in practical SNR regime, whenever the multiple-antenna transmitter does not receive any ACK, it assumes that with high probability a collision has happened.

An infrastructure network with a multiple-antenna AP, which can handle the transmissions from/to several STAs concurrently, is a typical example of such MIMO networks. The AP can easily form clusters by taking into account the PHY capability of the associated STAs, e.g., the number of antennas. This scenario is investigated as application A in Section 5.2. Figure 5.3 shows the uplink transmission in such a network, where the AP has two antennas and STA1 and STA2 belong to cluster $C_{s1}$ while STA3 is a member of cluster $C_{s2}$. In Figure 5.3(a) an example of a successful transmission is shown, where STAs in $C_{s1}$ have a shorter backoff than STAs in $C_{s2}$, hence they transmit first. In Figure 5.3(b) both clusters begin to transmit at the same time and hence a collision occurs. In this figure subscript $c_i$ above the data indicates the cluster from which the data packet is originated while subscript $j$ below the data or ACK specifies the respective STA.

**Cooperative MUIC:** In this thesis we consider realisation of the cooperative MUIC in an ad hoc network rather than an infrastructure network. Consequently there is no central node (AP) and decisions about updating the contention window should also be handled in
5.1 Cluster-Based CSMA/CA for Cooperative Wireless Networks

Figure 5.3: CB-CSMA/CA access mechanism in an infrastructure network with a multiple-antenna AP as the receiver. The common preamble is included in the data frame.
a distributed manner. However, a totally distributed exchange of packets implies lots of overhead. Therefore, in ad hoc scenarios, we assume that each cluster has a cluster-master. Additionally, it is assumed that in the ad hoc mode, destinations send back the ACK frames to the sources sequentially. According to the CB-CSMA/CA protocol, the nodes which need to jointly set parameters or decode, are allocated to the same cluster.

In ad hoc networks, one of the cluster-masters is responsible for detecting the source of transmission failure. Similar to the infrastructure case, the collision can be post-detected by the cluster-master, upon reception of the common preamble\(^3\).

The MUZFR network, where relays orthogonalise SD channels and allow concurrent communication among several SD pairs [27], is an example of such networks. We consider this scenario as application B in Section 5.2 and explain how CB-CSMA/CA can be utilised in such a scenario.

**DF Relaying:** A DF relay decodes a packet before forwarding it. As a result, it is able to check whether the packet is received correctly or not. For example, consider a DF multiple-antenna relay which receives several packets from a source-cluster in parallel. The relay forwards only those packets which are received correctly. If the common preambles are successfully decoded but none of the data packets is received correctly by the relay, the relay terminates the transmission by sending a short notification packet to all STAs in the source-cluster. Consequently STAs are noticed that a channel error has occurred during the first hop transmissions and they can update their backoff interval depending on the applied backoff model. It should be noted that according to the CB-CSMA/CA, in a two-hop link, the packets are forwarded by the relay after a SIFS. Hence, collisions may happen only during the first hop. However, channel errors can occur during both uplink or downlink transmissions.

**AF Relaying:** In general AF relays are less complex as compared to DF relays. However, AF relays cannot detect transmission errors and forward the received packets in all cases. In this way, the duration of an erroneous transmission is not limited to a single-hop rather it increases with the number of hops. Besides, there are some standard compatibility issues in case of AF relays which should be taken into account. Here we have assumed that the AF relays are able to decode the common preambles\(^4\).

**Frame Format and Channel Estimation:** As explained in Section 5.1.1 a common

---

\(^3\)If some STAs in a network cannot hear each other, then each cluster-master is responsible for post-detecting collisions and transmitting the CWUR to all members in its own cluster. In this case, in order to avoid interference among CWUR packets, the CWUR packets in each cluster should be transmitted on non-interfering (orthogonal) channels.

\(^4\)We need to consider some assumptions on the decoding capability of AF relays as explained in Section 5.3.2. The implementation of AF relays in a WLAN network needs further considerations.
5.1 Cluster-Based CSMA/CA for Cooperative Wireless Networks

The common preamble is transmitted prior to data transmission by all members of the source-cluster. After estimating CSI from the training sequences included in the common preamble, the SIGNAL part of the common preamble can be decoded by the individual nodes.

In the current WLAN, the receivers estimate the channels by receiving the training sequence in the PLCP preamble of each frame. In CB-CSMA/CA, each transmission attempt may consist of several simultaneous streams. Therefore, orthogonal training sequences are allocated to different STAs within a source-cluster. The orthogonal training sequences can be designed as proposed in the IEEE 802.11n for MIMO transmissions, i.e. for each transmit chain, the training field is cyclically shifted \[1\].

5.1.4 Impact of Control Frame Error

CWUR is the new control frame which has been introduced in CB-CSMA/CA. As explained before, STAs which are waiting for an ACK but instead receive CWUR, should update their backoff window. Upon reception of the CWUR, the backoff window is set to the next larger value if it has not yet exceeded its maximum size. Here we discuss the sources of the transmission or reception failure of a CWUR packet. Furthermore, we discuss the impact of this failure on the CB-CSMA/CA performance. In this section we deal with the infrastructure mode however extensions to ad hoc mode can be done in a similar way. In the infrastructure network the AP is responsible for transmitting the CWUR packets. First we describe the error factors in the transmission side, and then in the reception side.

Transmission Side: The AP transmits a CUWR upon detection of a collision. Hence, if the AP mis-detects collisions or falsely detects some of them, the CWUR is, not transmitted or transmitted by mistake, respectively.

The false-detection of collisions happens if no collision has been occurred but the AP is not able to decode the parallel preambles. This may happen when signal to noise and interference ratio (SNIR) of all links from a given source-cluster are too small. In this case CWUR is transmitted and STAs update their CW although no collision has occurred. It is expected that in practical SNR regimes, this does not happen often, since the common preamble is transmitted at the lowest rate using the most robust modulation.

The mis-detection of collisions happens when collision has occurred but the AP cannot detect it. This happens when the AP is able to capture at least one of the common preambles for example because of near-far problem\[5\].

\[5\] The near-far problem may occur if the received packets from different STAs with different power. In such cases, the strongest received signal may capture the decoding process at the destination [102, page 548].
As no CWUR will be transmitted in this case, the STAs involved do not update their CW. In case the number of clusters are not high, the chance that at the next transmission attempt no collision occurs is substantial. However, for large networks if mis-detection happens frequently, the collision avoidance procedure is not efficient anymore. To avoid this situation we can apply standard backoff procedure, for a large number of contending clusters. Consequently, even in case of channel errors the CW size is updated.

As it will be shown in Section 7.1, for medium and large number of clusters the standard backoff model outperforms the backoff model used in this chapter. In this case, to keep all STAs in a cluster event-synchronised, a control frame should be sent to STAs whenever transmission in one or more concurrent links has failed (regardless of the source of errors). It should be noted that, as long as an error occurs on the control frame, only at the transmission side, all receivers act in the same way and hence remain event-synchronised.

**Reception Side:** In this case we assume that the CWUR is correctly transmitted, but in one or more receivers the reception has failed due to poor downlink channels or strong interference. If all STAs within a cluster cannot decode the CWUR, then they all act in the same way. However, as soon as some STAs are able to decode the CWUR and some are not, they begin to become event-asynchronous. Therefore, depending on the new CW value, in their next transmission attempt they may begin to transmit individually. We will analyse the performance of CB-CSMA/CA in the presence of event-synchronisation errors in Section 8.1. As it will be shown, in the worst case scenario where all STAs in all clusters are asynchronous, the CB-CSMA/CA throughput gets close to that of the standard IEEE 802.11.

**Event-Synchronisation Error:** We denote all errors which cause different backoff window values at nodes within a same cluster, as event-synchronisation error. As mentioned, errors on reception of CWUR packets may result in an event-asynchronised cluster, i.e., a cluster which has one or more event-asynchronous STA(s).

Event-synchronisation error can also occur, when some cluster members sense the channel idle while other members of the same cluster sense it busy. Accordingly, the first group of members continue counting down their contention window while the second group of members freeze it. This situation can happen when some nodes are hidden from some cluster members but not from the others. Another situation which can lead to event-synchronisation error is the time synchronisation error. If the nodes in a cluster are not perfectly time synchronised some of them may begin to transmit before the others. The clusters should become event-synchronised occasionally. This can be done by the AP in infrastructure networks or by cluster-masters in ad hoc networks.

In this thesis, we assume perfect time synchronisation and a BSS where all nodes can hear
5.1 Cluster-Based CSMA/CA for Cooperative Wireless Networks

Each other (no hidden terminal). Moreover, in chapters 5 and 6 we assume a perfect CWUR transmission and perfect event-synchronisation. Time and frequency synchronisation issues are discussed in Section 5.1.7. We will study the impact of event-synchronisation error in details in Chapter 8.

5.1.5 Cluster-Based CSMA/CA with RTS/CTS

Throughout this thesis we consider the basic channel access method without exchanging RTS/CTS frames. The extension to the RTS/CTS case can be simply done as is is explained in this section. We distinguish between two modes: infrastructure mode and ad hoc mode.

Infrastructure Mode: In this case, STAs in the cluster, which has won the contention procedure, send RTS frames to the AP. If the multiple-antenna AP can decode the multiple RTS frames, it replies by multicasting a CTS frame to all STAs in the respective source-cluster. In this case, in order to inform other STAs in the BSS about the upcoming transmission, the cluster common preamble should be attached to the beginning of the RTS frame.

In [32] the authors have proposed to use RTS/CTS frames to support multipacket transmissions. In their proposed method, if accidentally more than one RTS is transmitted in a time slot and the AP can decode them, it sends a CTS frame to the respective transmitters. We will compare the performance of the CB-CSMA/CA protocol with their scheme in Chapter 8.

Ad Hoc Mode: In the ad hoc mode, the cluster master of the source-cluster which has access to the channel, transmits the RTS frame to the master of the destination/relay cluster. If the master of the immediate receiver cluster can decode the RTS frame, it multicasts a CTS frame to all STAs in the source-cluster. In this way, STAs know whether and when they should start transmission of the data packet. The common preamble packet should be attached to the RTS frame. STAs in the source-cluster should remain silent while the master transmits these control frames until they receive the CTS frame.

5.1.6 Cluster-Based Centralised Channel Access

The centralised coordination only applies to infrastructure networks, where a central unit, i.e., hybrid coordinator (HC), schedules STAs using a polling mechanism. Compared to the distributed case, the coordination is easier since the HC fully controls the channel access. The HC can set up the clusters, by simply polling each associated STA and asking it about its upcoming transmission.
The HC can adaptively change clusters based on the information obtained during the contention-free period. According to the IEEE 802.11e [39], STAs which do not have any queued packet to send, reply the CF-Poll by a Null frame. As a result, the HC (here the AP) can form clusters by taking only STAs with queued packets into account. As each CFP is followed by a CP, the clusters can be formed during CFP and can remain the same during the CP.

Once clusters have been formed, the HC should poll clusters for data transmission. The polling for the data transmission is performed in a similar way to the 802.11e HCCA. The main difference is that here, the HC polls one cluster after the other, instead of polling a certain node at each time instant, as it is defined in the standard. Consequently, the polling message should be multicast to all STAs belonging to the same cluster. All nodes, which have received the polling message and have data packets ready for transmission, start transmission in the following time slot concurrently. All other nodes in the network shall set the duration of their network allocation vector to the longest transmission duration [39]. Since the HC is not able to transmit and receive concurrently, it should belong to a single-member cluster.

5.1.7 Time and Frequency Synchronisation

In multiuser OFDM systems, where multiple STAs transmit/receive at the same time, the received signals should be aligned in time and frequency. In other words, proper time and frequency synchronisation at the PHY layer is required at all transmitters and receivers.

In OFDM systems as long as the sum of the maximum delay offset and the channel delay spread is shorter than the cyclic prefix, both the multiuser interference (MUI) and the inter-symbol interference (ISI) can be mitigated. The IEEE 802.11n specifies two possible values for the guard interval, the short one is $0.4 \mu s$ and the other is $0.8 \mu s$. Hence, taking into account the IEEE 802.11n parameters, the maximum delay offset should not exceed $0.8 \mu s$. Otherwise, the received packets cannot be decoded and a longer GI is necessary.

In indoor scenarios the maximum channel delay spread and distances between nodes are quite small. In this case, from the PHY layer aspect, in order to enable several users to transmit at the same time, it is enough to have an accurate synchronisation among STAs involved. Assuming an indoor scenario where the maximum distance-difference is around 40 m and the maximum channel delay spread is about 400 ns, then a GI of 800 ns leaves only small space for time synchronisation offset. However, some recent works have proposed

---

6The network allocation vector is maintained by each STA. It is an indicator of time periods when the STA will not initiate transmission, regardless of the sensed channel status (busy or idle).
techniques which can achieve synchronisation accuracy in the range of sub-microseconds among different nodes [103] [104].

In [103] IEEE 1588 synchronisation has been implemented over IEEE 802.11b. IEEE 1588 standard is defined for synchronising independent clocks running on individual nodes of a distributed system. It is a master-slave synchronisation protocol which provides higher accuracy. To achieve high accuracy, several control messages are exchanged periodically between a master and its slave in a way that the slave can calculate the offset of its clock with respect to the master. The authors evaluate the jitter introduced by the PHY layer of IEEE 802.11b and showed that the degradation of the synchronisation accuracy due to this jitter is less than a few hundred nanoseconds. In [104] the authors focused on localisation methods based on WLANs and describe common features between the localisation and precision clock synchronisation problems. They suggest a few modifications of current WLAN PHY products such as an additional counter to provide high synchronisation accuracy.

It should be noted that, in CB-CSMA/CA such an accuracy is required only among members of the same cluster. Besides such an accuracy is not required when AF relays are used to orthogonalise multiuser channels as the case in Application B in Section 5.3.2.

Assuming that the PHY layer is able to handle multiple streams efficiently, then a slot-based synchronisation is enough to utilise the CB-CSMA/CA protocol. According to the IEEE 802.11n the slot time is $9 \mu s$ long. In a WLAN system STAs are synchronised by a timing synchronisation function (TSF) with a resolution of one microsecond [5].

If for multiuser MIMO transmission, we use the same format for training fields as the one defined by the IEEE 802.11n, cyclic shift values in the range of several hundreds nanosecond are required. On the other hand, if there is not a time synchronisation method which achieves this accuracy, one should either apply longer cyclic shifts or for example use orthogonalisation in frequency domain instead of time domain. In the first case, longer cyclic shifts lead to larger delay spread and hence a longer GI is needed. In the second case, the goal can be achieved by allocating a subset of subcarriers to each user’s training filed. In the next chapter, we show that the CB-CSMA/CA enjoys a much higher throughput as compared to the IEEE 802.11n system even if it has a GI twice as long as in the IEEE 802.11n system.

In multiuser MIMO systems, beside the time synchronisation, frequency-synchronisation is necessary. In fact, for OFDM systems an accurate frequency-synchronisation is even more crucial. Frequency shifts of subcarriers may lead to non-orthogonal subcarriers and can degrade the system performance drastically. In multiuser MIMO systems, each user has a local oscillator and there can be a frequency offset between different local oscillators.
Chapter 5 A Novel Cooperative MAC Protocol

Hence, it is important to find a way to estimate these offsets and when possible track the residual offset during the data transmission phase at the receiver [105] [106].

In general the synchronisation in the downlink (point-to-multipoint) is easier than that in the uplink (multipoint-to-point). In the first case, each link can be considered individually while in the latter case, the simultaneously transmitted signals from multiple users should be aligned in time and frequency. Furthermore, synchronisation in infrastructure networks can be established more easily than in ad hoc case.

In the networks, where an accurate time synchronisation among STAs within a cluster cannot be achieved, in order to avoid event-synchronisation problems we take an alternative MAC protocol. Since this situation is more likely to happen in ad hoc networks, in the following paragraph we explain how this protocol can be applied to an ad hoc network:

According to this protocol only cluster-masters contend for the channel. However, once a cluster-master acquires the channel, it requests all members in its cluster to transmit concurrently. This can be done by sending a short packet to the members prior to data transmissions. In this way, STAs in a cluster do not need to be event-synchronous anymore. In the next chapters we investigate throughput performance of the original CB-CSMA/CA. However, throughput results of the alternative protocol will be similar to that of the original CB-CSMA/CA. This is due to the same cluster size, the same number of contending units and the same PHY layer requirements in both cases.

In this thesis we assume that all STAs are perfectly synchronised in time and frequency. In infrastructure mode the AP acts as the timing master and performs the TSF by sending the periodic beacons. Similarly, in the ad hoc mode the cluster-masters perform the TSF among their members. In this case, synchronisation between different clusters can be achieved in a way similar to that in the IEEE 802.11 ad hoc mode. Each cluster-master should maintain a TST timer and sends a beacon according to a defined beacon period parameter. Each cluster-master should adopt its time to a received TSF value which is later than its own TSF time.

It should be noted that the problem with time and frequency offset is not a specific problem of the proposed CB-CSMA/CA but a general problem in all multiuser MIMO OFDM systems. An overview of multiuser MIMO-OFDM can be found in [107]. Synchronisation techniques for multiuser MIMO-OFDM systems can be found in [105] [106] [108]. In [106] a blind carrier frequency offset estimator for OFDM systems have been proposed. The proposed estimator is a kurtosis-based estimator. It has been shown that the estimator can be applied in MIMO and multiuser OFDM systems. In [108] the authors consider both carrier frequency offset and sampling frequency offset in the uplink of multiuser OFDM systems.
A maximum-likelihood estimator has been proposed to estimate both carrier frequency and sampling frequency offsets and to track the phase. The proposed method can be applied in pilot-based as well as decision-directed methods. Bit error rate simulations have shown that if the proposed estimation is used the degradation due to frequency mismatches is negligible.

## 5.2 Outline of Representative CB-CSMA/CA Applications

The proposed CB-CSMA/CA has many applications in different cooperative wireless networks. Appropriate applications include cooperative scenarios like MUZFR [27], and two-way relaying [26], which require several STAs to transmit simultaneously and are not supported by the current WLAN MAC. Additionally, any point-to-point multiuser WLAN scenario that is supported by the standard MAC, is more efficient under the CB-CSMA/CA if multiuser interference can be cancelled.

Here we consider three representative applications to quantify the performance of the CB-CSMA/CA scheme. Taking into account the three relaying scenarios from Chapter 4, we investigate the MAC throughput and delay. All of these applications use two-hop traffic pattern, however the CB-CSMA/CA scheme can be applied to single-hop scenario as well. We consider a symmetric case, in which the number of sources and destinations are the same. In these applications we assume bidirectional communication between each source and its respective destination. Hence, a source- and its destination-cluster, interchange their functions at different time slots.

- **Application A: MIMO DF Relaying**

  The first application is carried out in an infrastructure network, where the AP acts as a wireless DF relay for STAs within its BSS. For each pair of nodes, there is a source-destination link via the AP, which establishes a wireless two-hop link\(^7\). The AP has \(N_a\) antennas while each STA is equipped with a single antenna. Therefore, at each transmission attempt \(N_a\) STAs will be able to transmit in parallel. In order to do so, sources and destinations are grouped into clusters and operate based on CB-CSMA/CA.

  **Reference System:** As reference system, we use the above-mentioned infrastructure network which operates based on the IEEE 802.11 DCF. In this reference system, at each time instant at most one node is allowed to access the channel. While our

\(^7\)We neglect the STAs which have packets for destinations outside of their BSS. If such STAs exist, they are also grouped into clusters but they establish wireless single-hop links.
Chapter 5  A Novel Cooperative MAC Protocol

reference system comprises intra-BSS transmissions in an infrastructure network, most existing works consider either ad hoc mode or only uplink and downlink transmissions in the infrastructure mode. Additionally, we consider the IEEE 802.11e DLS [39] where the AP initiates a direct link between two STAs.

- Application B: MUZFR

In application B, an ad hoc network is assumed, where no powerful central node exists which can cancel the multiuser interference. Since this scheme is not supported by the current standard, there is no equivalent reference system based on the IEEE 802.11 MAC. However, the MUZFR scenario can be supported by the CB-CSMA/CA ad hoc mode as described in Section 5.3.2.

- Application C: Two-way DF Relaying

In application C, an infrastructure network is considered where the AP acts as the two-way relay. As mentioned in Section 4.3, the two-way relaying scheme is not supported by the existing standard but it can be utilised when the CB-CSMA/CA is applied. Furthermore, the semi-two-way relaying as defined in Section 4.3.3 can be supported by conventional CSMA/CA protocol.

5.3 CB-CSMA/CA Throughput and Delay Investigation

In this chapter we assume a perfectly event-synchronised CB-CSMA/CA. Hence, we can apply the original Markov Chain model which has been explained in Chapter 3. However, from the MAC layer point of view, clusters replace individual STAs. Consequently, the total number of contending units should be set to the total number of contending clusters.

5.3.1 Application A

In this application, the AP defines the clusters. Since the AP has $N_a$ antennas, each cluster can at most include $N_{uc} = N_a$ STAs. Therefore, there are $N_c = 2\left\lceil \frac{N_a}{N_{uc}} \right\rceil$ source- and destination-clusters, while the AP itself constitutes a single-member cluster, $C_r$.

