Carsten Burmeister

Evaluation and Optimization of Wireless Internet Access at Hot Spots

Cuvillier Verlag Göttingen

Evaluation and Optimization of Wireless Internet Access at Hot Spots

Vom Promotionsausschuss der
Technischen Universität Hamburg-Harburg
zur Erlangung des akademischen Grades
Doktor-Ingenieur
genehmigte Dissertation

von
Carsten Burmeister
aus Hamburg

2007

Bibliografische Information Der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.ddb.de abrufbar.
1. Aufl. - Göttingen : Cuvillier, 2007
Zugl.: (TU) Hamburg-Harburg, Univ., Diss., 2007

978-3-86727-351-0

1. Gutachter: Prof. Dr. Ulrich Killat

2. Gutachter: Prof. Dr.-Ing. Klaus Jobmann

Tag der mündlichen Prüfung: 20. April 2007

Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de

1. Auflage, 2007
Gedruckt auf säurefreiem Papier

978-3-86727-351-0

Contents

List of Figures

Chapter 1

Introduction

The Internet has evolved to the main source of information and to a highly efficient communication means for millions of people all over the world. High bandwidth Internet access enables users to download World Wide Web and multi-media content and communicate via e-mails, Internet chat, voice or even video. By means of wireless communications this access is possible on the move. However, wireless bandwidth is a scarce resource and must be shared by all users that are located in some proximity. Especially in areas of high user density, so called Hot Spots, the existing resources must be used most efficiently to satisfy all communication demands.

Wireless communication technologies that enable Internet access can be split into two categories, namely cellular networks and Wireless Local Area Networks (WLANs). In cellular networks the mobile stations connect directly to base stations, which provide the access to the fixed network. The transmission power can be quite high leading to large coverage of a base station. The coverage of the network is further increased by allowing the mobile station to handover between base stations. The 2nd generation of cellular networks, which is represented by GSM systems in Europe, provides mainly voice communication and very low rate data access. On the way to the next generation of mobile communication systems, GPRS being a packet switched network was the first step in the direction of cellular Internet access. Data rates of up to 100kbps allow to retrieve World Wide Web content and communicate via e-mail. Multi-media communication is possible with the 3rd generation of mobile networks, namely UMTS in Europe and Asia. With a data rate of up to 2Mbps per cell, users could possibly access the Internet with a comparable Quality of Service (QoS) as via a fixed line at their home. However, as the data rate is shared between all users of that cell, data rates of up to few 100kbps are more common. With the advent of UMTS Release 5 the High Speed Downlink Packet Access (HSDPA) increases this bit rate to up to 10Mbps, shared between the users of a cell. While HSDPA is currently being deployed and it is started to be used, the access speed of wired links at homes is increased to few Mbps, which does not need to be shared with other users. Hence, comparable wireless access is again not possible in Hot Spots, where many users share the wireless bandwidth.

In parallel to the cellular networks Wireless Local Area Networks (WLANs) have been developed. The paradigm of these networks is completely different, as they aim at low cost high bit rate access. This comes at the cost of QoS, which cannot be guaranteed and efforts to differentiate at least between service classes have not yet found the way into the mass market. Instead, WLANs are deployed in many Hot Spots, usually consisting of one or few access points only. As in cellular systems, the users connect directly to the access point, but with less transmission power. Hence, the coverage is reduced, but also the spatial frequency reuse is increased. Handovers between access points are possible, but rarely used. The aim of these networks is to offer fast Internet access to users at that location; moving users that want to stay connected while traveling long distances have to use cellular systems. However, because of the comparable low operational cost and the high bit rates, WLANs build a strong alternative for Internet access at Hot Spots.

The channel bit rate of an access point to be shared between all users in a WLAN could be as high as 54Mbps. But, also at lightly loaded access points, the data rate of users is much lower, because of the probabilistic medium access protocol and the resulting overhead. The probabilistic channel access method poses also the problem of a variable channel access delay and further variabilities in the available bit rate to the user. The delay fluctuations could disturb real-time communication, because of the high delay jitter. The bit rate changes due to the channel access are usually not so problematic as long as the average bit rate is reasonably high. However, it depends on various other parameters than the channel access method. E.g., I will show in the next section that the bit rate available to a user of a WLAN can drop from reasonable values to a small fraction of that, only because other users have changed their locations. In many scenarios, even the high bit rates of WLAN today are not high enough or not efficiently enough exploited to offer users the Internet access quality in Hot Spots, which they are used from their wireline connection at home.

The aim of this thesis is to identify the main issues and performance influencing parameters of wireless Internet access at Hot Spots. Further, I will propose methods to increase the efficient use of todays WLAN systems and judge them by means of analytical and simulative performance evaluation.

The first task is to identify the main performance influencing parameters of WLAN. It is important to know, why users are confronted with such highly variable bit rates. By knowing the reasons, one could possibly find counter measures. To do so, I will first build an analytical model of WLAN, verify and evalute it in Chapter 2. The model includes a physical layer abstraction, that considers different modulation schemes and their error characteristics at bit level. The Medium Access Control (MAC) layer is modeled in great detail, as one could expect that access to the shared channel plays a significant role for the overall performance of users and system. This model is then used to analyze the QoS of different applications. First simple Constant Bit Rate (CBR) flows are considered from the access point to the users, a traffic characteristic that multi-media streaming applications could generate. Then the model is enhanced to Transmission Control Protocol (TCP) flows from the access point to the users. This represents the case where users download data from the Internet, still the most common application in the Internet today. Finally, bi-directional CBR traffic is modeled, representing e.g. voice communication. After verifying the model by means of simulations, I identify the QoS of users in a WLAN and the main parameters on which the QoS depends.

One of the main QoS influencing parameters is the location distribution of all other users that are connected to the same access point. I make use of that to control the QoS of users connected to a WLAN in Chapter 3. QoS control in WLAN is different from QoS control in fixed networks. Some requirements on a QoS control system are defined, which reflect the special properties of WLAN. Then two QoS control systems based on user locations are proposed and their superior performance is shown by means of simulations.

The QoS control systems allow to enhance the QoS of selected users and further to increase the efficiency of the access point. Still, the capacity of the network is not sufficient to satisfy all users. One way to increase the capacity is multi-hop communication. Thereby, stations act as source, destination and router in parallel, possibly forwarding flows for other stations. As the number of hops increases, the hop length decreases and more efficient communication takes place. I investigate how multi-hop communication can increase the network capacity in our access network scenario in Chapter 4.

Multi-hop communication has been researched for years and many different proposals have been made to increase the capacity further or make the communication more efficient. First, it is important to find out which of these mechanism should be used in the special case of multi-hop Internet access networks that is considered. Two basic wireless communication mechanisms, namely rate adaptation and power control are investigated to see their effect on the network capacity. These mechanisms are evaluated in Section 4.2 regarding a theoretical maximum performance gain.

Next, a suitable routing algorithm needs to be found that reaches the performance limits as good as possible. I develop an analytical evaluation methodology to compare selected routing algorithms from related work. Further two routing metrics are proposed that use the capacity of the access network very efficient and these are evaluated as well. The evaluation methodology is verified as well as the performance figures by means of simulations and it is shown that the new routing metrics outperform those of related work and nearly reach a theoretical upper performance limit.

Finally, I conclude this thesis in Chapter 5, where I summarize again the problems and issues of WLAN usage in areas with high user density. I highlight my findings and give suggestions on how to enable efficient Internet access in Hot Spots.

Chapter 2

An Analytical Model of Wireless LAN

Wireless LAN (WLAN) in general is able to offer users high bit rates and reasonable low delays, compared to other wireless communication technologies. However, the offered Quality of Service (QoS) is highly variable and usually not controlled. A user might get fast access at one point in time and the same user is offered a very low bit rate shortly afterwards. There is a significant dependence between the QoS of all users and a change in the wireless connection to one user could have performance implications to all others. All these factors make it hard to estimate the QoS of communication over WLAN. However, many applications could benefit from a more stable QoS or at least a more predictable one. E.g. users that are connected to cellular networks and get into coverage of a WLAN access point would need a reliable QoS estimation to make efficient handover decisions. In this section, the problem of identifying the main performance influencing parameters, i.e. the parameters that this QoS variability depends upon, is tackled.

It is a well known fact, that research, development and performance evaluation on network and transport layer requires the right models of the lower layers of the communication protocol stack (see for example [46]). The right models could thereby range from very detailed bit level modeling to rather abstract models at packet or flow level, depending on the issues that are to be investigated. E.g., a proof of concept for a signaling protocol might require only a high level packet error and delay model, whereas a performance evaluation of media transmission over a wireless channel requires detailed bit error modeling.