Clusters are scheduled according to the contention procedure. Figure 5.1 depicts this scenario for $N_{sd} = 8$ SD pairs, and $N_a = 4$. It should be noted that in contrast to this figure, adjacent STAs need not necessarily be within the same cluster. It is assumed that the AP has full CSI of the uplink and downlink channels, while STAs only have the CSI at the receiver side.
5.3 CB-CSMA/CA Throughput and Delay Investigation

**Throughput Investigation:** We start by stating the relevant formulae and then proceed by outlining the modifications, which are required by the CB-CSMA/CA protocol: for a fixed number of contending clusters, the probability that a cluster transmits in a randomly chosen time slot, is given by (3.9). The Markov chain model is similar to that of the standard case, cf. Figure 3.4. As it is mentioned in Section 3.1.4, for conventional CSMA/CA, the transitional probability from a backoff stage to the next one ($p$), is determined by the conditional collision probability $P_{\text{col}}$ and the probability of channel errors $P_e$. However, due to CB-CSMA/CA capability of distinguishing the source of packet loss, the channel errors do not influence the backoff duration and hence, for the systems based on the CB-CSMA/CA we can set $p$ in (3.9) according to the following equation:

$$p = P_{\text{col}} = 1 - (1 - \tau)^{n-1}. \quad (5.1)$$

It should be noted that contrary to the reference system, where $n$ is equal to the number of contending nodes, in a system based on the CB-CSMA/CA $n$ is reduced to the number of contending clusters, i.e. $N_c$. Taking this modification into account, the probability that there is at least one cluster that transmits in a slot, $P_{tr}$, and the probability that only one cluster transmits in a slot can respectively obtained from (3.12) and (3.13) by replacing $n$ with $N_c$. Assuming the same average channel conditions, data rates, and average payload size for all STAs, the throughput can be calculated similar to (3.14) according to the following formula:

$$S = \frac{P_s P_{tr} (1 - P_e) \tilde{L}}{T_{\text{slot}}}, \quad (5.2)$$

where $\tilde{L}$ is the average payload size transmitted by a cluster. Assuming cases where all clusters have $N_{nc}$ STAs, we have $\tilde{L} = L_c = N_{nc} \cdot L$ with $L$ being the payload size. $T_{\text{slot}}$ can be calculated from (3.15).

In the considered CB-CSMA/CA application, the AP forwards packets after a SIFS period. Thus collisions may occur only during the uplink. However, both downlink and uplink transmissions may fail due to link errors. In case of collision, the AP sends the CWUR packet to the clusters without forwarding packets to the final destinations.

Destinations confirm the successfully received packets by sending an ACK back to the AP simultaneously. According to CB-CSMA/CA the retransmissions, if required, are handled by the sources. Thus, ACK frames are forwarded by the AP back to the sources. Consequently, as it is shown in Figure 5.4, for CB-CSMA/CA application A, the average time required for
Figure 5.4: CB-CSMA/CA access mechanism in application A.
5.3 CB-CSMA/CA Throughput and Delay Investigation

Successful data transmission ($T_s$), collision ($T_c$) and channel error ($T_e$) are defined as follows:

\[ T_s = T_{\text{Data}}^1 + \text{SIFS} + \delta + T_{\text{Data}}^2 + \text{SIFS} + \delta + T_{\text{ACK}}^2 + \text{SIFS} + \delta + T_{\text{ACK}}^1 + \delta + \text{DIFS}, \]
\[ T_c = T_{\text{Data}}^1 + \delta + \text{SIFS} + T_{\text{CWUR}} + \delta + \text{DIFS}, \]
\[ T_e = \begin{cases} T_{\text{Data}}^1 + \delta + \text{SIFS} + T_{\text{Notification}} + \delta + \text{DIFS}, & \text{if all uplinks fail due to channel error} \\ T_{\text{Data}}^1 + \delta + T_{\text{Data}}^2 + \delta + T_{\text{ACK}}^A_{\text{timeout}} + \text{DIFS}, & \text{any other case} \end{cases} \]

where $T_{\text{Data}}^i$ and $T_{\text{ACK}}^A$ are the duration of the data (including the common preamble) and the ACK transmission at the $i$th-hop, respectively, and $T_{\text{CWUR}}$ is the duration of the CWUR frame. Since in practical SNR regime, the probability that all uplink transmissions fail due to the channel error is relatively small, for our numerical analysis, we set $T_e$ to:

\[ T_{\text{Data}}^1 + \delta + T_{\text{Data}}^2 + \delta + T_{\text{ACK}}^A_{\text{timeout}} + \text{DIFS}. \]

It is assumed that both hops have the same data rate, and hence, the same $T_{\text{Data}}$ and $T_{\text{ACK}}$. The ACK timeout for application A is:

\[ T_{\text{ACK}}^A_{\text{timeout}} = \text{SIFS} + T_{\text{ACK}}^2 + \delta + \text{SIFS} + T_{\text{ACK}}^1 + \delta + \sigma. \]

The duration of data, ACK and CWUR frames can be obtained from (6.8)(6.9) and (6.10), respectively.

In application A the uplink and downlink utilisation is equal and each destination on average enjoys the following throughput:

\[ S_D = \frac{S}{N_c \cdot N_{nc}}. \tag{5.3} \]

**Delay Investigation:** In this section we take into consideration the delay introduced by the MAC protocol. As described in [61], the average access delay $D$ can be calculated according to the following equation:

\[ D = \frac{N_c}{S/L}, \tag{5.4} \]

Where $\bar{L}$ should be set to $L_c$. Since in this application the AP does not introduce any queuing delay, the estimated packet delay represents the head-of-line (HoL) delay.

**Reference System:**

Throughput of the reference system per STA can be obtained from (3.14)-(3.18) and (3.41).
In the reference system, a packet on its way from a source to a destination has to additionally wait in the queue of the AP. Since, in the saturation mode the queueing delay at the AP is not bounded, we consider only the HoL delay of the uplink and the DLS scenario. For the DLS case we can apply the equations related to single hop networks in Section 3.1.4.

### 5.3.2 Application B

In this application, it is assumed that one of the relays, e.g., $r_i$, knows the number of available AF relays and it defines all clusters. It allocates all relays to a single-relay cluster $C_r$ and acts as the cluster-master of $C_r$. It groups sources and destinations into clusters such that, in a system with $N_r$ relays, each cluster has at most $N_{nc} = \left\lfloor \frac{1 + \sqrt{4N_r - 3}}{2} \right\rfloor$ stations, cf. (4.16).

It is assumed that the AF relays are simple nodes, which are equipped with a single antenna and do not generate their own data packets. Consequently, they do not need to participate in the contention procedure. The AF relays only scale and forward the received data packets to other STAs, without decoding them. However, we assume that they are able to decode an identification message and acknowledge them (as explained in the following paragraphs) and they are all synchronised to the timing synchronisation function [5].

The relay gain factors depend on the source and destination pairs. These values are calculated so that there is no interference between the data streams of different SD pairs when a specific cluster is transmitting. The relay gain factors depend on the source and destination channels. The CSI can be estimated locally at each relay or destination and disseminated to all relays in the initialisation phase. They have to be updated occasionally depending on how fast the channels are changing.

According to the CB-CSMA/CA, the source-cluster which has won the contention, acquires the channel. First, the cluster-master of that source-cluster sends an identification message. The identification message consists of the legacy preamble (as used in the common preamble in application A) but it is sent prior to the data transmission. It additionally includes a MAC header similar to that in the RTS frame but instead of the MAC addresses of the STAs, it includes the MAC addresses of the source- and the destination-cluster. Consequently, upon reception of the identification message the relays know which gain factors to apply.

Assuming an error-free identification message, if $r_i$ cannot decode the identification message it assumes that a collision has occurred. However, similar to the previous case, in this case $r_i$ cannot obtain the information about involved clusters. Hence it broadcasts the
5.3 CB-CSMA/CA Throughput and Delay Investigation

CWUR packet to all clusters. Again only the source-clusters which have been involved in the pending transmission update their CW size.

If the cluster-master can decode the identification message it acknowledges the reception by multicasting an acknowledgement packet (ACKid) to all members of the source-cluster. Then the STAs in the source-cluster send the data packets to all AF relays. The relays scale the received packets with the respective gain factors and forward them to the destination cluster. The destinations which successfully decode the packets, transmit the ACK packets one after the other, for example, in a pre-determined order.

In the case of channel errors, the sources which do not receive the ACK, retransmit their data packet in the next transmission attempt. Since AF relays do not decode the data and cannot detect channel errors, the channel error can only be detected after the second-hop transmission.

**Throughput Investigation:** In order to calculate the throughput of application B, similar to application A, we can apply (3.9)-(3.15) by setting \( n \) to the number of contending clusters in the network and replacing \( L \) by \( L_c \).

Figure 5.5 shows the CB-CSMA/CA access mechanism for this application. It is assumed that STA\(_1\) and STA\(_2\) belong to cluster \( C_{s1} \) while STA\(_3\) and STA\(_4\) are members of cluster \( C_{s2} \).

In general if AF relays cannot decode any packet, both link errors and collisions can only be detected after the second-hop transmission. However, as in this application they are able to decode the identification message the collisions can be post-detected upon reception of the identification message.

As it is shown in Figure 5.5(a), the ACK frames are transmitted by the destinations and forwarded by the relay \( r_i \) sequentially with a SIFS period in between. In this way, the ACK duration depends on the number of concurrent streams per cluster. To avoid collision at the beginning of the next transmission attempt, the LENGTH field in the identification message should be extended to include the data and all ACK frames durations. We assume that the maximum cluster size and therefore the ACK duration and ACK-timeout are fixed\(^8\).

As shown in Figure 5.5, the average time required for successful data transmission (\( T_s \)),

\(^8\)Otherwise, the value of the LENGTH field should be set to its largest value in a cluster
Chapter 5  A Novel Cooperative MAC Protocol

Figure 5.5: CB-CSMA/CA access mechanism in application B.
collision ($T_c$) and channel error ($T_e$) for a cluster with $N_{nc}$ members can be expressed as:

\[ T_s = T_{id} + \delta + SIFS + ACK_{id} + \delta + SIFS + 2T_{Data} + 2\delta + SIFS \]
\[ + N_{nc} \cdot (2SIFS + 2T_{ACK} + 2\delta) + DIFS, \]
\[ T_c = T_{id} + \delta + SIFS + T_{CWUR} + \delta + DIFS, \]
\[ T_e = T_{id} + \delta + SIFS + ACK_{id} + \delta + SIFS + 2T_{Data} + 2\delta + SIFS + T_{ACKtimeout} + DIFS, \]

respectively, where $T_{ACKtimeout} = N_{nc} \cdot (2SIFS+2T_{ACK}+2\delta)+\sigma$ and $T_{id}$ is the duration of the identification packet. Similar to application A, in this application the uplink and downlink utilisation is equal and throughput of each destination can be obtained from (5.3).

**Delay Investigation:** Like application A, the HoL delay of application B can be calculated using (5.4).

### 5.3.3 Application C

In this application we have a two-way relaying scenario in an infrastructure network where the AP is able to act as the two-way relay. In [93] we have studied the integration of two-way relaying in IEEE 802.11n systems and proposed a modification of the standard polling procedure to support two-way relaying. Here, we consider a two-way relaying scenario which operates according to the CB-CSMA/CA protocol.

Similar to application A, the AP defines clusters. However in this application, it allocates each SD pair into a single cluster, i.e., $N_{nc} = 2$. We consider a simple two-way relaying scenario with $N_a = 2$ and where the AP does not have CSIT. In this case, during downlink the AP simply divides its power equally between the two messages. Furthermore, in order to limit the required changes to the actual standard, it is assumed that in the downlink each antenna of the AP is allocated to a transmission towards one STA. The AP transmits two independent spatial streams using its two antennas as a multicast frame destined to both STAs.

**Throughput Investigation:** We assume again that the AP can forward the packets after a SIFS without participating in another contention procedure. As the AP is equipped with two antennas it is able to decode two ACK frames concurrently. Hence, contrary to the DCF where multicast frames are not acknowledged, in order to make the multicast phase

---

9Although in this case the performance is degraded compared to the cases where the AP performs superposition coding and allocates power optimally, it leads to a simple setup which can be realised by the existing 802.11n systems.
Chapter 5  A Novel Cooperative MAC Protocol

reliable, the STAs acknowledge successfully received packets. In this way, application C can be supported by the CB-CSMA/CA in a similar way to application A and we can apply directly (5.1)-(5.3) to calculate throughput of application C. Since in this application we only consider \( N_{nc} = 2 \), there are totally \( N_c = 2\lceil N_{sd}/2 \rceil \) clusters which contend for the channel.

**Delay Investigation:** Like application A, the HoL delay of application C can be calculated from (5.4).

**Semi-Two-Way Relaying:**

Since the two-way relaying scheme is not supported by the current standard, there is no equivalent reference system based on the IEEE 802.11 MAC. However, the semi-two-way relaying method (cf. Section 4.3.3) can be supported by the conventional CSMA/CA. The uplink transmissions in a semi-two-way relaying is the same as in the conventional uplink phase. However, during downlink the AP multicasts a frame to two STAs. Hence, to evaluate throughput of the semi-two-way relaying scenario, we need to extend expressions which are given in Section 3.1.4 due to the following reasons:

- According to the IEEE 802.11 MAC there is no ACK procedure for multicast transmission. Hence, there is not any retransmission, for downlink packets while packets can be retransmitted during the uplink. This leads to a different backoff model for STAs compared to that of the AP. It should be noted that in our investigation, it is assumed that the AP only forwards packets to other STAs and hence it only takes part in the multicast phase.

- Since each pair of destinations are served simultaneously, the efficiency of the downlink is increased as compared to that of the two-hop links in the IEEE 802.11 infrastructure networks.

It should be noted that, due to lack of ACK in downlink, the transmission duration is asymmetric in downlink and uplink phases. The collision durations, however, remains the same for all STAs. This happens since in each collision at least two nodes are involved but there is only one node with a shorter transmission duration, i.e., AP. Therefore, in each collision, there is at least one STA which has to wait for ACK-timeout.

As it is shown in Figure 4.5 and Figure 4.10, the two-way and semi-two-way relaying on average need two and three time slots, respectively, to carry a packet from one STA to another and the other way around. Therefore, compared to the conventional relaying scenarios, which need four time slots for the same packet transmission, higher spectral efficiency for two-way and semi-two-way relaying is expected. However, compared to the conventional relaying method, the cost of two-way relaying, are the required MAC modification, extra
antenna in the AP and self interference cancellation feature at the STAs. The CB-CSMA/CA provides solution to the first requirement. The last two requirements can be realised in a MIMO WLAN system by modifying IEEE 802.11n. In the rest of this section, we investigate throughput expressions of the semi-two-way relaying.

In semi-two-way relaying scenario, similar to the infrastructure network, the uplink and downlink throughput are not the same. However, since two STAs receive packets, each time the DF relay/AP attains the channel, the downlink throughput improves compared to standard infrastructure scenarios. It is assumed that the AP only forwards packets to the STAs as a multicast packet. Hence, to calculate transmission probability we need to distinguish between the AP with no retransmissions and all other STAs with an infinitely large retransmission limit. For the AP, the backoff window is always set to the initial value and consequently the transmission probability can be obtained from:

$$\tau_{AP} = \frac{2}{W + 1}. \quad (5.5)$$

The transmission probability of other STAs, $\tau$, can be obtained directly from (3.9). We can apply equations (3.25)-(3.26) to calculate collision probability and the probability that only one of the STAs or only the AP transmits in a time slot. It should be noted that we have $n_1 = 1$ and $n_2 = n - 1$. Hence, throughput of the AP (downlink) and aggregate throughput of all STAs in uplink can be calculated using:

$$S_{down} = S_{AP} = \frac{P_{sAP}(1 - P_{eAP})L}{T_{slot}},$$
$$S_{up} = \frac{P_{s}(1 - P_{e})L}{T_{slot}}, \quad (5.6)$$

where $T_{slot}$ can be obtained from:

$$T_{slot} = P_{id}\sigma + \left(P_{sAP}(1 - P_{eAP}) + P_{s}(1 - P_{e})\right)T_{a} + P_{sAP}P_{eAP}T_{eAP} + P_{s}P_{e}T_{e} + P_{c}T_{c}, \quad (5.7)$$

in which $P_{id}$ is the probability that the channel is idle and $P_{c}$ is the probability of collision. These values can be calculated from:

$$P_{id} = (1 - \tau_{AP})(1 - \tau)^{n-1},$$
$$P_{c} = 1 - P_{i} - P_{sAP} - P_{s}. \quad (5.8)$$
The average time the channel is sensed busy due to successful transmission, collision and channel error for STAs, i.e., $T_s$, $T_c$ and $T_e$, can be calculated from (3.16)-(3.18), respectively. However, for the AP we have:

$$
T_{eAP} = T_{sAP} = T_{Data} + \delta + \text{DIFS},
$$

$$
T_{cAP} = T_c.
$$

(5.9)

Each STA receives packets with the throughput:

$$
S_D = \min\{S_{up}/(n - 1), S_{down}/(n - 1/2)\}
$$

(5.10)

for large values of $n$, $S_D = \frac{1}{(n-1)/2}S_{down}$, where $n - 1$ is an even integer. Since in downlink the AP serves two STAs concurrently, in (5.10) for downlink the total number of STAs is divided by two.

### 5.4 Summary

In this chapter, we have introduced a new cluster-based CSMA/CA protocol which utilises distributed spatial multiplexing gain in an IEEE 802.11n network. After presenting the basics of the protocol, the impact of the signal processing on the MAC protocol have been discussed. Furthermore, feasibility of CB-CSMA/CA as a centralised channel access method and as a RTS/CTS access method in an infrastructure and ad hoc network have been briefly presented. It has been shown that CB-SMA/CA can reduce collision probability in a BSS significantly. Additionally, operation of the three considered cooperative protocols in Chapter 4 based on CB-CSMA/CA have been discussed and related throughput and delay equations are investigated.

Although, we have mainly focused on relaying networks, direct communication links can be easily supported by the CB-CSMA/CA. However in case of direct link communication, to realise distributed spatial multiplexing gain, there should be enough degrees of freedom available either at the source or at the destination.
Chapter 6

Network Performance of CB-CSMA/CA

In this chapter, we estimate the throughput and HoL delay of Chapter 5 applications with respect to PHY and MAC layers. First we obtain the outage probability of each scenario for a given outage rate via Monte-Carlo simulations. Then, the collision probabilities, throughput and delay values for a certain PHY rate are calculated from relevant equations in Chapter 5.

In this chapter we have assumed that an uncoded data rate of $R_{\text{out}} = 26 \text{ Mb/s}$ per hop is required. We consider a scenario where all STAs have the same distance to the relays/AP, hence the path loss has been neglected. However, in Chapter 7 we are dealing with a more realistic setup where STAs are positioned at different locations. We have assumed that all sources and destinations have always a non-empty buffer and all nodes have an infinite buffer size. All control frames like ACK and CWUR frames, are transmitted at the lowest rate, i.e., 6.5 Mb/s. It is assumed that the CWUR has a format similar to the ACK frame without the address field. The control frames are supposed to be error-free. Furthermore, all data frames have the same payload size. Except for the application-specific parameters which are defined in Section 6.1, other parameters used in our analysis are listed in Tables 3.2 and 3.3. The number of source-destination pairs, $N_{\text{sd}}$, in the network is increased from 4 to 8, 12, 16 and 20. The number of clusters $N_c$ grows proportionally.

In application A an AP with four antennas is assumed, while in application B, i.e., ad hoc case, the AP is replaced by 13 AF relays. Therefore, in both cases four STAs are grouped into a single cluster, i.e., $N_{\text{nc}} = 4$. As discussed in Section 5.2 the reference system for application A is the same infrastructure network which operates based on the IEEE 802.11 DCF.

We also consider Application C as our third representative scenario. In application C, a two-way DF relay with two antennas is used. Hence, each pair of STAs which have packets for each other are grouped into a single cluster with $N_{\text{nc}} = 2$. The first hop, similar to that
in the application A, is a multiple access phase. However, in the second hop the DF relay should be able to multicast the two received packets to the same STAs. The two-way relaying is not supported by the existing MAC. We consider a standard infrastructure network with an AP equipped with two antennas as our reference. In addition, the semi-two-way relaying method (cf. Section 4.3.3) is used as a second reference which can be supported by the standard MAC.

6.1 Application-Specific Parameters

Reference System: Since each STA is equipped with a single antenna and there is no spatial multiplexing gain, we take the required parameters from the standard IEEE 802.11n for a single spatial stream [1]. Here, all nodes including the AP contend for accessing the channel. Therefore, there are totally \( n = 2N_{sd} + 1 \) contending nodes in the network. The data and ACK durations are calculated using the following formulae [1]:

\[
T_{\text{Data}} = 36\mu s + \left\lceil \frac{40 + (16 + 6)/8 + L}{BpS(M)} \right\rceil \cdot T_{\text{sym}}, \tag{6.1}
\]

\[
T_{\text{ACK}} = 36\mu s + \left\lceil \frac{14 + (16 + 6)/8}{BpS(M')} \right\rceil \cdot T_{\text{sym}}, \tag{6.2}
\]

The ACK frame is sent at the basic data rate, \( R_s = 6.5 \text{ Mb/s} \), hence \( BpS(M') = 3.25 \), cf. Table 2.1.

Direct Link Setup: In this scenario in order to be consistent with the CB-CSMA/CA applications, the AP does not take part in the contention procedure and hence \( n = 2N_{sd} \). As the direct link setup requires a handshake prior to data transmission [39], we need to replace (3.16)-(3.18) by the following equations:

\[
T_s = T_{\text{Handshake}} + T_{\text{Data}} + \text{SIFS} + \delta + T_{\text{ACK}} + \text{DIFS} + \delta, \tag{6.3}
\]

\[
T_c = T_{\text{Req}} + \delta + T_{\text{Res},\text{timeout}} + \sigma + \text{DIFS}, \tag{6.4}
\]

\[
T_e = T_{\text{Handshake}} + T_{\text{Data}} + \delta + T_{\text{ACK},\text{timeout}} + \text{DIFS}. \tag{6.5}
\]

\[
T_{\text{Handshake}} = T_{\text{Req}}^{1} + \text{SIFS} + \delta + T_{\text{Req}}^{2} + \text{SIFS} + \delta + T_{\text{Res}}^{2} + \text{SIFS} + \delta + T_{\text{Res}}^{1} + \text{SIFS} + \delta,
\]

where \( T_{\text{Req}}^{1} \) and \( T_{\text{Res}}^{1} \) are the required transmission times for the DLS Request, and the DLS
6.1 Application-Specific Parameters

Response on the \(i\)th-hop, respectively. \(T_{\text{ACK}_{\text{timeout}}}\) is the same as in the reference system. The DLS Response timeout is obtained as follows:

\[
T_{\text{Res}_{\text{timeout}}} = \text{SIFS} + T_{\text{Req}}^2 + \text{SIFS} + \delta + T_{\text{Res}}^2 + \text{SIFS} + \delta + T_{\text{Res}}^1 + \text{SIFS} + \delta + \sigma.
\]

Each DLS Request and DLS Response frame body is 22 octets long. We have assumed that the DLS Request and DLS Response frames are transmitted at the data transmission rate and that they have MAC and PHY headers as the ones used in the ACK frame, cf. (6.2). The PHY and MAC headers of DLS data frames, are the same as the ones in the reference system. Throughput of each STA is given by:

\[
S_D = \frac{S}{2N_{sd}}, \quad (6.7)
\]

where \(S\) is the aggregate throughput.

**Application A:** In application A, the AP only forwards the packets for the other nodes without taking part in the contention procedure. Hence, for a network with \(N_c\) source- and destination-clusters, there are only \(N_c\) contending clusters. The common preamble consists of the legacy preamble and the source-cluster address. Assuming the mixed format preamble for PLCP headers, we only need to take an additional address field (8 \(\mu\)s). Each transmission in application A consists of four spatial streams. Consequently, the HT PLCP preamble is extended to contain four long training sequences to make channel estimation possible. Thus, the data and ACK duration are obtained from:

\[
T_{\text{Data}} = (8 + 48) \mu s + \left[\frac{40 + (16 + 6)/8 + L}{BpS(M)}\right] \cdot T_{\text{sym}}, \quad (6.8)
\]

\[
T_{\text{ACK}} = 48 \mu s + \left[\frac{14 + (16 + 6)/8}{BpS(M')}\right] \cdot T_{\text{sym}}, \quad (6.9)
\]

\[
T_{\text{CWUR}} = 36 \mu s + \left[\frac{8 + (16 + 6)/8}{BpS(M')}\right] \cdot T_{\text{sym}}. \quad (6.10)
\]

**Application B:** In this application, AF relays only forward the packets for the other nodes. Hence, the relay-cluster does not participate in the contention and there are again \(N_c\) contending clusters in the network. Here the source-clusters send an identification message prior to the data transmission. Since the relays do not decode the data packets and the ACKs are transmitted sequentially, it is sufficient to have one training sequence in the HT PLCP preamble of both data and ACK frames. Besides, as the identification message includes...
the legacy preamble here the green field preambles can be used. Hence, the data and ACK durations in application B are calculated according to the following equations:

\[
T_{\text{Data}} = 24 \mu s + \left\lceil \frac{40 + (16 + 6)/8 + L}{BpS(M)} \right\rceil \cdot T_{\text{sym}}, \quad (6.11)
\]

\[
T_{\text{ACK}} = 24 \mu s + \left\lceil \frac{14 + (16 + 6)/8}{BpS(M')} \right\rceil \cdot T_{\text{sym}}, \quad (6.12)
\]

\[
T_{\text{id}} = 20 \mu s + \left\lceil \frac{20 + (16 + 6)/8}{BpS(M')} \right\rceil \cdot T_{\text{sym}}. \quad (6.13)
\]

**Application C**: Application C is similar to application A with respect to frame format and header durations. Since here cluster size is equal to two, the channel estimation can be performed using two orthogonal training sequences. Thus, the data and ACK duration can be obtained from:

\[
T_{\text{Data}} = (8 + 40) \mu s + \left\lceil \frac{40 + (16 + 6)/8 + L}{BpS(M)} \right\rceil \cdot T_{\text{sym}}, \quad (6.14)
\]

\[
T_{\text{ACK}} = 40 \mu s + \left\lceil \frac{14 + (16 + 6)/8}{BpS(M')} \right\rceil \cdot T_{\text{sym}}, \quad (6.15)
\]

**Semi-Two-Way Relaying**: We consider a semi-two-way relaying scenario operating based on DCF. Consequently, same data and ACK durations are considered as those in the reference system. Similar to CB-CSMA/CA applications, we assume that a link is in outage whenever either its uplink or downlink is in outage. Consequently, we define the outage probability according to the performance of the weaker link and assume that both hops suffer from the same outage probability.

### 6.2 Outage Probability

The simulated outage probabilities versus \( \alpha = P/\sigma_n^2 \) for \( R_{\text{out}} = 26 \text{ Mb/s} \) are shown in Figure 6.1 and Figure 6.2, where in all scenarios and at each node we assume a white Gaussian noise with a variance of \( \sigma_n^2 \) and a transmit power of \( P \) per subcarrier.

In Figure 6.1 application A and B are considered. As it is seen, the SD link has a lower outage probability in the reference system than in application A based on the CB-CSMA/CA
scheme. The reason is that in both cases we have the same number of antennas at the AP. In the CB-CSMA/CA application, the antennas at the AP are used to cancel the multiuser interference but in the reference system they are all used to improve the individual link performance. However, the outage probability curves show only the PHY performance per individual link and they do not represent the overall performance. While the reference system provides the PHY uncoded rate of 26 Mb/s for a single STA, the CB-CSMA/CA applications support four STAs at the above rate concurrently.

As it is shown in Figure 6.1, for a constant $\alpha$, the direct link setup has a higher outage probability than the reference system. In the direct link setup the source-destination distance is assumed to be the same as the source-relay or relay-destination distance in other scenarios. However, in the DLS each communication link is a SISO link.

In Figure 6.2 we consider application C with CB-CSMA/CA. The semi-two-way relaying and reference system with $N_a = 2$ both operate based on the standard CSMA/CA are also included. It should be noted that there is no CSIT at the AP in all of these three scenarios. Due to lack of CSIT at the AP, in both semi-two-way and two-way scenarios the downlink phase is the limiting phase. Consequently, outage probabilities of both setups are the same. However, in the two-way scenario, two STAs are in one cluster and it has been assumed that a cluster is in outage whenever any of its members is in outage. Hence, cluster outage probability is higher in two-way setup, i.e., application C, compared to individual outage...
Figure 6.2: Outage probability vs. $\alpha = P/\sigma^2_n$. In application C each source and destination cluster has two STAs.

probability, cf. semi-two-way curve in Figure 6.2.

In order to evaluate the MAC throughput performance, we need to set a PHY rate. Figure 6.1 shows that in high SNR regimes the CB-CSMA/CA applications and DLS have similar outage values for the outage rate of 26 Mb/s. However the reference system has similar outage values for a much larger outage rate (here 78 Mb/s). Taking into account the outage behaviour, for our MAC throughput investigation we are taking an uncoded data rate of 26 Mb/s per link for the CB-CSMA/CA and DLS scenarios and an uncoded data rate of 78 Mb/s for the reference system. Assuming a coding rate of 3/4 into account, this leads to $BpS(M) = 9.75$ for the the CB-CSMA/CA and DLS scenarios and $BpS(M) = 29.25$ for the reference system.

In a similar way, we consider an uncoded data rate of 26 Mb/s for application C and the semi-two-way and an an uncoded data rate of 52 Mb/s for the reference system.

In the next sections, we investigate throughput of these scenarios for different number of STAs in the network.

Note that uplink and downlink in the reference system are 1x4 SIMO and 4x1 MISO links, respectively. Hence, compared to a SISO link, a spectral efficiency gain of at most $\log(4) = 2$ bps/Hz (or 40 Mb/s for a system with 20 MHz bandwidth ) can be expected. Here as an extreme case (in favour of the reference system), we have increased the data rate of the reference system to the next higher uncoded rate of IEEE 802.11n.
6.3 Collision Probability

Conditional collision probability versus the number of SD pairs in the network for $P_e = 0$ are plotted in Figure 6.3. In application A and B there are $N_{nc} = 4$ STAs per cluster, while in application C there are $N_{nc} = 2$ STAs in each cluster. As expected, $P_{\text{col}}$ grows for all scenarios when increasing the number of STAs, and consequently increasing the number of clusters in CB-CSMA/CA applications. However, it has always smaller values in CB-CSMA/CA applications as compared to the reference, DLS and semi-two-way scenarios, due to grouping STAs into clusters.

The small difference between the reference system and the DLS is due to the fact that in the reference system the AP needs to contend for the channel to forward data packets to destinations while in the DLS setup it does not contend for the channel. Although the semi-two-way relaying scenario is based on the standard CSMA/CA, its collision probability differs slightly from that of the reference, due to the multicasting at the AP. This leads to different $P_{\text{col}}$ for STAs and the AP. In Figure 6.3 for the semi-two-way scenario only the collision probability of the STAs is depicted.
6.4 Throughput

The end-to-end throughput which is calculated as the average throughput at destinations (cf. (3.41), (5.3) and (6.7)) is depicted in Figure 6.4. As it is shown in Figure 6.4 (a) despite a three times larger link data rate in the reference system, we observe a significant throughput gain for the CB-CSMA/CA applications. Although application A and B based on CB-CSMA/CA have slightly longer headers and a weaker PHY link performance than the reference system, they have higher throughput values than the reference system. This is mainly due to the clustering function which allows transmitting four packets in parallel. As it has been shown in Figure 6.3, the conditional collision probability is much smaller for CB-CSMA/CA protocol. Besides, the reference system significantly suffers from the downlink throughput. While the reference system throughput rapidly decreases with increasing number of STAs, the CB-CSMA/CA applications offer a reasonable throughput per STA even for 20 SD pairs, i.e., 40 STAs in a BSS. Application A has a higher throughput com-

![Figure 6.4: End-to-end throughput at each destination vs. number of SD pairs with $P_e = 0$.](image)

pared to application B because it benefits from shorter ACK transmissions as well as shorter $\text{ACK}_{\text{timeout}}$.

As it is shown in Figure 6.4 (b), application C has much higher throughput as compared to the reference system. Although both semi-two-way and reference system operate based on DCF, the semi-two-way scenario achieves a higher throughput than the reference system since it multicasts two packets to each SD pair in the downlink phase, cf. Figure 6.4 (b).
6.4 Throughput

Figure 6.5: Uplink aggregate throughput vs. number of STAs.

Figure 6.6: Uplink aggregate throughput vs. number of STAs for different values of GI in application A.
In order to exclude the performance difference due to distinct relaying methods and eliminate the impact of asymmetric uplink and downlink utilizations in the reference system, we also investigate the uplink throughput separately. Nevertheless, even by leaving the downlink transmissions aside, the reference system does not perform as well as the CB-CSMA/CA applications. Figure 6.5 shows the aggregate uplink throughput versus different number of STAs for the reference system and application A. For comparison, the aggregate throughput of DLS is also plotted, but since a single hop transmission in application B (AF relays) is not meaningful, we have excluded it from Figure 6.5. For this figure, it is assumed that there is no channel error. However, again the difference in PHY characteristics has been taken into account by keeping different PHY rates for application A and DLS as compared to that of the reference system. As shown in Figure 6.5, the CB-CSMA/CA application significantly outperforms both reference and DLS scenarios. The DLS has a lower throughput as compared to the reference system due to its lower PHY rate and additional handshake.

As discussed in previous chapter, in a multiuser MIMO system different signals may arrive at different time slots. Hence, the GI should be large enough to accommodate the delay offset. Up to this point we have always used the same GI in all applications (GI=0.8 $\mu$s). At this point, for application A we consider a GI duration of 1.6 $\mu$s. Consequently, its uncoded data rate drops to 18 Mb/s. Uplink throughput of application A with the new parameters is plotted in Figure 6.6. The throughput of the reference system (which has GI=0.8 $\mu$s)
and the throughput of the application A with GI=0.8 μs are also plotted. In spite of a GI
duration twice as long as in the reference system and a much lower PHY data rate, the
CB-CSMA/CA application achieves higher throughput than the reference system (see Fi-
gure 6.6). For example in a network with 40 competing STAs, the CB-CSMA/CA achieves
about 83% larger throughput than the reference system.

In order to take the impact of the channel error on throughput into account, we increase the
frame error rate from 0 to 0.2. The throughput of application A and B versus frame error rate
for 8 SD pairs are depicted in Figure 6.7. We set $P_e$ to be the cluster frame error probability.
For comparisons the throughput of the reference system with packet frame probability of
zero is also plotted.

Different backoff models are assumed for the CB-CSMA/CA applications as compared
to the reference systems based on the CSMA/CA. In CB-CSMA/CA application the backoff
window size does not increase after channel errors rather it resets to its initial value. Detailed
analysis of these backoff models can be found in Chapter 7.

As it will be explained in Section 7.1, we can also apply the standard backoff model to the
CB-CSMA/CA protocol. This leads to a better performance for large number of clusters but
degrades the performance when the number of clusters is small (cf. Section 7.1).
Chapter 6  Network Performance of CB-CSMA/CA

6.5 Delay

The HoL delay values versus number of SD pairs for the CB-CSMA/CA applications and that of the DLS case are plotted in Figure 6.8. The channel error probability is set to 0 for all scenarios. As explained in Section 5.2, in the reference system the queueing delay at the AP in the saturation mode is unbounded and it is not considered here. However, for comparison purposes, the HoL delay of the uplink in the reference system is also depicted. Surprisingly, despite a much lower link rate, the delay of the CB-CSMA/CA two-hop applications is in the same range as that of the reference system with single-hop links. As it is shown in Figure 6.8 applications A and B even achieve lower delay as that of the respective reference system with single-hop communication links.

6.6 Summary

In this chapter, throughput and MAC delay results for the considered CB-CSMA/CA applications and their respective reference systems have been presented. The results show that throughput and delay of an IEEE 802.11 system can be significantly improved when the CB-CSMA/CA is applied. For example, despite three times larger link rate in the reference system based on IEEE 802.11n, CB-CSMA/CA application A enjoys an uplink throughput over two times that of the reference system.

In this chapter, it has been assumed that cluster members are perfectly event synchronised and all STAs experience on average the same channel condition. In the next two chapters, we take more realistic setups into account, where for example, STAs in a cluster suffer from event-synchronisation error or they have different path loss values. Additionally, three different backoff models and their impact on throughput are studied in details.
Chapter 7

Cross-Layer Enhancements of CB-CSMA/CA

In this chapter we study the network performance of the CB-CSMA/CA protocol by taking into account the channel error probability and SNR. In the first section we consider three different backoff models and study their behaviour under different channel conditions. In the second section we look at a realistic scenario where STAs experience different SNR values and we examine different clustering approaches based on the individual PHY or higher layer conditions.

For the numerical results we consider an uplink scenario in an infrastructure network. However, the analysis and principles can be adapted for other applications, too. In the last section we include a two-hop link scenario.

7.1 Impact of Backoff Procedure on CB-CSMA/CA Performance

The backoff procedure of the IEEE 802.11 avoids collisions by an exponential increase of the CW after an unsuccessful transmission. As discussed in Section 3.1.4 and described in other literature ([74], [75]), the binary exponential backoff causes unfairness in the presence of channel errors. In Chapter 5, we have exploited the collision post-detection capability of the CB-CSMA/CA and proposed a backoff procedure for the CB-CSMA/CA which is different from that of the standard CSMA/CA. In this section, in addition to the standard exponential backoff and the one suggested in Chapter 5, we consider a third backoff procedure and analyse performance of all three cases under different conditions.
Chapter 7 Cross-Layer Enhancements of CB-CSMA/CA

Figure 7.1: Average channel error probability of a cluster for different numbers of STAs per cluster.

Assuming $P_e$ as the channel error probability, then in CB-CSMA/CA, $P_e$ should represent the channel error probability of a cluster and not that of the individual STAs. We set $P_e$ to its upper bound by assuming, that a cluster faces channel errors whenever any of its links faces an error. In practice channel error probability can be set to packet error rates. Packet error rates can be calculated at each STA by dividing the number of the erroneous packets to the total number of received packets.

We have assumed that a cluster experiences a channel error whenever any of its members’ transmissions fails due to channel errors. In order to find out how much a cluster channel error probability differs from that of its individual STAs, we study a worst case scenario, where different STAs in a cluster have different values of $P_e$ taken randomly from a uniform distribution. We calculate the probability of channel error of the cluster and plot it versus average $P_e$ of one STA. As shown in Figure 7.1, the cluster error probability is much higher than that of individual STAs but as the cluster size increases the difference between cluster error probabilities decreases.

7.1.1 Backoff Models

We consider the following backoff models:
7.1 Impact of Backoff Procedure on CB-CSMA/CA Performance

Model I-IEEE 802.11 Standard Backoff: In this model, after any unsuccessful transmission regardless of the source of error, the CW doubles up to a maximum predefined value. However, when the maximum CW is reached, the CW remains constant until the next successful transmission. Any successful transmission resets CW to its initial value irrespective of the current backoff stage.

This backoff model (contrary to the next considered backoff models) does not distinguish between losses due to collision and the ones due to channel error. Therefore, it can be applied to the CB-CSMA/CA in the following situations: (i) the collision detection is not possible anymore (for example when operating in the low SNR regime).(ii) Collisions can be distinguished from channel errors, however the CW values are updated after any unsuccessful transmission.

This backoff model has been widely studied and several mathematical models have been developed for it [6], [34] and [56]. As explained in Chapter 3, in this thesis we employ a model which has been extensively used in the literature i.e. the discrete-time Markov chain (MC) which is proposed in [6] and an extended version of that which additionally takes channel errors into account [57] [8].

To apply this model, we need the same approximation as in [6], i.e., at each transmission attempt collisions occur with constant and independent probability. It is assumed that collision and channel error events are statistically independent [8]. The Markov Chain for this backoff model is shown in Figure 3.4. The transition probability from one backoff stage to another one is [8]:

\[ p = 1 - (1 - P_{\text{col}}) (1 - P_e) = P_e + P_{\text{col}} - P_e P_{\text{col}}. \]

Model II: As the second backoff model we use the model which has been so far considered for the CB-CSMA/CA, i.e., applying the standard backoff procedure only in case a collision occurs and resetting the CW to its initial value in any other case (this includes successful transmissions as well as unsuccessful transmissions due to channel errors). In [109] this model has been applied to a CSMA/CA-based system. Note that contrary to the previous model, for this backoff procedure, a collision post-detection scheme is required. The Markov chain model for this backoff procedure can easily be obtained from the MC for the Model I (cf. Figure 3.4) by replacing \( p \) with \( P_{\text{col}} P_e + P_{\text{col}} (1 - P_e) = P_{\text{col}} \). Hence, the MC is simplified to the standard case with ideal PHY channel, i.e., the model shown in [6] with a transition probability \( p = P_{\text{col}} \).

Model III: Increasing the contention window value in case of channel errors is not in favour of the STA with high \( P_e \) and it is a waste of resource when there are only few clusters in the network. However, it may increase the efficiency of the STAs with low \( P_e \), specially
when there are several clusters in a network. Therefore, we consider a third backoff model in which again the same procedure as Model I and II is considered in case of collisions. However, the CW does not reset to its initial value after a channel error occurs rather it keeps its actual value until either, a successful transmission or a collision occurs. The MC for this backoff procedure is shown in Figure 7.2.

In [110] the above model has been used for a loss distinguishable WLAN and in [111] it has been applied to an IEEE 802.11b system. In both works, numerical values are calculated for the IEEE 802.11b parameters. Since the MC model and intermediate calculations for obtaining the transition probability are not given in those papers\(^1\), hence we study this model in detail in order to develop necessary equations for calculating the channel access probability, \(\tau\), and the conditional collision probability, \(P_{\text{col}}\), in this section.

**Channel Access Probability of Model III:** In order to calculate the channel access probability and the respective collision probability for model III, we depict the MC (cf. Figure 7.2), similar to the MC model in [6] which has been explained in Section 3.1.4.1, the backoff stage \(s(t)\) and the backoff timer counter \(b(t)\) are modeled as a bidimensional process using a discrete-time MC. Consider \(b_{i,k} = \lim_{t \to \infty} \Pr\{s(t) = i, b(t) = k\}\) to be the stationary distribution of the chain with \(i \in (0, m)\) and \(k \in (0, W_i - 1)\).