The first step now, is to build an analytical model of WLAN to get a good understanding of the system and be able to investigate the effect of parameter changes on the user or system performance. As said above, this model needs to have the right level of abstraction. The bit level needs to be considered, because different modulation schemes can be used. These in turn determine large parts of the transmission speed and error probability. The packet level must be considered in detail, because the MAC layer controls channel access by a probabilistic access scheme that is expected to have a significant performance implication. Further, packet retransmissions could be scheduled for corrupted or lost packets. Finally, the flow level has probably significant performance implications, because all flows will usually share the common bottleneck access point in order to reach the Internet or reach the users from the Internet. Hence, an analytical model of WLAN is required on bit, packet and flow level; it will be developed and presented in this chapter.

First, an overview of WLAN and especially the IEEE 802.11b system is given in the next section. WLAN performance has been studied for quite some time now, resulting in many different WLAN models at different levels of abstraction. The main related work in this area is presented in Section 2.2 and thereby it is shown how to differentiate the model to be newly developed from the prior art. The analytical model of WLAN at bit level is presented in Section 2.3. There, the assumed propagation and error model is presented and a transaction error probability depending on parameters such as the used modulation scheme and distance between sender and receiver is derived. Considering the channel access mechanisms of WLAN and finally deriving the MAC layer throughput of a station depending on its distance to the access point is done in Section 2.4, where the packet level model is presented. Assuming Constant Bit Rate (CBR) traffic from the access points to the users, the service time distribution of the access point is calculated and from that the throughput that each user will get is estimated in Section 2.5. The results are verified by simulations.

In Section 2.6 the scenario is enhanced to consider TCP traffic flows via the access point to the users. Thereby, the access point is modeled as an M/G/1/B queuing system. The system is analyzed by mean of a Markov chain and from that the loss rate and delay is calculated. This enables to determine the TCP throughput that users achieve. Again, the performance is evaluated in detail and the findings are verified by means of simulations. While the TCP throughput calculation already required to consider collisions on the wireless channel, because TCP data packets might be sent downlink at the same time as a TCP acknowledgement uplink, the collision model is further enhanced in Section 2.7. There, bi-directional voice traffic is modeled using the queuing model and Markov chain analysis from the previous sections.

2.1 Introduction to Wireless LAN

WLAN in general is a wireless data communication possibility in a locally limited area. Most WLAN technology used today is based on the IEEE 802.11 standard. All models and investigations in this dissertation are also based on this standard. Hence, I will describe the main properties and issues of WLAN according to IEEE 802.11 in this section.

2.1.1 A Short History

The Institute of Electrical and Electronics Engineers (IEEE) identified the need to have a wireless alternative to Ethernet LAN and started standardising WLAN in its activity group IEEE 802.11 in the early nineties. In 1997 the first standard, IEEE 802.11, was released. It specified two different raw data rates, namely 1 Mbps and 2 Mbps, to be transmitted via

Infra-Red (IR) signals or in the Industrial Scientific Medical (ISM) frequency band at 2.4 GHz. Further, the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) scheme was defined to access the channel.

Event though, the manufacturers adopted this first standard and designed a variety of IEEE 802.11 compliant devices, interoperability was not necessarily given. This was due to the fact that the standard allowed different implementation possibilities which could lead to problems in interoperability. In 1999 the first amendment to the initial standard was released as IEEE 802.11b. Here, the physical layer was enhanced by two more possible data rates, 5.5 Mbps and 11 Mbps. As the interoperability problems were still present, but the manufactures saw the significance of the WLAN market, the Wireless Fidelity (WIFI) alliance was founded in 1999. The WIFI alliance is an organization of about 200 companies that certificates WLAN products according to interoperability. The work of the WIFI alliance increased the acceptance of WLAN at end users and was the driving factor to the success of WLAN that we see today.

In parallel to the IEEE 802.11b standardization, the IEEE 802.11a amendment, released also in 1999, described also a new physical layer in a frequency band at 5 GHz; data rates of up to 54 Mbps are possible. In 2003 the IEEE 802.11g amendment was finalized, where it is possible to have these high data rates in the same 2.4 GHz band as the IEEE 802.11b standard.

2.1.2 IEEE 802.11 Architecture

The IEEE 802.11 architecture is based on the main building blocks Station, Basic Service Set (BSS), Extended Service Set (ESS) and Distribution System (DS). An overview on how these could be interconnected can be seen in Fig. 2.1.

The smallest IEEE 802.11 entity is a station. A station is a component that connects to the wireless medium. It consists of a MAC and a Physical Layer (PHY). A station can be portable, mobile, or embedded and offers fundamental services such as authentication, de-authentication, data delivery and privacy.

A BSS is a set of stations communicating with one another. A BSS does not refer to a sharply bounded area because of propagation uncertainties. If all stations within a BSS are mobile and there is no connection to another (wired) network, the BSS will be referred to as an ad-hoc network. An ad-hoc network is typically a short lived network with a relatively small number of stations created for a temporary purpose, e.g. exchange files during a group meeting. All stations communicate directly with one another. A BSS which includes an Access Point (AP) is called infrastructure BSS. The AP is a dedicated station, which provides additional functionality. Any communication among stations is routed via the access point. The AP is also defined as a gateway to the wired network or to a Distribution System (DS), which interconnects multiple BSSs to enlarge the WLAN networks.

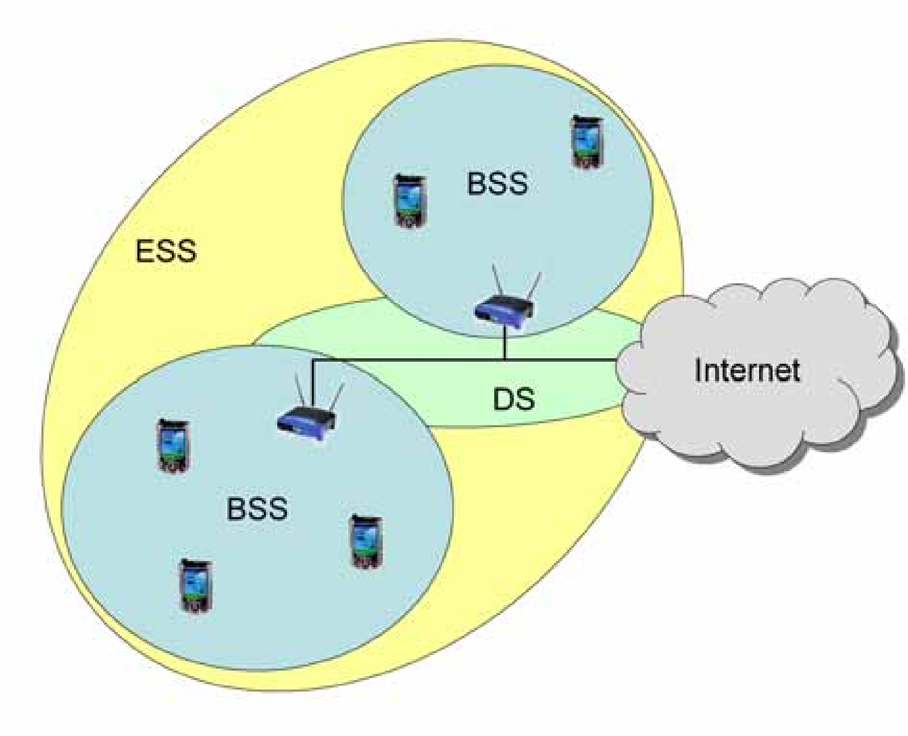


Figure 2.1: IEEE 802.11 architecture overview.

The DS and BSS allow forming a wireless network of arbitrary size and complexity. This type of network is referred to as an Extended Service Set (ESS). All stations within an ESS may communicate with each other and mobile stations may move between BSS. Thereby, BSS can be physically disjoint or co-located, or they can be partially or completely overlapping.

2.1.3 Protocol Architecture

The IEEE 802.11 protocol architecture is depicted in Fig. 2.2. Five different physical layers are available up to the time of writing, which are subdivided further into a Physical Medium Dependent (PMD) and a Physical Layer Convergence Protocol (PLCP) layer. The PMD provides the actual interface to send and receive data between two or more nodes. The PLCP allows the IEEE 802.11 MAC to work with a minimum dependence on the PMD sublayer. It opens the opportunity to use the same MAC protocol on top of several physical layers by offering the same interface.