Writing down balance equations for this MC we obtain:

\[
b_{i-1,0} = \frac{P_{\text{col}}}{W_i} + b_{i-1,0} \frac{W_i - 1}{W_i} P_{\text{col}} + \left(\frac{(1 - P_{\text{col}})P_e}{W_i} + \frac{W_i - 1}{W_i} (1 - P_{\text{col}})P_e\right) b_{i,0} = b_{i,0}. \tag{7.1}
\]

or

\[
b_{i,0} = \left(\frac{P_{\text{col}}}{P_{\text{col}} + \hat{P}_I}\right) b_{i-1,0}. \tag{7.2}
\]

where \(\hat{P}_I = 1 - P_I = 1 - \left(1 - (1 - P_{\text{col}})(1 - P_e)\right) = (1 - P_{\text{col}})(1 - P_e)\).

This results in:

\[
b_{i,0} = \left(\frac{P_{\text{col}}}{P_{\text{col}} + \hat{P}_I}\right)^i b_{0,0}. \tag{7.3}
\]

For \(k \in (1, W_i - 1)\) three regions are considered:

\(^1\)In [110] the authors consider four different backoff models but only provide the analysis of two of them.
7.1 Impact of Backoff Procedure on CB-CSMA/CA Performance

Figure 7.2: Markov chain model for the backoff model III.
i) $i = 0$

For $k = 1$ we obtain:

$$b_{0,1} = \frac{(1 - P_{\text{col}})(1 - P_{e})}{W_0} \sum_{j=0}^{m} b_{j,0} + 1 \times b_{0,2} + \frac{P_e(1 - P_{\text{col}})}{W_0} b_{0,0}$$

$$= \frac{(1 - P_{\text{col}})(1 - P_{e})}{W_0} \sum_{j=0}^{m} b_{j,0} + \left\{ \frac{(1 - P_{\text{col}})(1 - P_{e})}{W_0} \sum_{j=0}^{m} b_{j,0} + 1 \times b_{0,3} + \frac{P_e(1 - P_{\text{col}})}{W_0} b_{0,0} \right\} + \frac{P_e(1 - P_{\text{col}})}{W_0} b_{0,0} = ...$$

$$= \frac{W_0 - 1}{W_0} (1 - P_{\text{col}})(1 - P_{e}) \sum_{j=0}^{m} b_{j,0} + \frac{W_0 - 1}{W_0} P_e(1 - P_{\text{col}}) b_{0,0}$$

In a similar way we can calculate $b_{0,k}$ as:

$$b_{0,k} = \frac{W_0 - k}{W_0} (1 - P_{\text{col}})(1 - P_{e}) \sum_{j=0}^{m} b_{j,0} + \frac{W_0 - k}{W_0} P_e(1 - P_{\text{col}}) b_{0,0}. \quad (7.5)$$

ii) $0 < i < m$

Again we begin with $k = 1$:

$$b_{i,1} = \frac{P_{\text{col}}}{W_i} b_{i-1,0} + 1 \times b_{i,2} + \frac{P_e(1 - P_{\text{col}})}{W_i} b_{i,0} = ...$$

$$= \frac{W_i - 1}{W_i} P_{\text{col}} b_{i-1,0} + \frac{W_i - 1}{W_i} P_e(1 - P_{\text{col}}) b_{i,0}, \quad (7.6)$$

Similarly, we can obtain $b_{i,k}$:

$$b_{i,k} = \frac{W_i - k}{W_i} P_{\text{col}} b_{i-1,0} + \frac{W_i - k}{W_i} P_e(1 - P_{\text{col}}) b_{i,0}. \quad (7.7)$$

iii) $i = m$
For $k = 1$ we have:

$$b_{m,1} = \frac{P_{\text{col}}}{W_i} b_{m-1,0} + \frac{P_e(1 - P_{\text{col}})}{W_i} b_{m,0} + 1 \times b_{m,2} = $$

$$= \frac{W_m - 1}{W_m} P_{\text{col}} b_{m-1,0} + \frac{W_m - 1}{W_m} (1 - P_e)(1 - P_{\text{col}}) b_{m,0}$$

$$= \frac{W_m - 1}{W_m} \left( P_{\text{col}} b_{m,0} + P_f b_{m,0} \right) = \frac{W_m - 1}{W_m} b_{m,0},$$

(7.8)

and for $k > 1$:

$$b_{m,k} = \frac{W_m - k}{W_i} P_{\text{col}} b_{m-1,0} + \frac{W_m - 1}{W_m} P_f b_{m,0}.$$  

(7.9)

Using (7.5)-(7.9) we get:

$$\sum_{i=0}^{m} b_{i,0} = \frac{P_{\text{col}} + \tilde{P}_f}{P_f} b_{0,0}.$$  

(7.10)

Now by using the above equations for each $k \in (1, W_i - 1)$:

$$b_{i,k} = \frac{W_i - k}{W_i} \begin{cases} (1 - P_f) \sum_{j=0}^{m} b_{j,0} + P_e(1 - P_{\text{col}}) b_{0,0} = b_{0,0}, & i = 0 \\ P_{\text{col}} b_{i-1,0} + P_e(1 - P_{\text{col}}) b_{i,0} = b_{i,0} & 0 < i < m \\ P_{\text{col}} b_{m-1,0} + P_f b_{m,0} = b_{m,0} & i = m \end{cases}$$

(7.11)

The normalisation condition implies that:

$$\sum_{i=0}^{m} \sum_{k=0}^{W_i-1} b_{i,k} = 1,$$

(7.12)

From (7.10)-(7.12) and by performing some algebraic manipulations we obtain $b_{0,0}$:

$$b_{0,0} = \frac{2(1 - 2\beta)(1 - \beta)}{(1 - 2\beta)(W + 1) + \beta W(1 - (2\beta)^m)},$$

(7.13)

where $\beta$ is an auxiliary function which is defined by:

$$\beta = \frac{P_{\text{col}}}{P_{\text{col}} + \tilde{P}_f} = \frac{P_{\text{col}}}{1 - P_e + P_{\text{col}} P_e},$$

(7.14)

Similar to model I, in CB-CSMA/CA $P_e$ should represent the channel error probability per
Chapter 7  Cross-Layer Enhancements of CB-CSMA/CA

cluster. The channel access probability can be easily obtained from:
\[
\tau = \sum_{i=0}^{m} b_{i,0} = \frac{b_{0,0}}{1 - \beta} = \frac{2(1 - 2\beta)}{(1 - 2\beta)(W + 1) + \beta W (1 - (2\beta)^m)},
\]
(7.15)

\(P_{\text{col}}\) is given in (3.10) and recalled here:
\[
P_{\text{col}} = 1 - (1 - \tau)^{n-1},
\]
(7.16)

where \(n = N_c\).

We replace \(\beta\) in (7.15) by its equivalent from (7.14). In this way equations (7.15) and
(7.16) provide a nonlinear system of equations which can be solved applying numerical
methods. The unknowns are \(\tau\) and \(P_{\text{col}}\). Similar to the standard CSMA/CA [6], we can show
that this system of equations has a unique solution as shown below:

By inverting (7.16) we obtain: \(\tau^*(P_{\text{col}}) = 1 - (1 - P_{\text{col}})^{1/(n-1)}\) which is a continuous
function and increases monotonously for \(P_{\text{col}}\in (0, 1)\). We have \(\tau^*(0) = 0\) and \(\tau^*(1) = 1\).
In (7.15) \(\tau\) in the range \(P_{\text{col}}\in (0, 1)\) is also continuous and monotone decreases with \(\beta\) and
hence with \(P_{\text{col}}\). It should be noted that (7.15) can also be written as \(\tau = \frac{2}{1+W + \beta W \sum_{i=0}^{m-1}(2\beta)^i}\)
so as it is seen \(\tau\) is continuous at \(\beta = 1/2\) too. Assume any \(P_e \neq 0\) then we have:
\[
\beta = \begin{cases} 
0 & , P_{\text{col}} = 0 \\
1 & , P_{\text{col}} = 1.
\end{cases}
\]
(7.17)

which directly leads to \(\tau(0) = 2/(W + 1)\) and \(\tau(1) = 2/(1 + 2^m W)\). Since \(\tau(0) > \tau^*(0)\)
and \(\tau(1) < \tau^*(1)\) there is only one solution.

As an example we assume \(P_e = 0.05\) while \(n = N_c\) is set to 3, 10, and 30. Both \(\tau^*\)
and \(\tau\) versus \(P_{\text{col}}\) are plotted in Figure 7.3 using IEEE 802.11a or IEEE 802.11n system
parameters. As it is seen, for each \(n\) there is only one cross point which is the solution of the
two equations. For example for 10 contending clusters, \(N_c = 10\), we obtain \(\tau = 0.0513\) and
\(P_{\text{col}} = 0.3777\).

7.1.2 CB-CSMA/CA Modifications

So far we have applied the backoff model II in CB-CSMA/CA applications. According to
this model, STAs only have to update their CW in case of collisions and reset CW to the
initial value in all other cases. Hence, the CB-CSMA/CA has been designed to fulfill this
7.1 Impact of Backoff Procedure on CB-CSMA/CA Performance

Figure 7.3: Unique solution of the system equations given in (7.15) and (7.16) for IEEE 802.11a parameters and $P_e = 0.05$, $N_c = 3, 10, \text{and } 30$.

requirement and a CWUR packet is used to notify STAs whenever a collision is detected. On the other hand, backoff models I and III require that, STAs not only be aware of collisions but also they should be notified about PHY error on any of the parallel streams. Consequently, the CB-CSMA/CA should be modified to provide this information to cluster members. In this section, we distinguish between infrastructure and ad hoc networks and explain the required protocol changes for each case.

**Infrastructure Network:** As it has been discussed in Section 5.1.2, the CB-CSMA/CA can detect sources of transmission failures with high probability. In order to notify all STAs in a cluster about any present PHY error we should carry out the following modifications:

- The STAs which have transmitted a packet and neither received an ACK nor a CWUR should update their backoff.

- Additionally, the AP should include information about the presence of channel errors on any of the parallel links in the ACK frames. Hence, the STAs which have sent packets successfully are aware of the fact that some of the other cluster members could not transmit successfully. In this case they can also update the CW according to the backoff model, i.e., double CW if model I is used or keep the same CW as the one in the previous transmission if model III is used.

In the following analysis for simplicity, we consider the same ACK frame duration as in the previous chapters.
Chapter 7 Cross-Layer Enhancements of CB-CSMA/CA

**Figure 7.4:** Aggregate throughput for different backoff model vs. number of clusters (CB-CSMA/CA).

**Ad Hoc Network:** According to CB-CSMA/CA, in an ad hoc network the cluster-masters are responsible for detecting the source of transmission failure by checking the status of the received identification message. In order to facilitate backoff model I and III, the cluster-master should notify all cluster members in case of any channel error. In the considered MUZFR scenario (application B in Section 5.3.2), all ACKs are forwarded by the cluster-master of the relay-cluster. Hence if the relay cluster-master does not receive one or more ACKs, it can include this information in other ACKs. In this way, all STAs increase their CW value after a collision. After a channel error all STAs either keep their actual CW value or increase it, depending on the backoff model used.

**7.1.3 Performance Results**

In order to compare the performance of the CB-CSMA/CA scheme with different backoff procedures, the channel access probability for each case should be calculated. Using the channel access probability, the aggregate throughput under different conditions including different channel error probabilities are calculated. In all cases we assume a block fading channel model. Hence, STAs see a different channel each time they access the medium.

We consider an uplink scenario in an infrastructure network, where AP has four antennas.
and each source-cluster consists of four STAs. At first we consider a symmetric case where all clusters face the same channel error probability $P_e$. A data rate of 58.5 Mb/s (uncoded data rate of 78 Mb/s with a FEC coding rate of 3/4) and a basic rate of 6.5 Mb/s are assumed. The average aggregate throughput over different channel realisations with different $P_e$ is calculated and is shown in Figure 7.4. For each realisation, $P_e$ is chosen randomly within the range [0 1]. As it is seen in this figure, for a small number of contending clusters, the backoff procedure models II and III perform more efficiently than the standard one. In this regime, where the load is light, model I wastes resources by doubling the CW whenever a PHY error occurs. However, the backoff model I performs better, when the number of clusters is larger than a certain threshold. In this region, the load is heavy and the increase of CW (due to a channel error) helps to reduce the collision probability.

The thresholds, at which model I outperforms the other two models, depend on system parameters and setups. For example for lower data rates the threshold shifts to smaller number of clusters, while for higher data rates it moves towards larger number of clusters. This is mainly due to the fact that for the same number of clusters, the backoff waiting time occupies a larger fraction of the total time for higher transmission rates as compared to lower rates. The difference is considerably larger for model I than for the other two models, because on average model I makes STAs wait longer. Hence, for higher transmission rates both threshold values shift to the larger number of clusters.

To study the impact of the backoff model on throughput performance in presence of different channel error probabilities, we set the number of clusters once to $N_c = 2$ and another
time to 60 and calculate throughput while changing $P_e$ values. The results are shown in Figure 7.5. For a large number of clusters models I and III outperform Model II. However the additional cost (as it has been discussed in the previous section), is the need to communicate any channel errors to all STAs within a cluster in case of model I and III. As it is seen for small values of $P_e$, differences between the curves become small.

Now we consider an asymmetric case in which clusters experience different channel error probabilities. We study a case in which half of the clusters have no channel error and the other half have $P_e = 0.5$. We calculate throughput for both groups having different backoff models. Results are depicted in Figure 7.6. While the standard backoff procedure is in favour of clusters with strong channels, it drastically penalises the throughput of the clusters with weaker channels. This happens because according to the backoff model I, after each unsuccessful transmission the CW interval doubles (up to its maximum value) regardless of the type of errors. Hence, under backoff model I, clusters with high channel error probability can access the channel with lower probability while the ones with no channel error benefit from this situation and their transmission probability becomes larger.

To study the impact of channel error probability on the aggregate throughput in more details, four setups with different values of $P_e$ are considered. The aggregate throughput in the asymmetric scenario is plotted in Figure 7.7 where half of the clusters always have no channel error and the other half have $P_e = 0, 0.1, 0.5, \text{ and } 0.9$. 

**Figure 7.6:** Throughput per cluster in asymmetric case with different channel error probabilities and for different backoff model vs. number of clusters.
Figure 7.7: Aggregate throughput vs. number of clusters in the asymmetric case, half of the clusters have no channel errors and the other half have channel errors with probability $P_e$. 
Again we observe that for larger values of $P_e$, model I outperforms the other two models and model III outperforms model II. This is in line with what has been discussed so far. However, surprisingly, with backoff model I and for $P_e = 0.9$ the aggregate throughput is higher than that with $P_e = 0$. The reason is that in the asymmetric setup, in the extreme case, i.e., when $P_e$ is very large, the STAs with large $P_e$ hardly can access the channel and they are out of the competition. These STAs suffer from long contention windows while other STAs with $P_e = 0$ enjoy transmission in a "virtually" less loaded network.

To check this speculation, we consider again the same setup and plot the aggregate throughput using backoff model I, for $P_e = 0, 0.9$ and 1. We have also plotted the aggregate throughput for the case where only half of the STAs are contending for the channel. As it is seen, the throughput of the cases with $P_e = 0.9$ and 1 lies between the other two curves, both with $P_e = 0$.

Although we have investigated backoff performance for $0 \leq P_e \leq 1$, in practice we should keep $P_e$ values as low as possible. As it has been mentioned in Section 3.1 the IEEE 802.11 defines the receiver requirements such that the frame error ratios can be kept small. Nevertheless, study of extreme cases is interesting as it shows the performance under poor conditions (for example close to the border of an AP coverage area). Furthermore, it helps us to better understand the differences between the models.

According to the results obtained, we conclude that the backoff model should be selected...
7.1 Impact of Backoff Procedure on CB-CSMA/CA Performance

as explained in the following:

- If the number of clusters in a BSS is known to all clusters, then they can adaptively choose their backoff model, i.e., they should choose the standard backoff procedure when the number of contending clusters is large (e.g., in the considered symmetric setup cf. Figure 7.4 for number of clusters larger than 10). Otherwise, backoff model III can be used. Although the number of clusters can be broadcast to clusters for example by the AP in an infrastructure network, knowledge of the number of clusters in a BSS may not be feasible in some situations. For example, in a hot spot where many STAs may enter or exit the BSS within a short time or in an ad hoc network, where no central node exists.

- In cases where the STAs do not know the number of clusters in the network or due to system implementation there is not possibility to adaptively change the backoff model, the best option is to apply backoff procedure Model III all the time. In practice, often the network size is not very large. As it is shown in Figure 7.4, for a number of clusters below 10 (or a number of STAs below 40), Model III on average outperforms the other two models.

- In practice where link adaptations are possible, assuming that a system performs well in a given area, then we should focus on low values of $P_e$. In this regime, model III can be selected for a small number of clusters while model I should be chosen for medium to large number of clusters.

- In an asymmetric case, where different clusters may experience a large variety of channel error probabilities, the backoff procedure can be chosen depending on the application, cluster priority or fairness. For example, as it has been shown in Figure 7.6, using backoff model II reduces the performance difference between the clusters with poor channel and the ones with good channel. Although in this example the aggregate throughput of the model II is lower than that of model I and III, it can be used for cases in which the performance of the weakest clusters should be improved. Besides, assuming that the collisions can be post-detected, the backoff model II can be applied as a simple model which does not require any CW increment after channel errors.

For the rest of this chapter, depending on collision detection possibility, we consider either the backoff model I or III for the CB-CSMA/CA. It should be noted that backoff model III is more sensible for CB-CSMA/CA compared to CSMA/CA, where individual STAs contend for the channel, since the number of contending units in a CB-CSMA/CA system in general is smaller than that in a CSMA/CA system. Besides, in high SNR regime the CB-CSMA/CA is able to post-detect collisions.
Chapter 7 Cross-Layer Enhancements of CB-CSMA/CA

7.2 Impact of Clustering on CB-CSMA/CA Performance

So far in CB-CSMA/CA analysis, the same channel statistics (the same channel model and no pathloss), the same transmit power and hence the same average PHY rate and the same $P_e$ for all nodes are used. Therefore, the STAs were allocated to different clusters without any distinction. However, in practice some source-destination links may have on average stronger channels (e.g., due to shorter distance between a source and its destination). Furthermore, due to the nature of application, some STAs may transmit shorter packets than the others. In this case, different ways of grouping STAs into a cluster may influence the performance of the network.

Different clustering methods have been suggested in different works such as [99], [112–114]. However most of them consider the conventional clustering (hierarchy clusterings where nodes in a cluster only communicate with the cluster-heads) issues. Additionally, they consider different criteria including decreasing power consumption in sensor networks. An architecture called linked cluster architecture has been suggested for mobile radio networks in [112] where algorithms to determine the role of each node in the network such as cluster-head, gateway or ordinary, is described. In [113], the authors have suggested an access-based clustering protocol. They have focused on different clustering issues while their goal has been to provide a generic and flexible cluster structure for the upper-layer protocols. In [114] MAC throughput of a clustering algorithm based on CSMA/CA for ad hoc networks has been analysed. According to this algorithm if a node, which has a data packet to send, senses the channel as idle for a random time it broadcasts a periodic beacon, establishes a cluster and acts as a master of the cluster. The authors have proposed to apply any infrastructure high-performance MAC protocol for data transmissions within each cluster.

In [115] virtual grouping is used to reduce the collision probability. However, according to that paper the contention period is divided such that each group is only allowed to transmit in a certain part of it.

In CB-CSMA/CA systems the clustering may be performed based on application, different requirements and PHY layer characteristic. In this section we study different clustering methods for CB-CSMA/CA and the impact of them on the network throughput. We choose some examples to show the impact of clustering on the throughput in different scenarios. In order to improve other performance criteria, clustering can be performed differently. However, in all scenarios the clusters should be formed such that the receivers can decode the concurrent streams.

We mainly focus on two heterogenous scenarios. In the first scenario we look into clus-
7.2 Impact of Clustering on CB-CSMA/CA Performance

tering according to higher layer requirements. As an example a scenario where different STAs have different types of traffic and hence different packet sizes are considered. In the second scenario, we assume that all STAs have the same type of traffic but they face different channel statistics and hence on average different link SNRs.

7.2.1 Clustering According to the Higher Layer Parameters

STAs can be allocated to clusters according to the application requirements. For example we can allocate STAs with low-delay requirement into one cluster and give high priority to this cluster, while allocating the STAs with applications which are more robust to delay, to other clusters with low priorities. As it has been presented in Section 5.1.6 it is also possible to form the clusters in an adaptive way, by taking into account presence of the packets at the STAs. In this section we consider an example where different STAs have different payload sizes.

As it has been mentioned in Section 5.1.3, in CB-CSMA/CA the LENGTH field of each frame should be set to the maximum length of all concurrent packets. In order to numerically analyse the throughput performance, when STAs have different packet sizes, we consider a scenario where two types of applications are running, each with a certain payload requirement. It is assumed that one application needs short packets and the other one long packets.

We consider an uplink scenario in an infrastructure network where the AP has four antennas and each cluster has four STAs. Always half of the total STAs have short packets of $L_1 = 64$ bytes to transmit, while the other half, transmit packets of $L_2 = 2048$ bytes. It has been assumed that number of clusters in the network, $N_c$, is an even number. We look into two scenarios: in the first scenario, i.e., homogeneous scenario, STAs within the same cluster have either only short or long packets to transmit, while in the second scenario, i.e., heterogeneous scenario, in each cluster two STAs transmit short packets and the other two transmit long packets. We assume no channel error, i.e., $P_e = 0^2$. In this case, since payload size does not influence the backoff procedure, all clusters transmit with the same transmission probability. However, the throughput in each scenario is calculated differently.

In the first scenario we have two types of clusters, one with short packets and the other with long packets. The data packet duration and hence $T_{s_i}$ and $T_{c_i}$ are not the same for

\[2\text{If } P_e \neq 0 \text{ then depending on the backoff procedure STAs with different } P_e \text{ may have different } \tau. \text{ Besides, having the same bit error rate results in a larger packet error rate for longer packets compared to that of shorter packets.}\]
different cluster types. They can be calculated from the following equations:

\[ T_{s_i} = T_{\text{Data}_i} + \text{SIFS} + \delta + T_{\text{ACK}} + \delta + \text{DIFS}, \]  
\[ T_{c_i} = T_{\text{Data}_i} + \delta + \text{SIFS} + T_{\text{CWUR}} + \delta + \text{DIFS}, \]  

(7.18)  
(7.19)

where \( T_{s_i} \) and \( T_{c_i} \) are the average time the channel is sensed busy due successful transmission and collision of the cluster type \( i \), respectively and data frame duration is:

\[ T_{\text{Data}_i} = T_{\text{PLCP}_p} + T_{\text{PLCP}_{STG}} + \left\lceil \frac{40 + (16 + 6)/8 + L_i}{BpS(M)} \right\rceil \cdot T_{\text{sym}}, \]  

(7.20)

Hence aggregate throughput can be obtained from:

\[ S^{(\text{homogeneous})} = \frac{N_c \tau (1 - \tau)^{(N_c-1)}}{T^{(\text{homogeneous})}_{\text{slot}}} \times \left( N_{\text{mc}} L_1 + N_{\text{mc}} L_2 \right), \]  

(7.21)

where \( T^{(\text{homogeneous})}_{\text{slot}} \) can be obtained as in the following:

\[ T^{(\text{homogeneous})}_{\text{slot}} = P_{\text{id}} \sigma + \sum_{i=1}^{2} P_s T_{s_i} + \sum_{i=1}^{2} P_c T_{c_i}, \]  

\[ P_{\text{id}} = (1 - \tau)^{N_c}, \]  

\[ P_s = \frac{N_c}{2} \tau (1 - \tau)^{(N_c-1)}, \]  

(7.22)

In order to calculate the collision probability, \( P_{c_i} \), we need to distinguish between the collision among clusters of the same type and the collision between clusters of different types. It should be noted that the duration of collision in the basic access method is determined by the duration of the longest packet involved in collision. The calculation of the collision probability in this setup is similar to that in scenarios where different STAs have different data rates. For each cluster type \( i \) we have two collision probabilities: intra-collision or collisions among the same cluster type with probability \( P_{c_i}^{(\text{intra})} \) and inter-collision or collisions...
between different cluster types with probability $P_{c}^{(\text{inter})}$. Hence, we have:

$$P_{c_{i}}^{(\text{intra})} = \left( 1 - \left( (1 - \tau)^{N_{c}/2} + \frac{N_{c}}{2} \tau (1 - \tau)^{\frac{N_{c}}{2} - 1} \right) \right) \cdot (1 - \tau)^{N_{c}/2},$$

$$P_{c_{1}}^{(\text{inter})} = (1 - (1 - \tau)^{N_{c}/2}) \cdot (1 - (1 - \tau)^{N_{c}/2}),$$

$$P_{c_{2}}^{(\text{inter})} = 0,$$

$$P_{c_{i}} = P_{c_{i}}^{(\text{intra})} + P_{c_{i}}^{(\text{inter})}. \quad (7.23)$$

where the first term in $P_{c_{i}}^{(\text{intra})}$ indicates the probability that more than one STA from the same group transmit and the second term indicates that no other STAs from another group transmits. For inter-group collisions we only need to take the collision of longer packets with shorter ones into consideration. As we have already taken the collisions of both groups in $P_{c_{1}}^{(\text{inter})}$ into account, $P_{c_{2}}^{(\text{inter})}$ is set to 0.

In the second scenario each cluster has two STAs with long packets, and two with short packets. The long packets determine the duration of $T_{s}$ and $T_{c}$. These values remain the same for all clusters. In this scenario the aggregate throughput can be obtained from:

$$S^{(\text{heterogeneous})} = \frac{N_{c} \tau (1 - \tau)^{(N_{c} - 1)} \left( \frac{N_{uc}}{2} L_{1} + \frac{N_{uc}}{2} L_{2} \right)}{T_{\text{slot}}^{(\text{heterogeneous})}}, \quad (7.24)$$

where

$$T_{\text{slot}}^{(\text{heterogeneous})} = P_{\text{id}}\sigma + P_{s}T_{s_{2}} + P_{c}T_{c_{2}}, \quad (7.25)$$

in which $P_{\text{id}}$ can be calculated as in the first scenario and $P_{s}$ is $N_{c}\tau (1 - \tau)^{(N_{c} - 1)}$. In this scenario, the collision duration is always the same for all STAs and hence the fraction of the time slot which is wasted by collision is $(1 - P_{\text{id}} - P_{s})T_{c_{2}}$.

The aggregate throughput for both scenarios is plotted in Figure 7.9. As it is shown, the throughput degrades when STAs with different packet sizes are allocated to the same cluster. This happens since in each cluster STAs with shorter packets have to wait until transmission of the longer packets are completed.

When the STAs with shorter packets have always multiple packets in their queue, then we can improve the throughput of the heterogenous scenario by aggregating multiple short packets into one frame. The number of aggregated short packets can be determined such that the aggregated frame has similar length to the longest packet in a cluster.

For example in the considered scenario and by assuming that each short packet has the
same headers as in the previous case, the aggregated frame can include five short packets. We have analysed this case and included the aggregate throughput of the improved heterogeneous scenario in Figure 7.9.

The throughput loss could be further reduced if only one PHY header is used for the aggregated frame. In this way, the number of data packets per aggregate frame can be increased. It should be noted that although this method can compensate a fraction of the throughput loss, it is not possible to use it for all types of traffic and scenarios. For example, if there are short voice packets, aggregating them into one frame may increase the delay and the error probability. In general, it is beneficial to allocate STAs with similar packet types and requirements into the same cluster.

### 7.2.2 Clustering According to the PHY Layer Parameters

STAs can be grouped into distinct clusters by taking into account their PHY characteristics and/or requirements, e.g., their channel conditions, SNR, packet error probability, PHY data rate, possibility of joint decoding/signal processing, number of antennas per STA, etc. In this section we focus on scenarios where different STAs have different channel statistics. This
Table 7.1: Distinguishable ways of grouping STAs into clusters

<table>
<thead>
<tr>
<th>Name</th>
<th>Way of Clustering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clustering 1</td>
<td>(STA_1, STA_2), (STA_3, STA_4), (STA_5, STA_6)</td>
</tr>
<tr>
<td>Clustering 2</td>
<td>(STA_1, STA_2), (STA_3, STA_5), (STA_4, STA_6)</td>
</tr>
<tr>
<td>Clustering 3</td>
<td>(STA_1, STA_3), (STA_2, STA_4), (STA_5, STA_6)</td>
</tr>
<tr>
<td>Clustering 4</td>
<td>(STA_1, STA_3), (STA_2, STA_6), (STA_4, STA_5)</td>
</tr>
<tr>
<td>Clustering 5</td>
<td>(STA_1, STA_6), (STA_2, STA_5), (STA_3, STA_4)</td>
</tr>
</tbody>
</table>

is representative for realistic networks, where for example STAs are distributed at different locations and hence have different path loss. Our goal is to find out a clustering method which leads to the best possible aggregate throughput.

**Same Cluster Size:**

First we assume that the number of STAs per cluster is the same for all clusters. We begin our study by considering a constructed scenario, in which only three types of STAs each with a certain channel variance exist. It is assumed that there is no link adaptation and different clusters may have different error rates but they all transmit at the same PHY rate. Furthermore, it is assumed that all clusters have the same size. We try all possible clustering combinations to find out the configuration which leads to the highest aggregate throughput.

For the numerical analysis, \( n = 6 \) STAs are considered which are grouped into three clusters. An uplink scenario in the infrastructure network is considered where the AP has two antennas and hence the cluster size is set to two. If all STAs have different channel variances, there exist \( N_k \) possible ways of grouping STAs into indistinguishable clusters of the size \( N_{nc} \):

\[
N_k = \frac{1}{N_{c!}} \prod_{i=1}^{N_{c}} \left( n - (i-1)N_{nc} \right) = \frac{1}{N_{c!}} \left( \frac{n!}{(N_{nc})^{N_{c}}} \right) = \frac{1}{3!} \left( \begin{array}{c} 6 \\ 2 \end{array} \right) \left( \begin{array}{c} 4 \\ 2 \end{array} \right) \left( \begin{array}{c} 2 \\ 2 \end{array} \right) = 15.
\] (7.26)

In our case since each pair of STAs have the same channel variance, there are only five distinguishable ways of grouping them. A subcarrier channel variance of 0.25, 1, and 4 are assumed for the first pair of STAs (STA_1 and STA_2), the second pair (STA_3 and STA_4) and the third one (STA_5 and STA_6), respectively. This model represents a scenario where each two STAs have the same path loss.

We simulate the outage probability for all possible clusterings for different values of \( P/\sigma_n^2 \) in all STAs. We use the outage probabilities to calculate the aggregate throughput for each case. In this section, we assume that channel outages are the only source of channel errors. This corresponds to having either high SNR at receivers or strong channel coding.
Chapter 7  Cross-Layer Enhancements of CB-CSMA/CA

Figure 7.10: Aggregate Throughput for different ways of clusterings, STA₁ and STA₂ have the lowest channel variance, STA₃ and STA₄ the middle value and STA₅ and STA₆ the highest value.

The backoff Model I, cf. Section 7.1 is used. The five distinguishable ways of clusterings are shown in Table 7.1 and the throughput results are shown in Figure 7.10. The results show that in the low SNR regime, the highest aggregate throughput in the considered scenario can be achieved when STAs with strong channels are allocated to the same cluster, cf. Clustering 1 and 3. Consequently, the stronger STAs, which are grouped into the same cluster, benefit from higher error probabilities of the other clusters and hence lower transmission probabilities of those clusters. In this SNR regime, the performance of the STAs is degraded if each of those STAs is placed with a STA with lower channel variance in the same cluster. In the medium SNR regime, clustering methods 1 and 2 outperform the rest. Here, both types of clusters with high and medium channel variance perform better if the STAs with the worst channels are allocated to the same group. In the high SNR regime where \( P_{\text{out}} \) approaches zero, as the number of STAs per cluster is the same in all clusterings, they all achieve the same throughput.

The results show the potential of considering channel conditions in the clustering. It also shows that it is beneficial to allocate STAs with similar channel conditions into the same cluster. In order to have a closer look at the impact of the channel on clustering, we take a more realistic case into account as described in the following paragraphs.
Different Cluster Sizes:

So far, the same number of STAs in all source- and destination-clusters has been assumed. However, in practical systems there could be a different number of STAs per cluster, for example due to the absence of traffic at a certain source or entering/exiting STAs to/from a BSS.

Cluster size influences the throughput in two ways: it determines the maximum number of concurrent streams, and it can impact the link error probability and hence the transmission probability. The latter occurs if for example the cluster size is smaller than the maximum number of simultaneous streams which can be decoded by the receiver. In this case, the additional degrees of freedom can be used to enhance the link performance. Consequently, higher PHY rate or lower channel error probability can be achieved.

In this section we study the throughput of a realistic scenario in which STAs are distributed randomly in a certain area. Each pair of STAs forms a source-destination link via an AP with two antennas which acts as a DF relay. STAs are grouped into clusters. Each cluster may include either one or two STAs. We assume a unidirectional data communication from each source to its destination. The AP forwards, the received packets after a SIFS. Hence only source-clusters compete for the channel. The outage probability is simulated as explained for Application A in Chapters 5 and 6. However, now the number of antennas at the AP, $N_a$, and consequently the maximum number of STAs per cluster is reduced to two. Besides, instead of four orthogonal training sequences only two of them are needed. Accordingly a shorter header is applied. Since each STA may have different pathloss and due to the possibility of having clusters with one or two STAs there are $N_k$ possible clustering combinations where:

$$N_k = \sum_{i=1}^{N_{\text{conf}}} \frac{1}{N_c^{(1)}!N_c^{(2)}!...N_c^{(N_{\text{nc}})\max}!} \prod_{i=1}^{N_c} \left( n - \sum_{j=1}^{i-1} N_{\text{nc}}(j) \right) \frac{n!}{\prod_{i=1}^{N_{\text{nc}}} N_{\text{nc}}(i)!}$$

(7.27)

where $N_c^{(m)}$ is the number of clusters with $m$ STAs, $N_{\text{nc}}(i)$ is the number of STAs in cluster $i$th, and $N_{\text{conf}}$ is the number of possible configurations where each configuration leads to a
different number of clusters. For 6 STAs and a maximum cluster size of 2 we have:

\[
N_k = \frac{1}{3!} \binom{6}{2} \binom{4}{2} + \frac{1}{2!2!} \binom{6}{2} \binom{4}{2} \binom{2}{1} + \frac{1}{1!4!} \binom{6}{2} \binom{4}{1} \binom{3}{1} \binom{2}{1} \binom{1}{1} + \frac{1}{6!} \binom{6}{1} \binom{5}{1} \binom{4}{1} \binom{3}{1} \binom{2}{1} \binom{1}{1} = 76.
\] (7.28)

In the considered scenario there are four possible configurations: three source clusters each with two STAs, four clusters two of them with two STAs and the other two with one member, five clusters one of them with two STAs and the other four each with one STA, and six clusters each with a single STA.

We have assumed a square area with dimension of 50 m by 50 m. The sources and their respective destinations are placed at random locations within this area. The relay takes different positions on the x-axis: \(x_r = -40 : 5 : 40\) [m]. Position of sources and destinations are shown in Figure 7.11. For each location of the relay we simulate the outage probability for all possible clustering combinations. Then the throughput is calculated using the backoff model I and the simulated outage probabilities. For each possible configuration the maximum aggregate throughput is found by exhaustive search and depicted in Figure 7.12. A data rate of 19.5 Mb/s per hop and a basic rate of 6.5 Mb/s has been assumed. As it is seen in this figure, choosing the right configuration can improve the performance significantly.
7.2 Impact of Clustering on CB-CSMA/CA Performance

Figure 7.12: Maximum aggregate throughput vs. position of the relay on the x-axes.

However, the results in Figure 7.12 are obtained by exhaustive search which has a prohibitive complexity. Besides, in practice we usually cannot search all possible clusterings to find the best combination. Hence, it is desirable to define an algorithm for choosing a proper combination.

In this part we suggest a heuristic way to constitute clusters in the considered scenario. We take into account the same setup as before with $N_{cc} \leq 2$. In the previous scenarios it has been observed that putting STAs which have similar characteristic into the same cluster leads to better performance.

Here, we apply an algorithm to first find out the appropriate number of clusters and then search for the STAs which their distances from the relay are in the same range. In order to do this we take the following steps:

1. For a given position of the relay, distances between all STAs and the relay are calculated.

2. We find out if there is any SD pair which has either relatively large or small $d_{sr} + d_{rd}$ value, where $d_{sr}$ is the distance between source and relay and $d_{rd}$ is the distance between destination and relay. We refer to such SD pairs as "isolated" pairs. In order to find such pairs we compare $d_{sr} + d_{rd}$ value of each SD pair with a certain minimum and maximum threshold. The thresholds can be defined according to the position of the STAs and the environment dimension.
3. Since the transmission is performed over two hops, for each pair of SD, we consider the maximum of the uplink and downlink distances, $d_{\text{max}_i}$ as our metric.

4. If no isolated SD pair has been found, we constitute clusters with $N_{\text{nc}} = 2$ STAs. To do so we allocate first the STAs with the smallest $d_{\text{max}_i}$ into the same cluster, then we find the ones with the second most smallest $d_{\text{max}_i}$ and we repeat this step until all STAs are grouped.

5. If some isolated SD pairs have been found, they are allocated to the single-member clusters. Then we check if any of other SD pairs, has a small $d_{\text{max}_i}$. If such SD pair exists, it also forms a single-member cluster. For the remaining SD pairs, we again search for the pairs which have the smallest $d_{\text{max}_i}$ and we allocate them to the first group. We continue clustering STAs in this way until all STAs are allocated to the clusters. It should be noted that destinations of each source-cluster are also allocated into the same cluster.

In this scenario, during the uplink phase STAs within a cluster transmit simultaneously to the relay. It should be noted that in a multiple access phase, PHY layer performance of one STA impacts the other STA and vice versa. Besides, in such scenario the STA with a longer transmission duration, determines the total time the channel remains busy during the transmission. Therefore, allocating STAs with similar characteristic into one cluster or separating the isolated ones can enhance the aggregate throughput.

**Figure 7.13:** Maximum aggregate throughput vs. position of the relay on the x-axes.
7.3 Summary

For the setup shown in Figure 7.11, we select the appropriate configuration according to the above algorithm and calculate the throughput. The results are shown in Figure 7.13, where the optimum aggregate throughput found by exhaustive search is also plotted. For each position of the relay, we also pick a configuration randomly and plot its aggregate throughput. As it is shown, the heuristic method performs relatively good compared to the optimum case and in most cases better than the randomly selected case.

In general the parameters chosen for the algorithm can be set according to the setup and PHY signal processing. For example if the performance of a relaying scenario is mainly limited by its uplink, one can take only the uplink distances into account. In addition, the threshold(s) used to find out the isolated SD pairs, impact the number of clusters and accordingly throughput. Hence, one should try to find out the appropriate value for them. The considered heuristic method is simple and depends only on positions of the STAs and the relay. Although, it does not provide us with a general solution which can be directly applied to different setups, it shows the potential of the clustering for throughput enhancement.

7.3 Summary

In this chapter, we have investigated possible improvements of the CB-CSMA/CA by taking cross-layer information into account. We have studied the impact of the backoff model and different clustering techniques on the network and individual throughput. Furthermore, we have proposed how to choose the appropriate backoff model or clustering method in different situations in order to enhance the throughput.

In this chapter we have assumed that STAs in a cluster are all perfectly synchronised and all of them have always a packet ready to send. In the next chapter we analyse the performance bounds of the CB-CSMA/CA protocol in the presence of event-synchronisation errors. Furthermore, we compare the performance of the CB-CSMA/CA protocol with another protocol which also supports multipacket transmissions, under saturated as well as non-saturated conditions.
Chapter 8

Performance Bounds of CB-CSMA/CA

So far perfect event-synchronisation has been assumed in all clusters. As it has been pointed out in Section 5.1.4, event-synchronisation may fail for several reasons. In this chapter, we relax the assumption of the perfect event-synchronisation and analyse the throughput performance of CB-CSMA/CA in the presence of event-synchronisation errors. Specifically we consider a worst case scenario where all STAs in all clusters are event-asynchronous and compare the CB-CSMA/CA throughput with that of the CSMA/CA.

In the second section we compare the CB-CSMA/CA throughput with another multipacket protocol which has been proposed in [32]. We consider both saturated as well as unsaturated case.

8.1 Impact of Event-Synchronisation Errors on CB-CSMA/CA Performance

So far we have assumed perfect event-synchronisation in all clusters. Hence, in the saturated case each time when a cluster accesses the medium, all of its members transmit concurrently. As it has been mentioned in Chapter 5, the members of a cluster may become event-asynchronous when for example one or more of them cannot receive or decode the CWUR packet. This may happen if any of the members is in deep fade.

In the presence of an event-synchronisation error a subset of STAs in a cluster may be silent while the others are transmitting. Consequently, the number of parallel streams is not the same as the cluster size anymore.

In this section we study the impact of event-synchronisation errors on the throughput performance of the CB-CSMA/CA. It is assumed that the cluster has $N_{ncj}$ STAs. We begin the
analysis by considering a symmetric scenario where all clusters suffer from synchronisation errors in the same way. Then we consider an asymmetric setup, where only a certain cluster, $C_i$, out of $N_c$ clusters has a non-zero probability of event-synchronisation error, $P_{se_i}$. Finally, we extend the scenario to a general case, in which different clusters have different event synchronisation probabilities. For all scenarios, we investigate the throughput equations as a function of $P_{se}$.

### 8.1.1 Assumptions and Model

In this section we consider uplink of an infrastructure network where the AP has four antennas and each source-cluster contains four single-antenna STAs. We assume that all clusters can support a link data rate of 58.5 Mb/s regardless of the number of parallel streams. This could be the case in a high SNR regime where in spite of presence of multiuser interference this high rate could be supported on individual links. To focus on the impact of event-synchronisation errors on the numerical results, in this section we set $P_e$ in the numerical results to 0.