The MAC layer is situated on top of the physical layer and subdivides into a Distributed Coordination Function (DCF) and a Point Coordination Function (PCF). The DCF incorporates all basic MAC functionalities and provides for the fundamental contention service, which is similar to the asynchronous unreliable service offered by IEEE 802.3 networks.

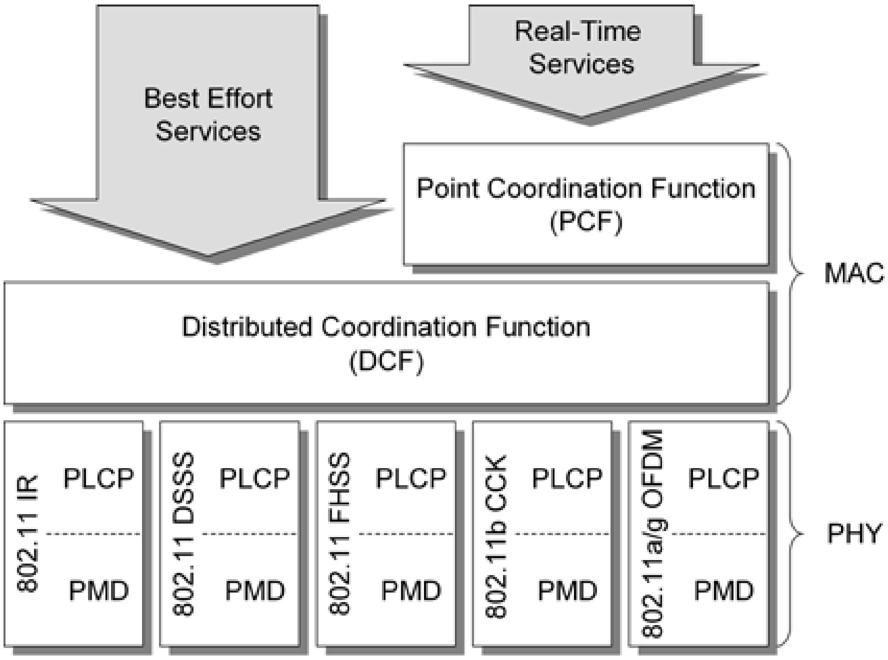


Figure 2.2: IEEE 802.11x protocol overview.

The optional PCF resides on top of the DCF and offers time bounded and contention free service.

The IEEE 802.11x standards provide five physical layers based on Infrared (IR), Frequency Hopping Spread Spectrum (FHSS), Direct Sequence Spread Spectrum (DSSS), DSSS with the use of Complementary Code Keying (CCK) modulation, and Orthogonal Frequency Division Multiplexing (OFDM), respectively.

In IEEE 802.11b, there are two possible PLCP frame structures, depicted in Fig.2.3. The long preamble type frame is backwards compatible with the 1 and 2 Mbps version. It consists of a SYNC field for receiver synchronization purposes, a Start Frame Delimiter (SFD) field marking the start of the frame, a Signal field indicating which data rate must be used to receive the Protocol Service Data Unit (PSDU), a Service field, which primarily indicates which type of modulation must be used to demodulate the PSDU, a Length field indicating the frame end to the receiver, and a Cyclic Redundancy Check (CRC) field used to check the correctness of the received header. The next field called PSDU contains the payload data. The preamble and the header are sent at 1 Mbps using DBPSK modulation. The PSDU is sent at the speed indicated by the Signal field.

Since the long frame preamble imposes a considerable overhead at higher data rates, a frame with a short preamble, which is not backwards compatible to the 1 and 2 Mbps frame version with the long preamble, was introduced. Therefore all stations in a BSS will have to support a short preamble frame type if it is used by one station. A frame with a short preamble primarily differs from its counterpart in the SYNC field, which is

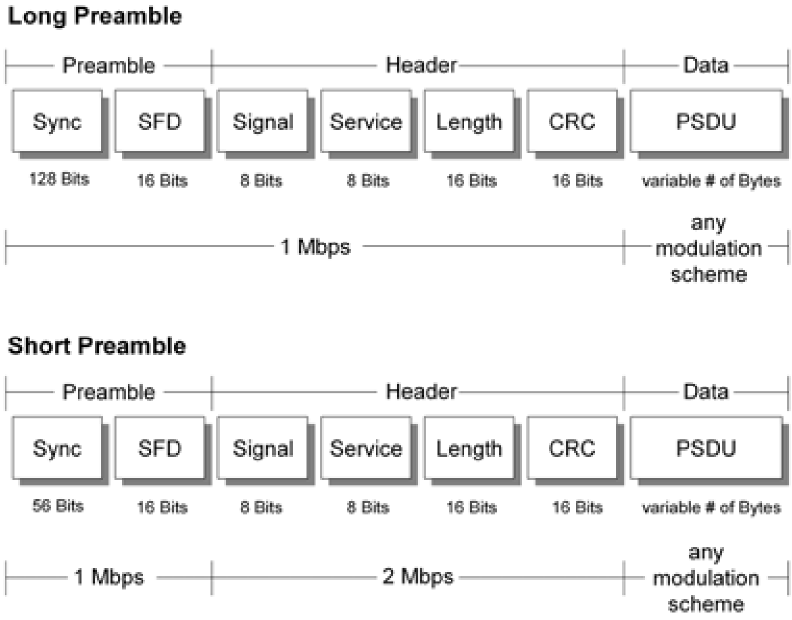


Figure 2.3: PLCP frame structures in for the long and short preamble mode.

considerably shorter. The preamble part of the frame is sent at 1 Mbps while the header part of the frame is sent at 2 Mbps. The PSDU is sent at 2, 5.5 or 11 Mbps.

2.1.4 IEEE 802.11 Medium Access Control

The primary task of the IEEE 802.11 MAC protocol is to provide an efficient, fair, secure and reliable data transfer service. Several problems arise because of the communication medium being a radio channel. The MAC protocol has to take the noisy and unreliable channel into account. This includes that the channel conditions may change rapidly.

The basic MAC protocol is similar to the 802.3 MAC. It belongs to the Listen Before Talk (LBT) protocol class, but uses additional features, e.g. immediate acknowledgements or channel reservation. The IEEE 802.11 MAC works in a distributed manner, although it also supports centralized control.

2.1.4.1 Frame Formats

An IEEE 802.11 frame as shown in Fig 2.4 basically consists of a frame header, a frame body and a Frame Check Sequence (FCS). Not all fields in the MAC header are used for all MAC packets. The actual frame composition using the fields described below and the exact meaning of these depend on the type and subtype of the frame.

Figure 2.4: General MAC frame format.

- The **Frame Control** field comprises information needed by the receiving MAC to interpret the subsequent fields of the MAC header. It contains information such as the type (data, control, management) and subtype (specific type of frame, e.g. Beacon) of the frame, whether the frame is a fragment and further fragments are outstanding, whether the frame is retried and encrypted or what the power mode of a station is or will be after a successful frame exchange.

- The **Duration/ID** field indicates how long an ongoing frame exchange lasts, where a frame exchange consists of a MAC data packet plus MAC acknowledgement in unicast mode or only the MAC data packet in broadcast mode. The value is used to update the Network Allocation Vector (NAV), which provides a virtual carrier sense capability. The virtual carrier sense mechanism is described in detail in the next section.

- The **Address** fields describe the originator (source address) and/or the receiver (destination address) of a particular frame. Two additional address fields can be incorporated, depending on the type of the frame. These two fields may be used for the BSS identifier (only probe request frames), for the transmitter address or for the receiver address (which can be different from source and destination address because the frame is sent via an AP into another BSS).

- The **Sequence Control** field identifies the frame by a particular number and enables the receiver to detect duplicate frames. Four Bits of this field are used to number fragments if the packet is fragmented.

- The **Frame Body** contains the information specific to each data or management frame. It may be as long as 2312 Bytes.

- The **FCS** field contains the result of the CCITT CRC-32 polynomial applied to the MAC header and the frame body. It is used to check the correctness of the entire MAC frame.

2.1.4.2 Distributed Coordination Function

The Distributed Coordination Function (DCF) is the basic operation mode. It is generally described as Carrier Sense Multiple Access / Collision Avoidance (CSMA/CA). This means it is a Multiple Access (MA) scheme with a Carrier Sense (CS) and a Collision Avoidance (CA) component.