As it has been explained in Section 5.1.4, event-synchronisation error can originate from different types of errors, including decoding errors on control frames and hidden nodes. In this chapter we assume that event-synchronisation errors occur due to decoding errors. Hence, the original backoff model can still be applied. In CB-CSMA/CA even if all STAs in all clusters become event-asynchronous, the conditional collision probability is reduced compared to the CSMA/CA case. This is due to the fact that simultaneous transmissions from the same cluster can still be decoded.

It should be noted that in CB-CSMA/CA the preambles are defined in such a way that only multiple streams originated from a single cluster can be decoded. On the other hand, if for example two STAs, each from one cluster, begin to transmit simultaneously, the training sequences are not orthogonal any more and hence the receiver cannot estimate the channel and decode the packets. This happens even if the receiver has multiple antennas. Therefore, it is assumed that collision occurs if at least two stations which belong to different clusters begin to transmit at the same time regardless of the number of the parallel streams.

### 8.1.2 Symmetric Scenario

We begin the analysis with a simple symmetric scenario, where all clusters have the same number of asynchronous STAs. Assuming each cluster has a total number of $N_{nc}$ STAs
8.1 Impact of Event-Synchronisation Errors on CB-CSMA/CA Performance

from which, \( k \) STAs are event-asynchronous. In this way, all clusters see the same number of contending units at each transmission attempt, i.e., \((k + 1)(N_c - 1)\) for \( k < N_{nc} \) and \( k(N_c - 1)\) for \( k = N_{nc} \). We distinguish between two cases in this scenario: (i) \( k = N_{nc} \) and (ii) \( k < N_{nc} \).

**case (i) - \( k = N_{nc} \):** First we assume that all STAs in clusters encounter synchronisation error. We study an extreme case where \( P_{se} \) is assumed to be equal to 1, (i.e., all members of clusters are with probability of one asynchronous). In this case, collision happens if one or more STAs in a cluster transmit while any other STA from any other cluster transmits at the same time. It should be noted that even in this case simultaneous transmissions of multiple STAs within a cluster can be resolved and hence do not lead to collision. The conditional collision probability can be calculated by taking into account the collision probability in an equivalent CSMA/CA, cf. (3.10) with \( n = N_cN_{nc} \), and then subtract the probability that any other node from the same cluster transmits while other clusters are silent:

\[
P_{\text{col}} = 1 - (1 - \tau)^{N_cN_{nc} - 1} - \sum_{i=1}^{N_{nc} - 1} \binom{N_{nc} - 1}{i} (1 - \tau)^{N_{nc} - 1 - i} (1 - \tau)^{N_{nc} - 1 - i}.
\]

(8.1)

\( P_{\text{col}} \) can also be directly obtained by taking into account the total number of contenders, i.e., the number of contending STAs which belong to other clusters:

\[
P_{\text{col}} = 1 - (1 - \tau)^{(N_c - 1)N_{nc}}.
\]

(8.2)

As expected, for the same parameters, equations (8.1) and (8.2) lead to the same result.

Since all members of clusters are asynchronous, the transmission probability of each cluster is: \( P_t = 1 - (1 - \tau)^{N_{nc}} \). Consequently, the aggregate throughput of all clusters is defined by:

\[
S = \frac{N_c \sum_{m=1}^{k} \binom{k}{m} \tau^m (1 - \tau)^{(k-m)mL}}{T_{\text{slot}}} \times (1 - P_{\text{col}})(1 - P_e),
\]

(8.3)

where \( k = N_{nc} \) and \( T_{\text{slot}} \) can be obtained using:

\[
T_{\text{slot}} = P_{\text{id}}\sigma + P_s(1 - P_e)T_s + (1 - P_{\text{id}} - P_s)T_e + P_e P_s T_e,
\]

(8.4)


**Figure 8.1:** Throughput for CB-CSMA/CA and CSMA/CA. In CB-CSMA/CA all STAs suffer from event-synchronisation error with probability $P_{se} = 1$

and $P_s$ and $P_{id}$ can be calculated from the following equations:

$$
P_s = N_c P_t (1 - P_{col})
$$

$$
P_{id} = (1 - \tau)^{N_{nc} N_c}
$$

As it is expected by setting $N_{nc}$ to 1 and $N_c$ to be equal to the number of STAs, $n$, (8.3) reduces to the throughput equation in the standard case, i.e., (3.14).

Throughput in the case where all STAs encounter synchronisation error with probability of one is plotted in Figure 8.1. We also depict throughput of a reference system based on the standard CSMA/CA. Contrary to the setups in Chapter 6, here it is assumed that both systems transmit at the same data rate. This is a sensible assumption since even for CB-CSMA/CA with high probability only one node transmits at each transmission attempt. Besides, in this chapter we focus on high SNR regime where the CB-CSMA/CA can also operate at the highest PHY data rate defined by the standard regardless of the number of parallel streams.

It can be observed that, when all STAs are event-asynchronous the CB-CSMA/CA throughput is reduced to that of the reference system. It should be noted that even in this scenario the collision probability of the CB-CSMA/CA is smaller than that of the standard CSMA/CA, however the headers in CB-CSMA/CA are longer than that in the reference system. For large network size, this leads to slightly smaller throughput in CB-CSMA/CA as
8.1 Impact of Event-Synchronisation Errors on CB-CSMA/CA Performance

![Figure 8.2: Comparison of normalised throughput results obtained using the model with those from simulations, $P_{se} = 1$](image)

compared to that of the standard CSMA/CA, cf. Figure 8.1 (b).

In order to validate the results obtained from the model, we have simulated the same scenario in MATLAB. The normalised throughput obtained from simulations are compared with those obtained from the analytical model in Figure 8.2. As it can be observed results from the model match the simulation results. The model becomes more accurate as the network size gets larger.

**Case (ii) - $k < N_{nc}$**: Similar to the previous case for asynchronous STAs, simultaneous transmissions of multiple STAs within a cluster may only happen accidentally. The conditional collision probability given that one or more STAs from a cluster transmit can be obtained from the following equation:

$$P_{col} = 1 - (1 - \tau)^{(k+1)(N_{c}-1)}.$$  

(8.6)

For each $k$ and $N_c$ the transmission probability can be obtained by solving numerically (3.9) and (8.6). In each cluster $N_{nc} - k$ STAs transmit at the same time while $k$ STAs transmit individually. We denote the first group of STAs by group $a$ and the latter one by group $b$. In each cluster, for a given $k$ the transmission probability of each group is defined
Figure 8.3: Throughput per STA for CB-CSMA/CA where k STAs from each cluster suffer from event-synchronisation error with probability \( P_{se} = 1 \) by:

\[
P_t^a = \tau \\
P_t^b = 1 - (1 - \tau)^k
\]  

(8.7)

When group \( a \) transmits, there are \( N_{nc} - k \) parallel transmissions, whereas STAs in group \( b \) transmits independently from each other. In this case the throughput can be calculated as follows:

\[
S = \frac{N_c \left[ P_{ta}(1 - P_{tb})(N_{nc} - k)L + (1 - P_{ta})\left(\sum_{m=1}^{k} \binom{k}{m} \tau^m(1 - \tau)^{(k-m)}mL\right) + P_{ta}\left(\sum_{m=1}^{k} \binom{k}{m} \tau^m(1 - \tau)^{(k-m)}(m + N_{nc} - k)L\right)\right]}{T_{slot}} \times (1 - P_{col})(1 - P_e) 
\]

(8.8)

where \( T_{slot} \) can be calculated using (8.4) and \( P_s \) and \( P_{id} \) are respectively given by:

\[
P_s = N_c \left[ P_{ta}(1 - P_{tb}) + P_{tb}(1 - P_{ta}) + P_{ta} P_{tb} \right] (1 - P_{col}) \\
P_{id} = (1 - \tau)^{k+1} N_c
\]

(8.9)
We set $N_{nc}$ to four and calculate throughput when $k$ STAs in each cluster are event-asynchronous. The results for $k = 0, 1, 2$ and $3$ are plotted in Figure 8.3. The throughput degrades as the values of $k$ increases. As it is expected the throughput in this case with $k = N_{nc} - 1 = 3$ and in the previous case are the same.

### 8.1.3 Asymmetric Scenario

In practice different clusters may have different channel conditions and hence we extend the study to cover asymmetric scenarios. In this section, we consider an asymmetric scenario where only one cluster suffers from synchronisation error. In this case we need to distinguish between two different categories of clusters: i) Category $i$ which has only one cluster, i.e., $C_i$. Each STA in $C_i$ has a non-zero $P_{se}$ independent from other STAs. ii) The second category includes all other clusters with $P_{se} = 0$. We denote the clusters in this category by $C_j$ where $j \neq i$.

Since the clusters in the latter category are all perfectly event-synchronised, the members of $C_i$ face collision only when $C_i$ begins to transmit and any other cluster transmits at the same time. Therefore, the collision probability of $C_i$ when $k$ of its members are asynchronous is defined by:

$$P_{\text{col}}^{(k)} = 1 - (1 - \tau_j^{(k)})N_{c} - 1.$$  \hspace{1cm} (8.10)

Assuming the backoff procedure model I or III, then the relation between $\tau_i$ and $P_{\text{col}}^{(k)}$ is defined by (3.9) and (7.15), respectively. However, for any $C_j$ in the second category, collision happens if it transmits and at the same time any other cluster from the same category or any of the STAs in $C_i$ begin to transmit. Accordingly, for this category collisions occur with different probabilities, depending on the number of clusters as well as the number of asynchronous STAs within the $C_i$.

$$P_{\text{col}}^{(k)} = \begin{cases} 1 - (1 - \tau_j^{(k)})N_{c} - 2(1 - \tau_i^{(k)}) & \text{w.p. } (1 - P_{se})^{N_{nci}}, \quad k = 0 \\ 1 - (1 - \tau_j^{(k)})N_{c} - 2(1 - \tau_i^{(k)})^2 & \text{w.p. } \left(\frac{N_{nci}}{1}\right)P_{se} (1 - P_{se})^{N_{nci} - 1}, \quad k = 1 \\ 1 - (1 - \tau_j^{(k)})N_{c} - 2(1 - \tau_i^{(k)})^3 & \text{w.p. } \left(\frac{N_{nci}}{2}\right)P_{se}^2 (1 - P_{se})^{N_{nci} - 2}, \quad k = 2 \\ \vdots & \\ 1 - (1 - \tau_j^{(k)})N_{c} - 2(1 - \tau_i^{(k)})^{N_{nci}} & \text{w.p. } \left(\frac{N_{nci}}{N_{nci} - 1}\right)P_{se}^{N_{nci} - 1} (1 - P_{se}), \quad k = N_{nci} - 1 \\ 1 - (1 - \tau_j^{(k)})N_{c} - 2(1 - \tau_i^{(k)})^{N_{nci}} & \text{w.p. } \left(\frac{N_{nci}}{N_{nci} - 1}\right)P_{se}^{N_{nci} - 1}, \quad k = N_{nci} \end{cases}$$  \hspace{1cm} (8.11)

For each $k$, equations (3.9), (8.10) and (8.11) can be solved numerically.
For each transmission attempt of $C_i$, there are different numbers of parallel streams. The number of parallel streams depends on the number of asynchronous STAs within $C_i$ and whether they transmit by chance at the same time or not. To calculate throughput of cluster $C_i$ we first need to determine the probability that $k$ STAs out of $N_{NC_i}$ STAs in $C_i$ are asynchronous, i.e.:

$$P_r(k) = \binom{N_{NC_i}}{k} P_{SE_i}^k (1 - P_{SE_i})^{N_{NC_i} - k}. \quad (8.12)$$

For a given $k$, similar to the previous case, the STAs in $C_i$ are divided into two groups:

i) first group consists of $N_{NC_i} - k$ synchronised STAs, these nodes access the channel at the same time and act as a single unit. The transmission probability of this group is denoted by $P_{ta}$.

ii) the second group consists of $k$ asynchronous members, these STAs access to the channel individually however, some of them, may by chance transmit at the same time. We denote the transmission probability of this group by $P_{tb}$.

For a given $k$ we have:

$$P_{ta}^{(k)} = \tau_i^{(k)}$$

$$P_{tb}^{(k)} = \sum_{m=1}^{k} \binom{k}{m} (\tau_i^{(k)})^m (1 - \tau_i^{(k)})^{k-m} = 1 - (1 - \tau_i^{(k)})^k \quad (8.13)$$

The probability that only STAs in $C_i$ transmit is:

$$P_{s_i}^{(k)} = \left[P_{ta}^{(k)} (1 - P_{tb}^{(k)}) + P_{tb}^{(k)} (1 - P_{ta}^{(k)}) + P_{ta}^{(k)} P_{tb}^{(k)} \right] (1 - P_{col}^{(k)})$$

$$= \left[1 - (1 - P_{ta}^{(k)}) (1 - P_{tb}^{(k)}) \right] \left(1 - P_{col}^{(k)} \right). \quad (8.14)$$

The probability that only one of the clusters from the second category transmits is:

$$P_{s_j}^{(k)} = (N_c - 1) \tau_j^{(k)} (1 - P_{col}^{(k)}). \quad (8.15)$$

Throughput of cluster $C_i$ and aggregate throughput of all other clusters are respectively ex-
8.1 Impact of Event-Synchronisation Errors on CB-CSMA/CA Performance

Figure 8.4: Throughput per cluster when each STA in $C_i$ suffers from event-synchronisation error with probability $P_{se_i}$

pressed as:

$$S_i = \sum_{k=0}^{N_{nci}} P_r^{(k)} P_{se_i}^{(k)} (N_{nci} - k) L + P_t^{(k)} (1 - P_{t_a}^{(k)}) mL + P_t^{(k)} (1 - P_{t_b}^{(k)}) (m + N_{nci} - k) L$$

$$\times (1 - P_{col}^{(k)})(1 - P_{e_i})$$

$$S_j = \sum_{k=0}^{N_{nci}} P_r^{(k)} P_{se_i}^{(k)} N_{ncj} L$$

$$\times (1 - P_{e_j})$$

$$T_{slot} = \left(1 - P_{id}^{(k)} \right) \left( \sum_{k=0}^{N_{ncj}} P_r^{(k)} P_{s_j}^{(k)} \right) T_s$$

$$+ \left(1 - P_{e_j}^{(k)} \right) \left( \sum_{k=0}^{N_{ncj}} P_r^{(k)} P_{s_j}^{(k)} \right) T_e + P_c T_c$$

(8.16)

(8.17)
where $P_{id}$ is the probability that the slot is idle, and $P_c$, the probability that collision occurs in the slot, are respectively given by:

$$P_{id} = \sum_{k=0}^{N_{sei}} P_r^{(k)} (1 - \tau_j^{(k)})^{N_c - 1} (1 - P_{ta}^{(k)})(1 - P_{ts}^{(k)})$$

$$P_c = 1 - P_{id} - \left( \sum_{k=0}^{N_{sei}} P_r^{(k)} P_{si}^{(k)} \right) - \left( \sum_{k=0}^{N_{sei}} P_r^{(k)} P_{sj}^{(k)} \right). \quad (8.18)$$

Throughput of each cluster category is shown in Figure 8.4. As the number of clusters increases, throughput of $C_i$ improves with increase in $P_{sei}$. When $P_{sei}$ increases, with high probability there are more asynchronous STAs in $C_i$. This increases only the collision probability of other clusters and hence $C_i$ may transmit with higher probability and benefits from longer backoff durations at other clusters. For the same reason throughput of any other cluster degrades when $P_{sei}$ increases.

The aggregate throughput versus number of clusters for different value of $P_{sei}$ has been depicted in Figure 8.5. The aggregate throughput decreases when $P_{sei}$ increases. However, as it is expected, the throughput values for different $P_{sei}$ get close to each other for large number of clusters. The impact of a single cluster $C_i$ almost disappears once the number of

\[\text{Figure 8.5: Aggregate throughput when each STA in } C_i \text{ suffers from event-synchronisation error with probability } P_{sei}.\]
clusters becomes very large.

### 8.1.4 General Scenario

In this section, we will consider a general and more realistic case, where each STA at any cluster may suffer from synchronisation error with a certain probability. It is assumed that each cluster $C_i$ has $N_{nci}$ STAs out of which $k_i$ are asynchronous. For each cluster, the collision probability depends on the number of total units, i.e., asynchronous and synchronous STAs, in other clusters. Consequently, for $0 \leq k_i \leq N_{nci}$ where $i \in \{0, ..., N_{nci}\}$ we have:

$$P_{\text{col}}^{(k_1, k_2, ..., k_{nc})} = 1 - \prod_{j=1, j \neq i}^{N_c} (1 - \tau_j^{(k_1, k_2, ..., k_{nc})})^{\bar{n}_j} \bar{\bar{n}}_j = \begin{cases} 1 + k_j & \text{for } k_j \leq N_{ncj} \\ k_j & \text{for } k_j = N_{ncj} \end{cases} \quad (8.19)$$

For any number of asynchronous STAs in each cluster the transmission probability can be obtained by solving the set of equations given in (8.19) and (7.15). Accordingly, transmission probability of synchronous and asynchronous STAs in each $C_i$ can be calculated respectively, as follows:

$$P_{t_{ai}}^{(k_1, k_2, ..., k_{nc})} = \tau_i^{(k_1, k_2, ..., k_{nc})}$$

$$P_{t_{bi}}^{(k_1, k_2, ..., k_{nc})} = \sum_{m=1}^{k_i} \binom{k_i}{m} \left( \tau_i^{(k_1, k_2, ..., k_{nc})} \right)^m (1 - \tau_i^{(k_1, k_2, ..., k_{nc})})^{k_i - m}. \quad (8.20)$$

From now on, for simplicity we denote the subscript $(k_1, k_2, ..., k_{nc})$ by $(\tilde{k})$. The probability that $k_i$ STAs, out of $N_{nci}$ STAs are asynchronous can be obtained from (8.12) however here $k$ may take different values in different clusters. The probability that only members of one cluster transmit also depends on $(k_1, k_2, ..., k_{nc})$ and can be obtained as follows:

$$P_{s_i} = \sum_{k_1=0}^{N_{nc1}} \sum_{k_2=0}^{N_{nc2}} \cdots \sum_{k_{nc}=0}^{N_{ncnc}} P_{t_{ni}}^{k_1} P_{t_{ni}}^{k_2} \cdots P_{t_{ni}}^{k_{nc}} \left[ P_{t_{ai}}^{(\tilde{k})} (1 - P_{t_{bi}}^{(\tilde{k})}) + P_{t_{bi}}^{(\tilde{k})} (1 - P_{t_{ai}}^{(\tilde{k})}) + P_{t_{ai}}^{(\tilde{k})} P_{t_{bi}}^{(\tilde{k})} \right] \times (1 - P_{\text{col}}^{(\tilde{k})}). \quad (8.21)$$
Chapter 8 Performance Bounds of CB-CSMA/CA

The aggregate throughput of each clusters is given by:

\[ S_i = \sum_{k_1=0}^{N_{nc1}} \sum_{k_2=0}^{N_{nc2}} \cdots \cdots \sum_{k_{nc}=0}^{N_{nc}} P_{k_1}^{(k_1)} (1 - P_{k_1}) (N_{nc} - k_i) L_i \]

\[ + (1 - P_{k_1}) \left( \sum_{m=1}^{k_1} \binom{k_1}{m} \tau_{i}^{(k_1-m)} (1 - \tau_{i}^{(k_1)}) m L_i \right) \]

\[ + P_{k_1}^{(k_1)} \left( \sum_{m=1}^{k_1} \binom{k_1}{m} \tau_{i}^{(k_1-m)} (1 - \tau_{i}^{(k_1)}) (m + N_{nc} - k_i) L_i \right) \times (1 - P_{k_1}^{(k_1)}) (1 - P_{k_1}) \]

\[ T_{\text{slot}} \]

(8.22)

where \( T_{\text{slot}} \) can be calculated from the following equations:

\[ T_{\text{slot}} = P_{id} \sigma + \left( \sum_{i=1}^{N_c} P_{s_i} (1 - P_{e_i}) \right) T_s + \left( \sum_{i=1}^{N_c} P_{s_i} P_{e_i} \right) T_e + P_c T_c, \]  

(8.23)

\[ P_{id} = \sum_{k_1=0}^{N_{nc1}} \sum_{k_2=0}^{N_{nc2}} \cdots \sum_{k_{nc}=0}^{N_{nc}} P_{k_1}^{k_1} P_{k_2}^{k_2} \cdots P_{k_{nc}}^{k_{nc}} (1 - \tau_{1}^{(k_1)}) \hat{u}_1 (1 - \tau_{2}^{(k_2)}) \hat{u}_2 \cdots (1 - \tau_{nc}^{(k_{nc})}) \hat{u}_{nc} \]

with \( \hat{u}_j = \begin{cases} 1 + k_j & \text{for } k_j \leq N_{nc} \\ k_j & \text{for } k_j = N_{nc} \end{cases} \)

\[ P_c = 1 - P_{id} - \left( \sum_{i=1}^{N_c} P_{s_i} \right). \]

(8.24)

Applying the above equations, we evaluate throughput of the general scenario where \( P_{se_i} \) of each cluster is randomly chosen, \( P_{se_i} \in [0, 1] \). Each time the throughput of each cluster and the aggregate throughput are calculated. Then \( P_{se_i} \) values are set to new random values. The results are shown in Figure 8.6. Although, throughput is degraded compared to the case where all STAs are event-synchronous, it is still much higher than that of the CSMA/CA system.

### 8.2 Comparison with the Multipacket Reception Protocol

In [32] and [116] the authors have proposed a multipacket reception (MPR) protocol and investigated its performance, for various types of networks and parameters. In [32] uplink transmissions in an infrastructure network are considered, where the AP has \( N_a \) antennas.
Figure 8.6: Throughput of the general case where STAs in each cluster $C_i$ suffer from event-synchronisation error with randomly chosen probability $P_{se_i}$

STAs compete for the channel according to DCF RTS/CTS mechanism. However, if accidentally more than one STA transmits in a time slot and the AP can decode the RTS frames, it sends the CTS frame to the senders. This can be done as long as the number of concurrent transmissions are less or equal to $N_a$. Afterwards the transmitters send their data packets simultaneously to the AP. The proposed MAC closely follows the standard MAC, however some modifications are required. For example as the AP does not have any a priori knowledge about the transmitters and their channels, it should apply blind techniques to decode the RTS frames. Furthermore it has to allocate orthogonal training sequences to the transmitters once it sends the CTS frames.

In this section we compare throughput values for the CB-CSMA/CA and that of MPR. To start with we consider the saturation throughput and then we analyse performance in non-saturation mode for different numbers of STAs. In the latter case, we distinguish between two types of clusterings: adaptive clustering where STAs which have packets to transmit are allocated to the same clusters and non-adaptive clustering where clusters are formed regardless of the presence of queued packets. As it has been explained in Section 5.1.6 adaptive clustering can be achieved by using the polling information in the CFP.
Table 8.1: The analysis parameters which are used in this section.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>CB-CSMA/CA</th>
<th>MPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIFS</td>
<td>16 $\mu$s</td>
<td>16 $\mu$s</td>
</tr>
<tr>
<td>DIFS</td>
<td>34 $\mu$s</td>
<td>34 $\mu$s</td>
</tr>
<tr>
<td>Propagation delay $\delta$</td>
<td>1 $\mu$s</td>
<td>1 $\mu$s</td>
</tr>
<tr>
<td>$CW_{\text{min}}$</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>$CW_{\text{max}}$</td>
<td>1023</td>
<td>1023</td>
</tr>
<tr>
<td>Slot time $\sigma$</td>
<td>9 $\mu$s</td>
<td>9 $\mu$s</td>
</tr>
<tr>
<td>Payload size $L$</td>
<td>1024 Byte</td>
<td>1024 Byte</td>
</tr>
<tr>
<td>Basic rate</td>
<td>6.5 Mb/s</td>
<td>6.5 Mb/s</td>
</tr>
<tr>
<td>Data rate</td>
<td>58.5 Mb/s</td>
<td>58.5 Mb/s</td>
</tr>
<tr>
<td>MAC Header (Data)</td>
<td>40 Byte</td>
<td>34 Byte</td>
</tr>
<tr>
<td>MAC Header (ACK, CTS)</td>
<td>14 Byte</td>
<td>20 Byte</td>
</tr>
<tr>
<td>RTS MAC$_{H}$</td>
<td>20 Byte</td>
<td>20 Byte</td>
</tr>
</tbody>
</table>

8.2.1 Saturation Throughput

In this part we consider an uplink transmission in an infrastructure network where the AP has $N_a$ antennas. The MPR headers are taken from [32]$^1$, however the data rate, SIFS and DIFS durations are set to the ones defined in the IEEE 802.11n. In the MPR the PHY header is set to 26 $\mu$s, while in CB-CSMA/CA application we have used the PHY headers as defined in Chapter 6 for application A. Table 8.1 shows the parameters used in numerical analysis.

To start with we consider a saturation mode. We set $N_a$ once to 2 and another time to 4. The throughput of both systems versus number of STAs in the network is depicted in Figure 8.7. It can be observed that the CB-CSMA/CA achieves much higher throughput than MPR scenario. This is because in CB-CSMA/CA, at each transmission attempts $N_a$ STAs transmit while in MPR there are $N_a \leq N_a$ concurrent transmissions. The larger the $N_a$ is the higher the throughput gain of the CB-CSMA/CA as compared to the MPR.

In [32] it has been shown that there is a maximum throughput value for MPR. In order to compare both systems, throughput versus the probability that a node transmits in a randomly chosen time slot, $\tau$, is plotted for both CB-CSMA/CA and MPR. The results for $N_a = 2$ are shown in Figure 8.8 and the ones for $N_a = 4$ are plotted in Figure 8.9. As expected when there are only $N_a$ STAs in the network, the CB-CSMA/CA achieves a much higher throughput than the MRP for all values of $\tau$. Besides, for both $N_a$ values, the maximum throughput of CB-CSMA/CA is higher than the maximum throughput of MPR. However, the

$^1$The parameters in this paper are given for 2 antennas at the AP, nevertheless we use them for the case with 4 antennas too.
8.2 Comparison with the Multipacket Reception Protocol

![Graph showing throughput comparison between CB-CSMA/CA and MPR vs. number of STAs]

**Figure 8.7:** Throughput of the CB-CSMA/CA and MPR vs. number of STAs for $N_a=2$, 4 antennas at the AP.

CB-CSMA/CA reaches its maximum value at lower $\tau$. In case of no channel error, for the IEEE 802.11n parameters we have $\tau \leq 0.1176$. In this regime CB-CSMA/CA considerably outperforms the MPR.

Slightly different behaviour can be observed for $N_a = 2$. This occurs since the two considered protocols have different overheads and transmission procedures.

Similar to MPR, for CB-CSMA/CA applications, if we are allowed to choose different backoff parameters from those defined by the IEEE 802.11 standard, we can adjust the parameters such that the maximum throughput can be achieved.

The above values are obtained for perfect event-synchronisation. However, we also consider a scenario with event-synchronisation error. In this scenario in CB-CSMA/CA application, all STAs are event-synchronous with probability $P_{se}$. Figure 8.10 shows the results for $P_{se} = 0.1, 0.25, \text{and } 0.33$. As it is seen, using the parameters given in Table 8.1, for small to medium number of STAs, the CB-CSMA/CA outperforms the MPR but for large values of $P_{se}$ and large network size the MPR outperforms the CB-CSMA/CA. This shows that even in the presence of low event-synchronisation error, it is beneficial to apply CB-CSMA/CA for small networks. However, the CB-CSMA/CA throughput degrades as the $P_{se}$ becomes large.
Figure 8.8: Aggregate throughput vs. $\tau$ for CB-CSMA/CA and MPR with $N_a = 2$.

Figure 8.9: Aggregate throughput vs. $\tau$ for CB-CSMA/CA and MPR with $N_a = 4$.  

188
8.2 Comparison with the Multipacket Reception Protocol

Figure 8.10: Throughput of the CB-CSMA/CA and MPR vs. number of STAs for $N_a = 4$ antennas at the AP. In the CB-CSMA/CA application each STA faces event-synchronisation with the probability of $P_{se}$.

8.2.2 Non-saturation Throughput

In non-saturated networks STAs may or may not have a packet to transmit, depending on the traffic arrival rate. In order to calculate the transmission probability under non-saturated condition, we can apply the MC model as proposed in [8]. In this model, compared to the original model explained in Chapter 3, a new state is introduced, which indicates the probability that there is at least one packet to be transmitted in the buffer. This probability is denoted by $q$. As it is shown in [8], in this case the transmission probability can be obtained from:

$$
\tau = \frac{2(1 - 2p)q}{q[(1 - 2p)(W + 1) + pW(1 - (2p)^m)] + 2(1 - q)(1 - p)(1 - 2p)}; \quad (8.25)
$$

Assume $\lambda$ as the average packet arrival rate at each STA. For a traffic model with Poisson packet arrival process and for small buffer size the probability $q$ can be calculated from [117]:

$$
q = 1 - e^{-\lambda T_{\text{slot}}}; \quad (8.26)
$$

where $T_{\text{slot}}$ can be obtained from (3.15).
CB-CSMA/CA with Non-Adaptive Clustering:

Here we consider a scenario where clusters are defined independently of the presence of packets in their queues. Let us assume a worst case scenario where a cluster begins to compete for the channel as soon as all of its members have packets ready to send. Although in practice members of a cluster should be able to transmit even if other members remain idle, this scenario shows an extreme inefficient way of clustering under non-saturated condition. Accordingly, for the CB-CSMA/CA throughput analysis we have to replace $q$ in (8.25) by $q^{N_{nc}}$.

Non-saturation throughput of CB-CSMA/CA protocol with the considered non-adaptive clustering is depicted in Figure 8.11. For comparison purpose, throughput of the MPR as well as that of the IEEE 802.11n DCF basic access mechanism are also shown. At $q = 1$ each STA has a non-empty buffer with probability one and the system is saturated. It can be observed that the CB-CSMA/CA achieves higher throughput above a certain threshold. For very low values of $q$, CB-CSMA/CA throughput is below that of the other systems. This happens since clusters are defined independently of the packet arrival rate and each cluster transmits only if all of its members have packets to transmit. The results in Figure 8.11 are calculated for $n = 24$.

The CSMA/CA achieves slightly higher throughput for very low value of $q$ as compa-
8.3 Summary

In this chapter first we have investigated the CB-CSMA/CA performance bounds in the presence of event-synchronisation errors. The analysis of event-synchronisation errors, shows that the CB-CSMA/CA is relatively robust to the event-synchronisation error and in the worst

![Graph showing non-saturation throughput of the adaptive CB-CSMA/CA, the MPR and the CSMA/CA vs. probability of having non-empty buffer, for Na=2 antennas at the AP and n = 24.](image)

**Figure 8.12:** Non-saturation throughput of the adaptive CB-CSMA/CA, the MPR and the CSMA/CA vs. probability of having non-empty buffer, for Na=2 antennas at the AP and n = 24.

to MPR. While MPR is performed using RTS/CTS handshake, the CSMA/CA is based on the basic access mechanism. Hence, in the region where collision probability is small CSMA/CA benefits from smaller overheads.

**CB-CSMA/CA with Adaptive Clustering:**

Here we assume an adaptive clustering method. According to this method, for each transmission attempt, we select STAs which have packets in their buffer. In this way, assuming n STAs and Nnc = 2, for a given STA we can find at least one other STA, which has a packet ready to transmit, with probability \(1 - (1 - q)^{(n-1)}\).

Throughput results for this scenario with \(n = 24\) is depicted in Figure 8.12. It can be observed that the CB-CSMA/CA protocol with adaptive clustering outperforms both the MPR and CSMA/CA for most values of \(q\).
case it performs similar to the standard CSMA/CA. In the second section we have studied the CB-CSMA/CA throughput having different backoff parameters, and compared it with another multipacket reception protocol. Both saturation and non-saturation throughput are calculated. The CB-CSMA/CA outperforms both MPR and CSMA/CA in saturation as well as non-saturation mode with medium and high probability of non-empty buffers. However, to benefit from CB-CSMA/CA for low packet arrival rates, we should apply an adaptive clustering. The adaptive clustering takes the presence of packets in the STAs into account and it can be performed by using the information obtained in the polling phase, as it has been explained in Section 5.1.6. Furthermore, the probability of event-synchronisation errors should be kept as small as possible.

Throughout this thesis we have focused on single- and two-hop links, which are typical patterns in infrastructure and many ad hoc networks. However, in mesh or sensor networks a source may communicate with a destination over multihop links. In the next chapter, we will deal with multihop links.
Chapter 9

Outlook for Multihop Communications

So far we have focused on single- and at most two-hop communication links. These links are typical ones in infrastructure and most ad hoc scenarios. The fast growth of WLAN deployments has led to many different configurations such as mesh and sensor networks. A survey on mesh networks is presented in [118] and a survey on sensor networks is given in [119]. In a mesh network, each node may act as a router. Mesh networking has already been considered for WLANs [120–122]. A mobile mesh network, called mobile ad hoc network (MANET), has also been widely considered in the literature [123, 124]. Wireless sensor network is another type of network which has also been investigated in different papers [125, 126]. In a wireless sensor network, nodes are usually spatially distributed and they monitor some physical or environmental parameters.

In the above-mentioned networks and many other types of networks, a source and destination may be separated by several hops. In this chapter we focus only on multihop links (more than two hops).

In this chapter, first the main differences between multihop links with single hop links are presented. Then some of the existing multihop techniques are briefly described. Finally, we discuss the possibility of extending the CB-CSMA/CA to support multihop transmissions, the challenges and the required modifications.

9.1 Conceptual Discussion

To study multihop communication systems we have to take into consideration different aspects, assumptions and new features as compared to that of single- and two-hop links such as:
Chapter 9  Outlook for Multihop Communications

- As soon as we have more than one possible path between a source and its destination, the routing becomes important. For example if we can redirect a packet and find a new path so that it bypasses an intensively loaded relay, we can avoid the bottleneck problem. This may also help to avoid congestion and buffer overflows.

- So far we have assumed that all STAs belong to the same BSS and they are all able to hear each other, hence the spatial reuse factor has been equal to one. However, in many multihop networks we need to direct the traffic over multihops due to the large physical distance between the source and its destination. In such networks, finding an smart way to increase the space reuse factor can help us to use the resources more efficiently.

- If the same packet is forwarded by several intermediate nodes, the end-to-end delay may become quite large. An appropriate channel access method and scheduling could help to keep the delay as small as possible.

- For throughput analysis, the saturation assumption may not be true anymore even if sources are saturated. As it is shown in [84] the arrival packet rate at an intermediate node, along a path is determined by the maximum service times of preceding nodes.

- Control frames including ACK transmission need to be adapted. In multihop links, the ACK can be an end-to-end ACK or it can be sent per hop.

Some of these points have been referred to by the IEEE task group s. The IEEE 802.11s draft standard specifies the MAC and PHY layers for mesh networking [120]. It defines new protocols including hybrid wireless mesh protocol (HWMP), mesh beacon collision avoidance (MBCA) and mesh coordination function (MCF). The first protocol is defined as the default path selection protocol providing both proactive and the reactive path selections1. Once the mesh path of a particular SD pair is found, it is propagated by the mesh STAs in the whole BSS. Since in a mesh network beacon frames are transmitted by multiple STAs, which may even be located out of communication range of each other, beacon frames may face the hidden node problem. Therefore, the IEEE 802.11s specifies MBCA as a mechanism to avoid collisions among beacon frames [120].

The IEEE 802.11s specifies mesh BSS (MBSS) as a BSS which forms a network of mesh STAs. Each MBSS can be used as a distribution system. Figure 9.1 shows an example of such networks. As it is seen in this figure the mesh BSS can form a whole or part of the DS using the wireless medium [120].

1Reactive routing protocols search for a path between two nodes only if there is a data packet to send while proactive routing protocols search and define paths between nodes regardless of the presence of data packet.
Mesh STAs use mesh coordination function to access the channel [120]. MCF consists of EDCA, which is the access method defined by the IEEE 802.11e, and MCF controlled channel access (MCCA). MCCA is a reservation-based access method, which can be used optionally by mesh STAs. According to this access method, mesh STA uses management frame to make a reservation for transmission in advance. The reservation is advertised by the mesh STA among its neighbours using MCCA reports. All other MCCA capable STAs which their transmission can interfere with the transmission during the reserved time period, do not transmit during the reserved time period. The mesh STA needs to obtain a TXOP by winning EDCA contention procedure. In this way, at reserved time periods, mesh STAs can access the channel with lower collision probability, since they do not face any competition from other MCCA capable neighbouring mesh STAs.

According to the IEEE 802.11s, Mesh STAs can use ACK, explicit ACK, and Block ACK procedure as defined by the IEEE 802.11e. However, there is another type of ACK which has been proposed for multihop communications [127]. This ACK, known as passive or implicit ACK, can be utilised due to the broadcast nature of wireless medium as explained in this paragraph. When a STA forwards a packet to the next immediate receiver it indirectly acknowledges reception of that packet to the previous transmitter. In this way, there is no
Figure 9.2: An example of a network where source and its destination communicates over \( m \) hops.

![Figure 9.2](image1)

Figure 9.3: Ratio of overhead due to ACK to data frame size for per hop ACK (dashed lines) and for passive ACK (solid lines).

![Figure 9.3](image2)

We consider a source S and destination D which communicates via \( m - 1 \) relays over \( m \) hops, cf. Figure 9.2. We compare the overhead of passive ACK with that of the per hop ACK transmission. To do so, the ACK overhead used for transmission of a data packet from S to D is calculated and divided by the duration of a data frame. The MAC and PHY header values are taken from Table 3.3. The duration of data and ACK frames are obtained from (6.1) and (6.2), respectively. The data payload size varies from 128 to 1024 Byte. ACK frames are sent at 6.5 Mb/s. Two cases are taken into consideration: in the first case data packets are transmitted at 6.5 Mb/s and in the second case data packets are sent at 58.5 Mb/s. Results for both cases are shown in Figure 9.3.

As it is shown in this figure the ACK overhead, when ACK is transmitted per hop, grows linearly with number of hop. However, for passive ACK, the overhead does not increase with number of hops since only the last hop needs to transmit the ACK. For both cases, the ratio of ACK overhead to the data packet size is higher for higher data rate (since in both cases...
ACK rate is set to the lowest rate) and smaller payload size.

Similar idea can be applied in multihop multicast/broadcast setups. According to the IEEE 802.11 standard, there is no ACK transmission available for multicast and broadcast transmissions [5]. However, if a node is equipped with \( N_a > 1 \) antennas it can decode \( N_a \) independent ACK frames simultaneously.

In addition to applications of sensor and mesh networks, there are other wireless multihop applications which are proposed in the literature. MUZFR over multihops [128], MIMO tunnelling using AF relays and multipath transmission using DF relays [129], are some examples of such applications. In the next section we consider the CB-CSMA/CA protocol and discuss extensions of CB-CSMA/CA required to support multihop applications.

9.2 A Specific Proposal

In order to support multihop communication links, using the CB-CSMA/CA protocol, we need to reconsider some of our assumptions. For example the assumption that both source and destination clusters belong to the same BSS can easily be violated in multihop scenarios. In this section we consider a WLAN where members of the same cluster are in the communication range of each other, however the source- and destination- clusters may be out of the radio range of each other.

We distinguish between two groups of multihop setups:

- The first group is a general setup, where any STA in the network can act as an intermediate node and forward a packet towards a destination (either directly or via some other intermediate nodes). The mesh networks in Figure 9.1 is an example of such networks.

- The second group is a more structured setup where only dedicated relays, which are allocated to multiple stages, forward data packets for source-clusters. In this way, the relay stages form multihop links.

Adapting CB-CSMA/CA to support the general case, i.e., first group, requires further investigation and is beyond the scope of this thesis. For example, routing and traffic pattern can highly impact the cluster establishment in such networks. At the same time, an efficient way of clustering may lead to smaller collision probability and larger spatial reuse factor. Although different aspects of clusterings have widely been considered in other publications,
Figure 9.4: An example of a multihop network where source and its destination communicates over \( m - 1 \) DF relay-stages.

some clustering issues of CB-CSMA/CA are protocol-related and should be analysed exclusively. Here, we briefly study the second group and discuss the possible extensions of CB-CSMA/CA to support these types of applications.

Let us study the network shown in Figure 9.4 where two source-destination pairs communicate via several DF relay-stages. Source \( l \) transmits packet \( P_l \) to destination \( l \) over \( m \) hops. In each relay-stage, there are two relays which decode the packets received from the preceding stage and forward them to the next immediate stage. In [129] achievable rates for a similar multipath network but with one source-destination pair is studied. Here, we assume that each STA has a single antenna while relays are equipped with \( N_a \geq 2 \) antennas and hence at each stage the relay on the first path can cancel the interference from the second path and vice versa. In this chapter, we focus on the MAC protocol and spatial reuse factor.

The following assumptions are taken into consideration:

- The same transmit power per sources and relays are assumed.
- On each path distance between two neighbours (including source or relays) is \( d \).
- The relays are used for coverage extension. The radio range of each node is larger than \( d \) but smaller than \( 2d \).
- Each relay forwards a packet only if it can decode it.

We allocate nodes on the same stage to a single cluster. In this way, according to CB-CSMA/CA access method, they can transmit concurrently. In the example shown in Figure 9.4 sources and destinations are allocated to clusters \( C_s \) and \( C_d \), respectively while relays on stage \( i \)th are grouped into clusters \( C_{r}^{i} \). We exploit the back-propagate packet, i.e., the received packet at stage \( i \) from the transmission at stage \( i + 1 \), as the passive ACK. Assuming that all errors can be detected, then relays either forward the correctly decoded packets or become silent when an error is detected. In the latter case, the preceding node retransmits the packet. Accordingly, retransmissions only use one hop.
Two cases are taken into consideration: i) \( N_a = 2 \) and ii) \( N_a = 4 \).

**First Case: \( N_a = 2 \)**

In this case, both antennas at each relay are needed to cancel the interference from other path and decode the intended stream. Therefore, the same nodes cannot decode the back-propagate signal at the same time. In this case either we should apply explicit ACK frames or reduce the spatial reuse factor. In the latter case, for the considered setup, only clusters with a distance of at least three hops, i.e., clusters on a stage \( i \)th and \( j \)th where \( j \geq i + 3k \), \( 0 \leq i + 3k < m \) and where \( k \) is a non-negative integer, can transmit concurrently.