If the MAC receives a request to transmit a frame, it first checks the availability of the radio channel. If the channel is sensed idle for a Distributed Inter-Frame Space (DIFS) time interval, the frame transfer starts immediately. There may be an additional random delay introduced by the CA mechanism if there was an ongoing transmission while the transmit request was received. The CA mechanism is explained later in more detail. If no other station starts a frame transmission before the CA phase is over, the MAC will start the transmission of the frame. Afterwards the MAC awaits an Acknowledgement (ACK) from the receiver of the frame within a Short Inter-Frame Space (SIFS) time interval. Upon reception of the ACK the whole procedure starts over for a new request to transmit a frame, if any. If the station has not received the ACK after time Extended Inter-Frame Space (EIFS), the frame is retransmitted until it is successfully received or the maximum retry counter expires. Retransmissions influence the CA mechanism as explained later. The basic access mechanism is depicted in Fig. 2.5.

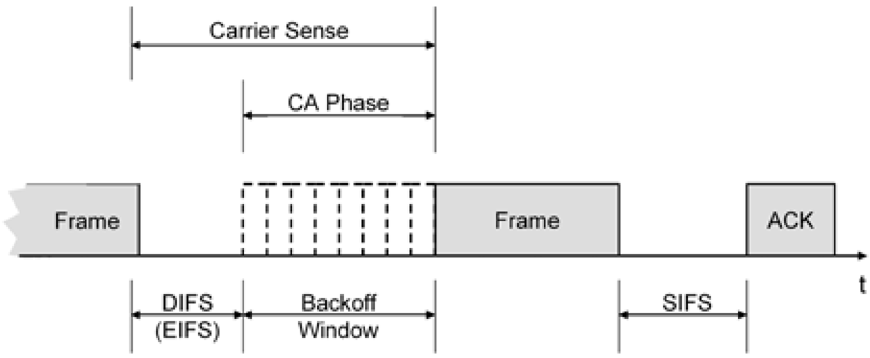


Figure 2.5: IEEE 802.11 basic access mechanism.

As said above, a frame is usually transmitted without any delay after sensing the channel idle for DIFS time. If the channel was not silent during the DIFS listen interval, contention for the radio channel is assumed, which can lead to a higher collision probability. The MAC counteracts contention with the CA mechanism, which basically introduces an additional delay before the actual frame transmission to spread the channel access of different stations in time. If the radio channel remains idle within in this random interval, the frame transmission is started. The interval is called backoff window and is calculated as:

$$T_{bo} = b \cdot T_{slot}, \tag{2.1}$$

where b denotes the backoff value drawn from the contention window. b is a discrete random variable uniformly distributed between 0 and CW; T_{slot} is the duration of a time slot, which is a physical layer dependent parameter. The contention window, CW, is initially set to a value CW_{min}, which is 31 in the initial IEEE 802.11 standard. Each time a transmission fails, the MAC layer assumes that this had happened due to a collision and thus further reduces the collision probability by increasing the backoff window interval. Thereby, the contention window is increased as follows:

$$CW_{new} = 2 \cdot (CW_{old} + 1) - 1. \tag{2.2}$$

However, a maximum contention window, CW_{max} is defined, which is 1023 in the initial IEEE 802.11 standard. In case of a successful transmission the contention window is reduced to $CW = CW_{min}$.

A station, which receives a signal from another station while listening to the radio channel during the backoff interval, defers until the channel was idle for a DIFS interval again. Instead of computing a new backoff value it uses the old value minus the time already spent during the backoff interval of the previous channel access attempt. Therefore a station that lost the competition for channel access has a higher priority for the next channel access attempt.

Besides physical layer channel sensing, additionally virtual channel sensing is used by means of the Network Allocation Vector (NAV). It indicates usage of the channel, independent of the outcome of the physical layer channel sensing. The NAV is set according to the duration field of a received frame. The duration field contains the time in microseconds that is needed for a successful completion of the frame exchange sequence. Even if a station cannot detect the entire frame exchange, e.g. due to range limitations, it will not disturb an ongoing transmission by accessing the channel itself, since the set NAV indicates an unavailable channel.

Usually not all stations within a BSS are aware of an ongoing transmission because of range limitations. As a result collisions can occur at the receiver. This problem is commonly referred to as the Hidden Terminal problem. To combat this, a four-way handshake transmission mode, namely Ready to Send (RTS)/ Clear to Send (CTS), can be used optionally. By transmitting an RTS message prior to the data all stations in the vicinity of the sender are informed about the start of a transmission. Although we could assume collisions only to occur at the receiver, the sender has to inform its vicinity as well, since it has to receive the ACK frame to complete a frame transfer successfully. Accordingly the intended receiver of a frame sends a CTS to inform its vicinity. The surrounding stations update their NAV and will not transmit as long as the NAV is set. The RTS frame may collide with other frames as well since the basic channel access rules apply. But the collision time, i.e., the total time between the start of the transmission of at least two frames and the end of the transmission of the frames, is significantly reduced as a secondary effect. Only short RTS instead of the longer MAC Protocol Data Units (MPDU) collide.

Fragmentation is a MAC feature to combat packet errors. Generally, long packets are more likely to be hit by an error than shorter packets. Therefore, a frame can be split into several fragments if the frame size exceeds a certain adjustable threshold value. Fragmentation may be combined with the RTS/CTS mechanism. Fragmentation is transparent to the upper protocol layers. In order to transmit a fragmented frame efficiently, the frame is sent as a fragment burst. Normal access rules apply for the first fragment. The remaining

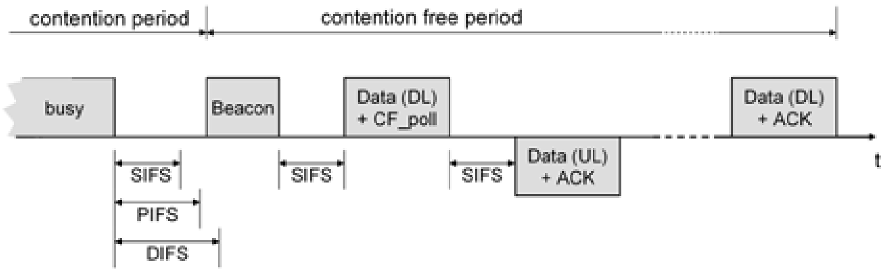


Figure 2.6: IEEE 802.11 point coordination function.

fragments are sent immediately after a SIFS, if the previous frame was acknowledged. If one fragment remains unacknowledged, the transmission of the frame is resumed with unacknowledged fragments following the basic access rules. In other words, instead of transmitting the entire frame again, which can incur a considerable overhead, only the missing part of the frame is sent. The packet size plays a major role regarding the performance. The largest payload of a MAC packet is 2312 Bytes and the smallest is 0 Byte. Assuming an ideal radio channel, large packets would result in a high throughput performance because of the small overhead payload ratio. However, the collision window and the channel access delay may be increased, in turn reducing the throughput. Small packets a priori lead to a lower throughput because of the higher overhead-payload ratio, but the packet error probability is reduced. Therefore the likelihood of a successful transmission over an impaired radio channel can be increased to a certain extent if shorter packets are used.

2.1.4.3 Point Coordination Function

The IEEE 802.11 MAC may also incorporate an optional access method called a Point Coordination Function (PCF), which is only usable on infrastructure network configurations. This access method uses a point coordinator (PC), which operates at the access point of the BSS, to determine which station currently has the right to transmit. The operation is essentially that of polling, with the PC performing the role of the polling master. The functionality of the PCF is depicted in Fig. 2.6.

The PCF uses a virtual carrier-sense mechanism aided by an access priority mechanism. As can be seen in Fig. 2.6, the time is divided into a contention period and a contention free period. During the contention period the basic access scheme is used, mainly for stations using the best effort service. The contention free period is governed by the PCF mechanism.

To start a contention free period, the PC sends a Beacon management frame that sets the NAV in all stations to the length of the contention period. This beacon is sent after sensing the channel idle for a PCF Inter Frame Space (PIFS). The PIFS is shorter than DIFS to have priority over basic access data frames, but longer than SIFS in order to allow ACKs to be sent before the contention free period starts.

After reserving the channel with the beacon frame, the PC polls the stations for uplink data and sends the downlink data scheduled for this contention free period. Therefore, the downlink data for the first station is sent and/or the first station is polled. The frame is sent a SIFS period after the beacon frame. If data was sent, the PC waits for an ACK. If a station was polled, the PC waits for data. In both cases the PC has to wait for a SIFS period only, after which it either receives data or ACK or polls the next station. The ACK of the previous data packet sent is always piggy-backed onto the data frame to the next station, as is the poll to the same station the data is sent to.

2.2 Related Work

The performance of WLAN has been investigated for years now. The result of this effort is a variety of analytical WLAN models, simulation studies and test bed measurements. For the aim of finding the main performance influencing parameters of WLAN, a bit level, packet level and flow level model is required to come to reliable conclusions. In this section the related work is described and it is shown why the existing models were not sufficient for this work.

First, related work on modeling WLANs assuming saturation conditions is described. Saturation conditions mean that the stations always have a packet to send. After that, models are described that do not make this assumption, but rather consider some kind of application or packet arrival process at the stations. These models are usually described by queuing systems.

2.2.1 Packet Level Models under Saturation Conditions

Bianchi was the first to describe an analytical model of the DCF at packet level in [13], [14] and [15]. The average throughput is evaluated by means of modeling the channel access behavior with a two-dimensional Markov chain. A number of sources are assumed, which are saturated, i.e. always have packets to send. Transmission errors are not considered and all packets are sent with the same modulation scheme. As the aim of this thesis is to investigate the effect of all parameters on the user performance, especially the users' locations, a more accurate model of the packet transmission itself is required. Further, looking at applications that are in use at Hot Spots today, e.g. surfing the Internet, multimedia streaming, voice communication, saturated sources cannot be assumed. Still, this quite simple model has proven to be accurate in predicting the performance of WLAN under the above mentioned assumptions and is the reference to nearly all following work done on WLAN modeling.

In [23] Chatzimisios et al. extend the above mentioned Bianchi model to calculate the average transmission delay of packets, including channel access delay, but excluding queuing delay. In [24] the authors further enhance the model to consider a finite number of retransmissions and calculate the transmission delays in that case. Further a packet dropping probability is calculated, where packet drops occur only if the maximum number of transmissions is exceeded. The same model is used in [26] and [27] to investigate the effect of the backoff parameters on the delay and throughput performance and further in [114] to investigate the effect of allowing stations to send packet bursts instead of single packets. Another enhancement of that model is described in [25], where transmission errors are considered by means of a fixed bit error rate. All of these models have in common that saturated sources are assumed and rate adaptation is not considered.

Ziouva et al. extend the Bianchi model to consider a state in which a station could send without invoking the backoff procedure in [130]. However, according to the IEEE 802.11 standard this is only allowed, if a station wants to send a packet, has not sent a packet immediately before and senses the channel idle for time DIFS. As Bianchi considers saturated sources only, this state would not be reached. In [130] the authors falsely assume that this state is also entered when a station has just sent a packet and immediately after that senses the channel idle for time DIFS. Still, the analysis of considering the additional state without backoff is useful for investigations of non-saturated sources. However, the proposed Markov chain is solved under the assumption of saturated sources.

In [19] Cali et al. use a similar model to calculate the saturation throughput for different numbers of competing stations. Further, a theoretical upper bound of the throughput is calculated assuming a perfect channel access scheme. It is shown that by tuning parameters of the backoff scheme this maximum throughput can nearly be reached. However, the optimum parameters depend on the environment and scenario and thus would be need to be changed dynamically. The authors prove their results by simulations. Xiao et al. calculate a theoretical throughput limit of 802.11 networks in [121]. They also consider different modulation schemes and come to the conclusion that regardless of the used transmission bit rate a theoretical throughput limit exists, which depends on the channel access scheme and transmitted overhead. A similar investigation with the same results was conducted by Jun et al. in [65]. Even though different modulation schemes are considered, I cannot make use of these models for my work. In these models, all stations use the same modulation schemes and work under saturation conditions. Bit errors are not modeled and neither are queuing losses, which means that packets can get lost only because of collisions.

In [47] and [49] the authors extend the Markov chain of the Bianchi model to consider transmission errors. Thereby, a fixed bit error rate is assumed. Further, the authors consider a finite number of retransmission attempts, which also has an influence on the WLAN performance in terms of throughput and delay. In [50] the authors extend the two dimensional Markov chain analysis of Bianchi to a three dimensional Markov chain. The third dimension represents the retransmission count of a packet and hence a finite number of retransmissions could be modeled here as well.

In contrast to his original work with the Markov chain analysis, Bianchi proposed a different methodology to model DCF in [16]. Here, conditional probabilities are used to calculate throughput and delay. This procedure can be extended to account for other backoff operations, but still considers saturated sources and ideal channel conditions. A similar methodology was used by Tay et al. in [107], were mainly the dependence between contention window size, number of sources and collision probability was investigated. In [73] Kumar et al. use the Bianchi assumptions and show that the packet transmission rate of a station depends on its collision probability and the collision probability depends on the packet transmission rate, resulting in a fixed point problem. Solving that problem the authors come to similar performance results as Bianchi.

Heusse et.al. were the first to describe the performance anomaly of WLAN by means of an analytical model in [51]. The performance anomaly is the fact that users which transmit their data using a lower bit rate modulation scheme degrade the throughput of all users. The model used for this work is however quite simple: Saturated sources are considered and an average throughput is calculated. Further, the authors consider ideal channel conditions and define each user to transmit with a certain modulation scheme.

In [78] Li et al. enhance the Bianchi model to consider different traffic service classes. However, still only saturated sources and ideal channel conditions are considered. A similar investigation is performed by Xiao et al. in [122]. In [74] Kuo et al. calculate the average delay and throughput for different service classes using conditional probabilities. Similarly, the packet transmission delay is calculated for the IEEE 802.11e enhancements under saturation and ideal channel conditions by Banchs et al. in [9]. In [10] the authors extend this work to model admission control decisions to guarantee delay limits for priority packets.

Wu et al. propose a DCF enhancement to increase the performance of reliable transport protocols over WLAN in [120]. Collisions between transport layer acknowledgements and the next data packet are avoided by allowing sending the transport layer acknowledgement directly after the MAC layer acknowledgement. The authors show an improved performance by means of an analytical model considering again saturated sources and ideal channel conditions.

In [113] Vishnevsky et al. consider the seizing effect, which describes the fact that a station which has just sent a packet has a slightly increased probability of being able to send the next packet as well and thus seizing the channel. The significance of the seizing effect is shown to be depending on the backoff parameters. In the model that is proposed in this work, the seizing effect is implicitly considered, as the DCF is modeled in detail. However, because competing saturated sources are rarely considered, this has no effect on the performance.

Kim et al. show in [70] that throughput of stations is depending mainly on the collision probability, which in turn mainly depends on the number of stations that want to send. The authors propose a scheduling strategy, where nodes first calculate the network utilization according to the model and the observed collision probability. The calculated network utilization is used to determine a waiting time, by which the transmission is delayed. Using this mechanism, the collision probability can be decreased and the throughput increased.

In [75] Kwon et al. investigate WLAN performance by means of a simple model concentrating on the backoff mechanism only. A new collision resolution algorithm is then proposed, which is shown to perform well in simulations.

2.2.2 Modeling WLAN by Queuing Systems

A WLAN station can well be modeled by a queuing system as it logically consists of a queue and a server. The parameters that describe the queuing system are the arrival rate and distribution, the service time and distribution and the buffer space. There are many investigations published on modeling WLAN by means of queuing systems. However, the assumptions and abstractions are quite different, e.g. finite buffer space is only rarely considered, and the service time calculation usually does not include physical layer issues, such as different modulation schemes and bit errors. We will describe here the main related work in this area.

The Bianchi model and all other models described above consider saturated sources only. In [17] Bruno et al. present results of a simulation study of TCP downloads in a WLAN Hot Spot. The main result is that the saturation condition is not fulfilled and the performance changes significantly from results assuming saturation. As TCP downloads can still be considered as the main application in the Internet today, the need to consider the flow level in addition to the packet level is obvious. The authors build an analytical model of TCP over WLAN in [18]. By means of conditional probabilities they calculate the channel utilization, assuming that the access point has always a packet to send. The stations reply with TCP acknowledgements. Infinite buffers are assumed in the access point as well as the stations. Hence, queuing losses cannot occur. Further, ideal channel conditions and only one modulation scheme are considered, meaning that transmission errors or different transmission delays do also not occur.

Tickoo et al. model WLAN stations as G/G/1 queuing models in [108]. The service time distribution is calculated from the DCF packet dynamics. Ideal channel conditions and no rate adaptation are considered. As the authors consider infinite buffers at the stations, queuing losses are not considered. Evaluating the performance of these stations under variable traffic load, the authors note a significantly reduced collision probability if the network is not saturated. The authors extend their model in [109] to consider different service classes.

In [82] Litjens et al. model besides the packet level also the flow level. The authors first calculate the throughput for saturated sources from the Bianchi model. Assuming that all users get the same share of the throughput, this serves as input to a processor sharing model at flow level. Assuming a flow size distribution, they calculate the average flow duration. The authors could reproduce the typical processor sharing performance, i.e. the flow durations are small in low load situations, but increase rapidly from a certain load value onwards. In [30] the authors extend their model to support different service classes according to the IEEE 802.11e standard.