**Second Case: \( N_a = 4 \)**

In this case, there is enough degrees of freedom to decode both direct and back-propagate packets at the same time. Hence, clusters with at least two-hop distances, i.e., clusters on stages \( i \)th and \( j \)th where \( j \geq i + 2k \), can transmit simultaneously. For example, sources transmit new packets, while their previously sent packets are forwarded by the relays on the second stage to the third stage. Each relay on the first stage receives sum of four signals: the two new packets sent by two sources and the back-propagated packets forwarded by the two relays on the second stage. However, as there are four antennas at each relay, the relays on the first stage can decode all received packets with high probability. They can use the decoded back-propagated packets as the passive ACK by comparing them with those they have forwarded before.

In Figure 9.5 we show this setup for \( m = 3 \) hops and 5 time slots. The \( n \)th data packet which is transmitted by the \( l \)th source is denoted by \( P_n^l \). It should be noted that in this example, we simply assume that clusters access the channel one after the other. However, in each IBSS clusters access the channel according to the CB-CSMA/CA contention procedure or, as it has been briefly presented in Section 5.1.6, using the CB-CSMA/CA contention-free centralised procedure.

In this example we benefit from additional antennas at the relays to reduce number of required time slots for forwarding data packets. The same concepts can be applied using an additional frequency band. A similar approach can be used for more than two source-destinations, provided that there are enough degrees of freedom to decode the intended data packets. The source-cluster size is defined by the available degrees of freedom at relays and by the ACK method. If concurrent transmissions of several sources cannot be efficiently handled by the relays, then they should be allocated to more than one cluster.

In such networks where several IBSS are involved in transmission of a certain packet from a source to its destination, the contention procedure can be performed in each IBSS. If the
nodes belonging to more than one IBSS, do not have enough degrees of freedom to decode streams from several clusters, then inter-IBSS communications are needed. In the above example, where \( N_a = 2 \) this is needed to avoid simultaneous packet receptions at the relay-cluster on the first stage from the source-cluster and the relay-cluster on the second stage. This can easily be handled by exchanging RTS/CTS frames prior to data transmission.

For real-time applications, delay should be small. For such applications it is recommendable to keep the number of hops as small as possible, and to have a larger spatial reuse factor. In order to guarantee that a certain relay-cluster can access the channel at a certain time, contention-free channel access methods are preferable. In this case, in each IBSS a cluster-master forms clusters and handles the scheduling.

In the considered network, we simply allocated nodes on the same stage to one cluster. In practice forming clusters can be more complicated. A simple way is to allocate nodes which are physically close to each other into the same clusters. However, to enhance performance some other features like available degrees of freedom at each node, application requirements, channel conditions, etc. should be taken into consideration.

---

**Figure 9.5:** An example of a multihop network where two source-destination pairs communicate over two DF relay-stages. Each source and relay has two and four antennas, respectively.
9.3 Summary

In this chapter, multihop communication links have been considered. After a short review of the related IEEE 802.11 standard, possibility of CB-CSMA/CA extension to support multihop links have been discussed. In order to support generic multihop networks, further investigations are required. However, a method has been proposed for a simple multihop network, where multiple dedicated relays assist the communication between a source and its destination.
Chapter 10

Summary and Conclusion

The IEEE 802.11n specifies MIMO techniques to enhance data rate in WLANs. However, it can support only MIMO point-to-point links. On the other hand, it is known that multiuser MIMO techniques can significantly increase the spectral efficiency of networks.

In this thesis, to realise MIMO spatial multiplexing gain in networks with single-antenna STAs and to enhance network performance, cooperative protocols are considered with respect to both PHY and MAC layers. Integration of three different cooperative relaying scenarios into WLAN have shown the potential gain in PHY layer. Three representative cooperative applications have been studied in details: multiuser MIMO DF relaying, two-way DF relaying and multiuser zero-forcing AF relaying. The first application can be considered as a typical application in infrastructure networks. The PHY layer of the second application can also be easily integrated into the existing systems. However, the third application has been considered as a future application.

In order to utilise such signalling in WLANs, the MAC protocol has to be modified. The existing CSMA/CA-based MAC avoids any interference and supports only point-to-point links. To solve this problem and enhance IEEE 802.11n system in a way that multiple users can transmit simultaneously, a novel cluster-based CSMA/CA has been proposed. We have explained the CB-CSMA/CA protocol in details and presented the impact of the PHY layer on the MAC protocol design. Additionally, we have discussed some of the challenges confronted to realise such a multiuser MIMO system.

Throughput results have shown significant gain of CB-CSMA/CA over IEEE 802.11n. Despite three times larger link rate for a reference system operating based on IEEE 802.11n, the CB-CSMA/CA, even with a GI twice as long as in the reference system, shows about two times larger throughput in an infrastructure network. The study has shown that the CB-CSMA/CA can significantly improve throughput and reduce delay in different types of infrastructure or ad hoc networks with single-hop or two-hop links.
Chapter 10 Summary and Conclusion

The throughput of CB-CSMA/CA, has further been improved by taking cross-layer parameters into account, when forming the clusters or defining the backoff models. Study has shown that in many situations allocating STAs with similar requirements or conditions, such as channel conditions, average SNR and packet size, into the same cluster is advantageous.

The CB-CSMA/CA protocol is based on synchronous transmissions of group of STAs. This synchronisation may be impaired in practice due to different reasons. The analytical study of cases with event-synchronisation errors has shown that the CB-CSMA/CA is robust. In the worst case scenario, where all clusters encounter even-synchronisation error with probability of one, performance of CB-CSMA/CA approaches that of the IEEE 802.11n system.

Furthermore, throughput comparisons between the CB-CSMA/CA and a multiuser MIMO protocol have been made. It has been shown that the CB-CSMA/CA outperforms the other protocol in both saturation and non-saturation cases except in a low loaded regime. However, even there, one could achieve higher throughput by defining the clusters in an adaptive way. Furthermore, numerical analysis has shown that a maximum throughput by adjusting the backoff parameters can be reached.

The analyses have shown that the CB-CSMA/CA is a promising approach which can be applied in different types of wireless networks.

The focus of this thesis has been on networks with single- or two-hop communication links, which are the most typical ones in infrastructure and ad hoc networks. On the other hand, with increase in demand for wireless communications other types of configurations have also received a lot of attention. Wireless mesh and sensor networks are two examples of such configurations. In Chapter 9 outlook for multihop communications and possible extensions of CB-CSMA/CA to support multihop links have been presented. Further investigations are needed to apply this protocol efficiently to a generic multihop scenario.

Furthermore, it is expected that upcoming wireless systems operate in very high frequency bands such as 60 GHz. Moving towards higher frequencies brings in new issues. In this work, simulations have been done for NLoS situations. As much larger path loss is expected at 60 GHz band compared to 5 GHz, one needs to take into account LoS situations. Besides, in this work we have referred only to indoor applications. For outdoor scenarios mobility should be considered too.
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-MPDU</td>
<td>Aggregate MAC Protocol Data Unit</td>
</tr>
<tr>
<td>A-MSDU</td>
<td>Aggregate MAC Service Data Unit</td>
</tr>
<tr>
<td>AC</td>
<td>Access Category</td>
</tr>
<tr>
<td>ACK</td>
<td>Acknowledgement</td>
</tr>
<tr>
<td>ADPCM</td>
<td>Adaptive Differential Pulse Code Modulation</td>
</tr>
<tr>
<td>AF</td>
<td>Amplify-and-Forward</td>
</tr>
<tr>
<td>AIFS</td>
<td>Arbitration Interframe Space</td>
</tr>
<tr>
<td>AP</td>
<td>Access Point</td>
</tr>
<tr>
<td>APEP</td>
<td>Average Pairwise Error Probability</td>
</tr>
<tr>
<td>ARQ</td>
<td>Automatic Repeat Request</td>
</tr>
<tr>
<td>AWGN</td>
<td>Additive White Gaussian Noise</td>
</tr>
<tr>
<td>BPSK</td>
<td>Binary Phase Shift Keying</td>
</tr>
<tr>
<td>BS</td>
<td>Base Station</td>
</tr>
<tr>
<td>BSS</td>
<td>Basic Service Set</td>
</tr>
<tr>
<td>CB</td>
<td>Cluster-based</td>
</tr>
<tr>
<td>CB-CSMA/CA</td>
<td>Cluster-based Carrier Sense Multiple Access with Collision Avoidance</td>
</tr>
<tr>
<td>CF</td>
<td>Contention-Free</td>
</tr>
<tr>
<td>CFP</td>
<td>Contention-Free Period</td>
</tr>
<tr>
<td>CP</td>
<td>Contention Period</td>
</tr>
<tr>
<td>CRC</td>
<td>Cyclic Redundancy Check</td>
</tr>
<tr>
<td>CSIT</td>
<td>Channel State Information at Transmitter</td>
</tr>
<tr>
<td>CSMA</td>
<td>Carrier Sense Multiple Access</td>
</tr>
<tr>
<td>CSMA/CA</td>
<td>Carrier Sense Multiple Access with Collision Avoidance</td>
</tr>
<tr>
<td>CTS</td>
<td>Clear-to-Send</td>
</tr>
<tr>
<td>CW</td>
<td>Contention Window</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>CWUR</td>
<td>Contention Window Update Request</td>
</tr>
<tr>
<td>DA</td>
<td>Destination Address</td>
</tr>
<tr>
<td>DAS</td>
<td>Distributed Antenna System</td>
</tr>
<tr>
<td>DCF</td>
<td>Distributed coordination function</td>
</tr>
<tr>
<td>DECT</td>
<td>Digital Enhanced Cordless Telecommunications</td>
</tr>
<tr>
<td>DF</td>
<td>Decode-and-Forward</td>
</tr>
<tr>
<td>DIFS</td>
<td>DFC Interframe Space</td>
</tr>
<tr>
<td>DLS</td>
<td>Direct-Link Setup</td>
</tr>
<tr>
<td>DS</td>
<td>Distribution System</td>
</tr>
<tr>
<td>EDCA</td>
<td>Enhanced Distributed Channel Access</td>
</tr>
<tr>
<td>ETSI</td>
<td>European Telecommunications Standards Institute</td>
</tr>
<tr>
<td>FCS</td>
<td>Frame Check Sequence</td>
</tr>
<tr>
<td>FEC</td>
<td>Forward Error Correction</td>
</tr>
<tr>
<td>FER</td>
<td>Frame Error Ratio</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast Fourier Transform</td>
</tr>
<tr>
<td>GI</td>
<td>Guard Interval</td>
</tr>
<tr>
<td>HC</td>
<td>Hybrid Coordinator</td>
</tr>
<tr>
<td>HCCA</td>
<td>HCF Controlled Channel Access</td>
</tr>
<tr>
<td>HCF</td>
<td>Hybrid Coordination Function</td>
</tr>
<tr>
<td>HoL</td>
<td>Head-of-Line</td>
</tr>
<tr>
<td>HT</td>
<td>High Throughput</td>
</tr>
<tr>
<td>HWMP</td>
<td>Hybrid Wireless Mesh Protocol</td>
</tr>
<tr>
<td>IBSS</td>
<td>Independent Basic Service Set</td>
</tr>
<tr>
<td>IEEE</td>
<td>Institute of Electrical and Electronics Engineers</td>
</tr>
<tr>
<td>IP</td>
<td>Internet Protocol</td>
</tr>
<tr>
<td>IP</td>
<td>Internet protocol</td>
</tr>
<tr>
<td>LoS</td>
<td>Line of Sight</td>
</tr>
<tr>
<td>Acronyms</td>
<td>Description</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>MAC</td>
<td>Medium Access Control</td>
</tr>
<tr>
<td>MANET</td>
<td>Mobile Ad hoc Network</td>
</tr>
<tr>
<td>MBCA</td>
<td>Mesh Beacon Collision Avoidance</td>
</tr>
<tr>
<td>MBSS</td>
<td>Mesh Basic Service Set</td>
</tr>
<tr>
<td>MC</td>
<td>Markov Chain</td>
</tr>
<tr>
<td>MCCA</td>
<td>Mesh Coordinated Channel Access</td>
</tr>
<tr>
<td>MCF</td>
<td>Mesh Coordination Function</td>
</tr>
<tr>
<td>MCS</td>
<td>Modulation and Coding Schemes</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple-Input Multiple-Output</td>
</tr>
<tr>
<td>MISO</td>
<td>Multiple-Input Single-Output</td>
</tr>
<tr>
<td>MPDU</td>
<td>MAC Protocol Data Unit</td>
</tr>
<tr>
<td>MSDU</td>
<td>MAC Service Data Unit</td>
</tr>
<tr>
<td>MUIC</td>
<td>Multiuser Interference Cancellation</td>
</tr>
<tr>
<td>MUZFR</td>
<td>Multiuser Zero Forcing Relaying</td>
</tr>
<tr>
<td>NLoS</td>
<td>None Line of Sight</td>
</tr>
<tr>
<td>OFDM</td>
<td>Orthogonal Frequency Division Multiplexing</td>
</tr>
<tr>
<td>PAL</td>
<td>Protocol Adaptation Layer</td>
</tr>
<tr>
<td>PBX</td>
<td>Private Branch eXchange</td>
</tr>
<tr>
<td>PCF</td>
<td>Point Coordination Function</td>
</tr>
<tr>
<td>PER</td>
<td>Packet Error Rate</td>
</tr>
<tr>
<td>PHY</td>
<td>Physical Layer</td>
</tr>
<tr>
<td>PLCP</td>
<td>Physical Layer Convergence Procedure</td>
</tr>
<tr>
<td>PPDU</td>
<td>PLCP Protocol Data Unit</td>
</tr>
<tr>
<td>PSMP</td>
<td>Power Save Multi-Poll</td>
</tr>
<tr>
<td>QAM</td>
<td>Quadrature Amplitude Modulation</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>QPSK</td>
<td>Quadrature Phase Shift Keying</td>
</tr>
<tr>
<td>RSSI</td>
<td>Receive Signal Strength Indicator</td>
</tr>
<tr>
<td>RTS</td>
<td>Real-Time Transport Protocol</td>
</tr>
<tr>
<td>RTS</td>
<td>Request-to-Send</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>SA</td>
<td>Source Address</td>
</tr>
<tr>
<td>SD</td>
<td>Source-Destination</td>
</tr>
<tr>
<td>SER</td>
<td>Symbol Error Ratio</td>
</tr>
<tr>
<td>SIFS</td>
<td>Short Interframe Space</td>
</tr>
<tr>
<td>SIMO</td>
<td>Single-Input Multiple-Output</td>
</tr>
<tr>
<td>SINR</td>
<td>Signal to Interference and Noise Ratio</td>
</tr>
<tr>
<td>SISO</td>
<td>Single-Input Single-Output</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal to Noise Ratio</td>
</tr>
<tr>
<td>STA</td>
<td>Station</td>
</tr>
<tr>
<td>STBC</td>
<td>Space-Time Block Code</td>
</tr>
<tr>
<td>TDMA</td>
<td>Time Division Multiple Access</td>
</tr>
<tr>
<td>TG</td>
<td>Task Group</td>
</tr>
<tr>
<td>TXOP</td>
<td>Transmission Opportunity</td>
</tr>
<tr>
<td>UDP</td>
<td>User Datagram Protocol</td>
</tr>
<tr>
<td>VoIP</td>
<td>Voice over IP</td>
</tr>
<tr>
<td>WF</td>
<td>Water-filling</td>
</tr>
<tr>
<td>WINDECT</td>
<td>Wireless LAN with Integration of Professional-Quality DECT</td>
</tr>
<tr>
<td>WLAN</td>
<td>Wireless Local Area Network</td>
</tr>
</tbody>
</table>
# Notation

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$BW$</td>
<td>Bandwidth</td>
</tr>
<tr>
<td>$BpS$</td>
<td>Byte per OFDM symbol</td>
</tr>
<tr>
<td>$C$</td>
<td>Capacity</td>
</tr>
<tr>
<td>$CW_{\text{max}}$</td>
<td>Maximum contention window size</td>
</tr>
<tr>
<td>$CW_{\text{min}}$</td>
<td>Minimum contention window size</td>
</tr>
<tr>
<td>$D$</td>
<td>Delay</td>
</tr>
<tr>
<td>$L$</td>
<td>Data packet size</td>
</tr>
<tr>
<td>$M_r$</td>
<td>Number of antennas at the receiver</td>
</tr>
<tr>
<td>$M_t$</td>
<td>Number of antennas at the transmitter</td>
</tr>
<tr>
<td>$NF$</td>
<td>Noise figure</td>
</tr>
<tr>
<td>$N_c$</td>
<td>Number of clusters</td>
</tr>
<tr>
<td>$N_{nc}$</td>
<td>Number of nodes per cluster</td>
</tr>
<tr>
<td>$N_{sd}$</td>
<td>Number of source-destination pairs</td>
</tr>
<tr>
<td>$N_{sub}$</td>
<td>Number of data subcarriers</td>
</tr>
<tr>
<td>$P_{\text{out}}$</td>
<td>Outage probability</td>
</tr>
<tr>
<td>$R$</td>
<td>Data rate</td>
</tr>
<tr>
<td>$S$</td>
<td>Throughput</td>
</tr>
<tr>
<td>$T_0$</td>
<td>Temperature in Kelvin</td>
</tr>
<tr>
<td>$T_c$</td>
<td>Average time the channel is sensed busy due to collision</td>
</tr>
<tr>
<td>$T_e$</td>
<td>Average time the channel is sensed busy due to channel error</td>
</tr>
<tr>
<td>$T_s$</td>
<td>Average time the channel is sensed busy due to successful transmission</td>
</tr>
<tr>
<td>$T_{\text{ACK}}$</td>
<td>ACK transmission duration</td>
</tr>
<tr>
<td>$T_{\text{Data}}$</td>
<td>Data transmission duration</td>
</tr>
<tr>
<td>$T_{\text{GI}}$</td>
<td>Guard interval duration</td>
</tr>
<tr>
<td>$T_{\text{sym}}$</td>
<td>OFDM symbol duration</td>
</tr>
<tr>
<td>$A$</td>
<td>Matrix $A$</td>
</tr>
</tbody>
</table>
### Notation

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A^H$</td>
<td>Hermition of matrix $A$</td>
</tr>
<tr>
<td>$A^T$</td>
<td>Transpose of matrix $A$</td>
</tr>
<tr>
<td>$H$</td>
<td>Matrix of channel coefficients</td>
</tr>
<tr>
<td>$I_n$</td>
<td>The $n \times n$ identity matrix</td>
</tr>
<tr>
<td>$a$</td>
<td>Vector $a$</td>
</tr>
<tr>
<td>$a^T$</td>
<td>Transpose of vector $a$</td>
</tr>
<tr>
<td>$</td>
<td>x</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Propagation delay</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Path loss exponent</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Wavelength</td>
</tr>
<tr>
<td>$\lceil \cdot \rceil$</td>
<td>Ceiling function</td>
</tr>
<tr>
<td>$\lfloor \cdot \rfloor$</td>
<td>Floor function</td>
</tr>
<tr>
<td>$\mathbb{E}[\cdot]$</td>
<td>Expectation operator</td>
</tr>
<tr>
<td>$|A|_F^2$</td>
<td>Squared Frobenius norm of $A$</td>
</tr>
<tr>
<td>$\odot$</td>
<td>Hadamard (element-wise) product</td>
</tr>
<tr>
<td>$P_{col}$</td>
<td>Conditional collision probability</td>
</tr>
<tr>
<td>$P_e$</td>
<td>Packet error probability</td>
</tr>
<tr>
<td>$P_{id}$</td>
<td>Probability that the channel is idle</td>
</tr>
<tr>
<td>$P_{se}$</td>
<td>Probability of synchronisation error</td>
</tr>
<tr>
<td>$P_s$</td>
<td>Probability that only one node transmits in a slot</td>
</tr>
<tr>
<td>$P_{tr}$</td>
<td>Probability that at least one node transmits in a slot</td>
</tr>
<tr>
<td>$Pr$</td>
<td>Probability</td>
</tr>
<tr>
<td>$\det[A]$</td>
<td>Determinant of matrix $A$</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Duration of a time slot</td>
</tr>
<tr>
<td>$\sigma_n^2$</td>
<td>Noise variance</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Transmission probability</td>
</tr>
<tr>
<td>$a$</td>
<td>Scalar $a$</td>
</tr>
<tr>
<td>$d$</td>
<td>Distance</td>
</tr>
<tr>
<td>$k$</td>
<td>Boltzmann’s constant</td>
</tr>
<tr>
<td>$n$</td>
<td>Number of stations</td>
</tr>
<tr>
<td>$p$</td>
<td>Transition probability in Markov chain model</td>
</tr>
</tbody>
</table>
Bibliography


[38] Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, Amendment 4: High-speed Physical Layer in the 5 GHz Band, Further Higher Data Rate Extension in the 2.4 GHz Band, IEEE Std. 802.11g, 2003.


[47] A. Ettefagh, M. Kuhn, and A. Wittneben, “Relaying to enhance the coverage range of a voice over WLAN system,” International ITG Workshop on Smart Antennas, WSA, 2006.


Publications