In [126] Yin et al. consider non saturated sources and further transmission errors. The transmission errors are assumed to be the same for all stations. To model non-saturated sources, the authors assume each station to have a fixed probability to be backlogged, i.e. having to send the next packet immediately after the previous one. While this lets the authors investigate the network and channel state in more detail, it does not give an insight to the user performance. Queuing losses and individual station dependent loss rates and modulation schemes are important factors that would need to be considered.

TCP over WLAN is modeled by Miorandi et al. in [88]. It is assumed that all stations download data via the access point and further that the access point has always a packet to send. Hence, the TCP dynamics are not modeled. Assuming ideal channel conditions the time to send a TCP data packet and receive a TCP acknowledgement is calculated and from that the access point throughput is determined. The effect of delayed acknowledgements is considered, by modifying the model such that an acknowledgement is sent only after the reception of two data packets. They make use of fact that the terminals in WLAN receive a fair share of data from the access point in [89] as it was done in the processor sharing model in [82]. In [89] they consider HTTP downloads to multiple users via an 802.11 access point. Assuming that each user gets a fair share, the authors calculate the mean delay of short lived TCP connections. They focus on investigating the effect of different TCP enhancements. Ideal channel conditions and no rate adaptation are assumed.

In [39] Foh et al. calculate the throughput of WLAN under ideal channel conditions and again, it is assumed that each station in the network gets the same share of that throughput. The authors then construct a state dependent M/M/1/k queue, where the number of customers in the queue represent the number of users in the system. Hence, the queue is here not used to model the station itself, but rather used to model the number of active users in the WLAN. With the users in the system the service time is changed according to the available throughput share to the users.

Lei et al. extend the Bianchi model to power managed WLANs in [77]. Thereby, packets that are destined to a station that is currently in sleep modus are buffered. The authors model that by means of an M/G/1 queuing system with bulk service. Even though buffering of packets is explicitly considered and the main issue of the investigation, only infinite buffers are considered. Further, transmission errors or different modulation schemes are neglected and hence packets can only be lost because of collisions.

An M/G/1/k queuing model is developed by Özdemir et al. in [93]. The service time is thereby described by a Markov Modulated General Independent (MMGI) process. The number of currently busy nodes determines thereby the service time. The authors show by means of simulations that their model approximates WLAN behavior quite well. However, ideal channel conditions and only one modulation scheme are considered. Hence, packet failures could only happen because of collisions. Integrating a transmission error model is not trivial, because an important assumption for their model is that all stations suffer from the same packet error probability. The authors consider collisions as the only source of packet loss and thus this assumption is valid for their cases.

In [12] Bellalta et al. model the stations as M/M/1/k queues. Hence, finite buffers in stations are considered and queuing losses could happen. The average throughput is calculated from the Bianchi model under saturation conditions. Assuming that all stations get the same share of throughput, the average service time is calculated and the distribution is assumed to be exponential. The packets are assumed to arrive according to a Poisson process. An ideal channel is assumed. Considering the finite buffers lets the authors calculate a loss probability, but the service time distribution cannot be assumed to be exponential in rate adaptive WLANs as will be shown below.

In [21] Cantieni et al. model the WLAN stations by M/G/1 queuing systems. Markovian arrivals are considered and hence also non-saturated sources. The service time is calculated considering multiple modulation schemes, where the authors however simply assign certain stations to use certain modulation schemes. Packet losses could still only happen because of collisions, because neither transmission errors due to corruption, nor queuing losses are considered. The latter comes because the authors derive only a mean service time and not the service time distribution. Having only the mean service time does not allow to enhance the model to finite buffers and calculate a dropping probability.

Summarizing it can be said that a variety of analytical WLAN models exists, but these models all lack some level of detail. Most of the models assume saturated sources. As it was shown in [17], the collision probability is significantly different in scenarios where the sources are saturated, in contrast to situations in which application behavior is considered. Especially in the case of TCP transmissions, the collision probabilities are only a small fraction of what we see from saturated sources. As TCP is the main transport protocol used in the Internet today, I think that saturated sources cannot be assumed for my work.

There are models from related work that do consider non-saturated sources. In this case, calculating performance parameters usually involves solving a queuing system. Most of the models approximate the arrival process by a Poisson process, which is a reasonable assumption and simulation results from related work verify that. The service time is in some related work assumed to be exponentially distributed as well. It will be shown later that this cannot be assumed. Some authors calculate the average service time only. However, in order to calculate a loss probability, assuming finite buffer space, the service time distribution is required. An important thing to consider for the service time distribution is rate adaptation, transmission errors and resulting retransmissions, because these factors influence the service time significantly. This is however done in very few publications and not detailed enough. Individual users should be modeled using modulation schemes independently from each other and thus facing individual error probabilities as well.

Summarizing it can be said that even though a variety of analytical models exists, none of these can be used to determine the performance influencing parameters as it is aimed to do. Therefore a new model is required that considers non-saturated sources, i.e. a detailed flow level model. Further the packet dynamics have to be considered to calculate the service time distribution of a resulting queuing system. The service time calculation must include physical layer details, such as transmission errors and different modulation scheme characteristics.

2.3 Bit Level Model

In this section an analytical WLAN model is described that fulfills the requirements to model the relevant details at flow, packet and bit level. Thereby, first the bit level model is developed and presented. The bit level model describes the packet transmissions and receptions. The distance between sender and receiver, the modulation scheme to be used and the data packet size serve as the input parameters. Taking certain assumptions as described in this section, the transaction error probability can be calculated from that. I define the transaction error probability as the probability that the MAC data packet or the MAC acknowledgement corresponding to that packet contain at least one bit error.

2.3.1 Propagation and Path Loss

The signal strength decays on the path from sender to receiver. This path loss is depending on several parameters, e.g. distance between sender and receiver, environment, antennas, etc. While it is not possible to describe the exact signal propagation for all environments, different models exist, that can be applied to typical scenarios, such as signal propagation in micro or macro cells, in urban street canyons, etc. The COST-231 Walfish-Ikegami model [33] describes the path loss of a radio wave in micro cells. This model is adopted here, as it considers an environment that is likely to be seen in WLAN Hot Spot deployments. The path loss is given as

$$P_L(d) = 42.6 + 26 \cdot \log d + 20 \cdot \log f_c, \quad d \geq 20m, \tag{2.3}$$

where f_c is the carrier frequency. For distances $d < 20$ m free space attenuation is assumed, where the signal strength decays exponentially with a path loss exponent of 2. The mean received power at distance d is then given in dB as

$$\overline{P_r}(d) = P_t - P_L(d), \tag{2.4}$$

where P_t is the transmitted power, given in dB.

Obstacles between sender and transmitter additionally reduce the received signal power. This can be well modeled with a log-normally distributed random variable ([44], [101] and [115]) leading to the received signal power of

$$P_r(d) = P_t - P_L(d) + X_\sigma, \tag{2.5}$$

where X_σ is a zero-mean Gaussian distributed random variable (in dB) with standard deviation σ (also in dB).

2.3.2 Bit Error Rate

From the received signal power, the signal to noise ratio is derived. Thereby, thermic noise at the receiver is assumed and interference is considered only from other WLAN terminals. IEEE 802.11 allows only one station to send at the same time on the same frequency and thus the noise at a station is assumed to be additive with constant spectral density N_0. The received noise power can then be calculated as $N = N_0 \cdot B$, where B denotes the bandwidth of the received signal. With the time required to send one chip, T_c, and the chip-rate, R_c, the energy per chip is given as

$$E_c(d) = P_r(d) \cdot T_c = \frac{P_r(d)}{R_c}. \tag{2.6}$$

The chip-rate is defined in IEEE 802.11b as $R_c = 11 \cdot 10^6$ chips/s. The transmission rate, R_t, however, varies for different modulation techniques between 1 Mbps, 2 Mbps, 5.5 Mbps and 11 Mbps. The processing gain can be calculated depending on the used modulation scheme, m, as

$$G(m) = 10 \cdot \log \left(\frac{n_c \text{ chips}}{n_b \text{ bit}} \right), \tag{2.7}$$

where n_c denotes the number of chips per n_b bits using modulation scheme m.

By assuming only white Gaussian noise at the receiver, bit errors occur independent from each other. While this is not very realistic in real implementations, it serves here as a kind of worst case estimation, because a packet is already discarded if only one bit is corrupted.

The IEEE 802.11b standard describes four different modulation schemes, with different transmission rates and error characteristics, namely Differential Binary Phase Shift Keying (DBPSK) transmitting with 1 Mbps, Differential Quadrature Phase Shift Keying (DQPSK) with 2 Mbps and two Complementary Code Keying (CCK) schemes with 5.5 and 11 Mbps. The bit error probabilities can be calculated theoretically as a function of the modulation scheme and signal to noise ratio. However, these calculations do not consider possible losses from implementation, which change the performance significantly. As an example Fig. 2.7 shows the theoretical and measured bit error performance of the Intersil HFA 3861B WLAN processor, taken from [62]. We see that the theoretical values for all modulation schemes are much better (up to 4 dB) than the measured ones.

These measured bit error rates are used for our investigations. The bit error rate performance vs E_c/N_0 follows from adding the processing gain, which can be calculated from the received signal strength as described above.

2.3.3 Transaction Error Probability

As described above, MAC layer data packets are acknowledged. A transmission attempt fails if either the data packet or the acknowledgement is corrupted by at least one bit

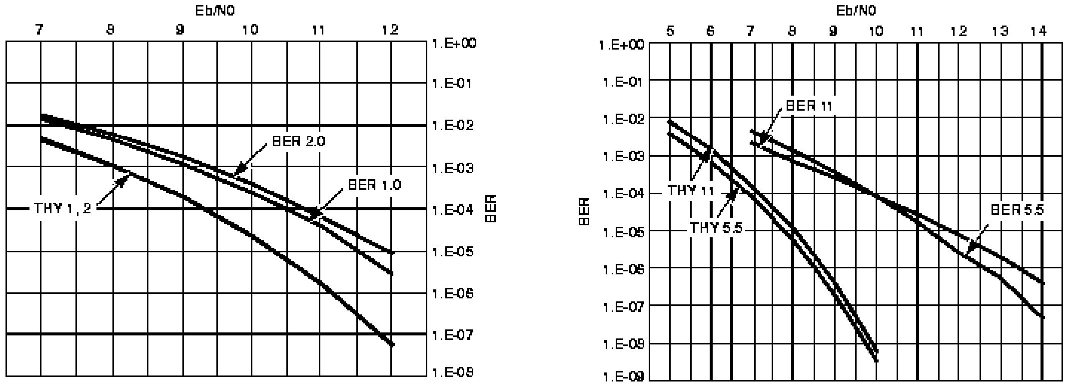


Figure 2.7: Theoretical and measured bit error rates vs. signal to noise ratio for WLAN modulation schemes. On the left side the measured (BER) and theoretical (THY) bit error rates for the 1 Mbps DBPSK and 2 Mbps DQPSK mode are shown; the right figure shows the 5.5 Mbps and 11 Mbps CCK modes.

error. As described in Section 2.1, the data and acknowledgement packets have a PLCP preamble and header of length L_{plcp}, which is transmitted at 1 Mbps. Additionally, the data packet has MAC header of size L_{mac} and the IP payload of size L. The acknowledgement has MAC data of size L_{ack}, but no further payload. The probability of a failed transmission attempt for an IP packet of length L is then given as:

$$p_e(L, m, d) = 1 - \left((1 - p_b(m, d))^{(L+L_{mac}+L_{ack})} \cdot (1 - p_b(m_1, d))^{2L_{plcp}}\right), \tag{2.8}$$

where $p_b(m, d)$ denotes the bit error probability at distance d and transmission mode m. m_1 is the 1 Mbps transmission mode.

Summarizing the bit level model, the error probability of a data packet's transmission and acknowledgement is described to be depending on the length of the data to be sent, the used modulation scheme and the distance between sender and receiver.

2.4 Packet Level Model

In this section the packet level part of the WLAN model is described. Here, the channel access according to the DCF is modeled. Therefore, first the message exchange is described, that is involved in transmitting an IP packet from the access point to a mobile station. Because these messages can get lost and might need to be retransmitted, the delay of data transmission is variable depending on the error probability. The channel access delay, before the station can start transmitting the packet is depending on its history, i.e. the more transmission attempts have failed, the higher the mean channel access delay. Having calculated the transmission and channel access times, it is shown how the throughput can be optimized, by dynamically choosing the modulation scheme depending on the location of the user.

2.4.1 Channel Access Time

The DCF is here considered as the only MAC protocol, as this is the most commonly configured channel access method. As described in Section 2.1, the DCF in WLAN coordinates the channel access of all stations. Thereby, all stations that want to transmit will first wait a fixed time T_{difs}. If after that time the channel is still idle, they draw a random backoff time. This time is continuously being reduced, while the channel is still sensed idle. After expiration, the station is allowed to access the channel, i.e. send the data packet. After a station has sent data, it waits for an acknowledgement from the receiving station. If this is not received, the data packet is assumed to be lost and scheduled for retransmission. For each retransmission the interval, from which the random backoff value is drawn is increased. The number of transmission attempts is limited. If the maximum number of transmission attempts is exceeded, the packet is discarded.

Internet Protocol (IP) packets that are to be transmitted via WLAN are encapsulated: The MAC layer adds a MAC header of size L_{mac}, containing fields such as source and destination addresses. These MAC PDUs are further encapsulated by a preceding PLCP header consisting of a PLCP preamble, a PLCP header and an end-delimiter having a size of L_{plcp}. The PLCP header contains the information at which data rate the MAC PDU will be transmitted. The transmitting station decides the transmission mode. There are two different PLCP modes defined: Long and Short Preamble. Here, the Long Preamble mode is assumed only, where the PLCP preamble and header are transmitted at 1 Mbps. In this mode, the transmission of the PLCP preamble and header needs the time $T_{plcp} = 192\ \mu$s. Having a data packet of size L, the transmission time for transmission mode m is given as

$$T_{tx}(L, m) = T_{plcp} + \frac{L + L_{mac}}{R(m)}, \tag{2.9}$$

where $R(m)$ is the bit rate of transmission mode m.

Assuming the previous packet was transmitted immediately before this packet the channel first needs to be sensed idle for time T_{difs}. Then the backoff value b is drawn from the contention window. b is a discrete random variable uniformly distributed between 0 and CW_{min}. Every time slot, T_{slot}, that the channel is sensed idle, the backoff value is reduced by one. The channel access time for the first transmission attempt can then be described as

$$T_{ca}(1) = T_{difs} + b \cdot T_{slot}. \tag{2.10}$$

Thus it follows for the expectation value of the channel access time for the first transmission attempt

$$\overline{T_{ca}(1)} = T_{difs} + \frac{CW_{min}}{2} \cdot T_{slot}. \tag{2.11}$$

After the packet is received, the receiver waits for time T_{sifs} and then sends an acknowledgement. The acknowledgement has a length of L_{ack} plus the PLCP preamble and header. The acknowledgement is transmitted at the same rate as the data frame was received and the PLCP preamble and header is again transmitted at 1 Mbps, requiring the transmission time T_{plcp}, as given above. Thus the transmission time of the acknowledgement is given as

$$T_{tx_{ack}}(m) = T_{plcp} + \frac{L_{ack}}{R(m)}. \tag{2.12}$$

Now, that the data packet's and MAC acknowledgement's transmissions times and the channel access times before and in between these are calculated, we can sum them up to get the overall channel busy time for the first transmission attempt as

$$T_b(1, L, m) = T_{ca}(1) + T_{tx}(L, m) + T_{sifs} + T_{tx_{ack}}, \tag{2.13}$$

or its expectation value as

$$\overline{T_b(1, L, m)} = \overline{T_{ca}(1)} + T_{tx}(L, m) + T_{sifs} + T_{tx_{ack}}. \tag{2.14}$$

If after this time the station has not received an acknowledgement, it assumes that the data packet was lost and initiates a retransmission. The retransmission attempt is similar to the first transmission attempt, only that the maximum contention window is increased to

$$CW_{new} = 2 \cdot (CW_{old} + 1) - 1. \tag{2.15}$$

For further transmission attempts the contention window is further increased. However, a maximum of five transmission attempts is allowed, before the packet would be discarded.

The mean channel access delay for the k-th transmission attempt is thus given as

$$\overline{T_{ca}(k)} = T_{difs} + \frac{2^{k-1}(CW_{min} + 1) - 1}{2} \cdot T_{slot}, \tag{2.16}$$

where $k = 1..5$ is the number of the transmission attempt. The mean channel busy time for k transmission attempts is then given as:

$$\overline{T_b(k,L,m)} = k \cdot (T_{tx}(L,m) + T_{sifs} + T_{tx_{ack}}(m)) + \sum_{i=1}^{k} \overline{T_{ca}(k)}. \qquad (2.17)$$

2.4.2 Throughput and Rate Adaptation

The transmission attempt duration was calculated above and from the bit level model the transaction error probability is given. Putting that together, the expected throughput, that users with a distance d to the access point can achieve can be calculated. The throughput is here defined as the maximum MAC layer goodput that a user can get, if he offers a load higher than the capacity of the link and is the only sending station. The expectation value for the throughput is then given as:

$$\overline{TP(L,m,d)} = \sum_{k=1}^{5} \frac{L}{\overline{T_b(k,L,m)}} \cdot p_k(L,m,d), \qquad (2.18)$$

where p_k denotes the probability that a packet is transmitted in the k-th transmission attempt and is given as follows:

$$p_k(L,m,d) = (1 - p_e(L,m,d)) \cdot p_e(L,m,d)^{(k-1)} \qquad (2.19)$$

The calculated expectation value for the throughput is shown in Fig. 2.8. All packets are assumed to have a data length of 1000 Bytes. The highest bit rate modulation scheme, which can provide 11 Mbps physical bit rate, offers under perfect channel conditions a MAC layer throughput of less than 5.5 Mbps. This is due to the channel access protocol on the one hand and the preamble transmission at a bit rate of only 1 Mbps on the other hand. At distances of about 70 m from the access point, the throughput of this modulation scheme decreases until it reaches zero at about 90 m distance. This decrease in throughput happens, because the packets are received with a lower signal strength and thus the bit and packet error probability increases. First, this is compensated with retransmissions, and thus the throughput is degraded slightly, but at distances of more than 90m the throughput drops to zero, because all packets have to be discarded. Similar results are shown for the other modulation schemes, where the drop distances however increase with a decreasing bit rate.

From Fig. 2.8 it is obvious, how rate adaptation works: The modulation scheme is chosen such, that the throughput is maximized. Hence, in near distances to the access point, terminals can achieve high throughput with the highest possible modulation scheme. If the terminal is located further away from the access point, a different modulation scheme is used, that has a lower channel bit rate, but a lower error rate as well.

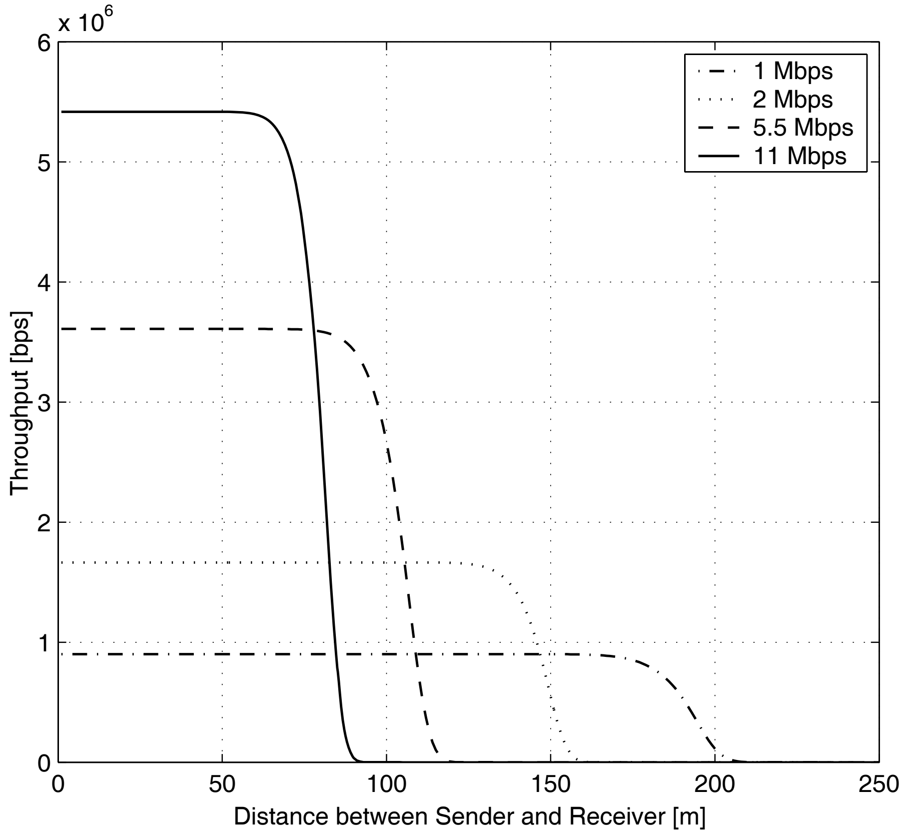


Figure 2.8: Throughput depending on the used modulation scheme and distance between sender and receiver

In the coming analytical investigations, perfect rate adaptation regarding throughput is assumed, i.e. the sending station always chooses the modulation scheme for the transmission that offers the highest throughput.

2.5 Constant Bit Rate Traffic over WLAN

Having described WLAN at bit and packet level, the throughput that a user could achieve under saturation conditions was calculated. Now, multiple users and different traffic conditions are considered. The first considered applications are non-adaptive constant bit rate sources that stream data via the access point to users.

2.5.1 Scenario and Assumptions

Several stations are being served by one access point. All users get data via the access point from non-adaptive sources, e.g. multimedia streaming applications. The packets for each user are assumed to arrive according to a Poisson process, where the same mean packet rate is sent to each user. The arrival rates are chosen such, that the sum over all stations is larger than the access point's capacity. Hence the access point will always have data to send, which means that the access point operates under saturation condition here. The packet size is assumed to be the same for all arriving packets. Then, we are interested in the fraction of bandwidth that each user gets.

As we have seen in the previous section, the distance between the access point and mobile station determines the received signal strength and thus influences the packet transmission delay significantly. This means that the user locations' will have a strong influence on the overall performance. N users are assumed to be randomly located as follows: All users are located within a circular area with radius R around the access point. Within this area, the users are located randomly according to a spatial uniform distribution.

2.5.2 Queuing Model

In order to model the scenario as described above, the access point is substituted by a queuing system. The access point serves all stations and is the only server of the system. The access point has only one queue, which has a capacity of B. The arrival process is the sum of N Poisson processes with the same arrival rate λ_i. Thus we have a Poissonian arrival process with arrival rate $\lambda = N \cdot \lambda_i$. The only non-trivial parameter of the queueing system is the service time distribution. If the locations of the users are known, the service time distribution can be calculated, which will be done in the next section. We can thus model the access point by an M/G/1/B-queueing model.

Now, the service time distribution of this queuing model is derived. Thereby, first the service time distribution of one user located at distance d is calculated. Then the service time distribution for a random user distribution in a location area of a certain radius is determined from that. The service time distribution depends on the user locations and because of the user locations being random, the service time distribution is random as well. Hence, the expectation value for all service time probabilities is calculated assuming the users to be randomly distributed according to a uniform distribution.

First, a single user is assumed to be served, which is located at distance d to the access point. The distance to the access point determines the mean received signal strength and thus the used modulation m according to Fig. 2.8. The probability p_k that the k-th transmission attempt is successful is given in Eqn. 2.19. For each transmission attempt, the overall transmission time is increased by a fixed part and a random part as follows (see also Eqn. 2.17):

$$T_b(k, L, m) \quad = \quad k \cdot T_{fix}(L, m) + \sum_{i=1}^{k} T_c(i) \tag{2.20}$$

Here, $T_c(i)$ is a discrete random variable describing the contention time before the i-th transmission attempt. The variable is uniformly distributed in the current contention window $CW(i)$ according to the following probability mass function:

$$\begin{aligned} f_{T_c(i)}(\tau) &= P\left(T_c(i) = \tau \cdot T_{slot}\right) \\ &= \begin{cases} \frac{1}{CW(i)} & \text{if } 0 \leq \tau < CW(i) \\ 0 & \text{else} \end{cases} \end{aligned} \tag{2.21}$$

The sum over these random variables, $\sum T_c(i)$, is also a discrete random variable. The probability mass function can be found by the convolution of each of the probability mass functions of the contention times:

$$f_{\sum_{i=1}^{k} T_c(i)}(\tau) = f_{T_c(1)} * f_{T_c(2)} * \cdots * f_{T_c(k)}. \tag{2.22}$$

The probability mass function of the busy time, $f_{T_b(k,L,m)}(\tau)$, can be found by shifting by the fixed delay part.

Now, we have calculated the probability mass function of the channel busy time if exactly k transmission attempts are required. The probability that exaclty k transmission attempts are required, p_k was given in Eqn. 2.19 in dependence of the used modulation scheme, the length of the packet and the distance between sender and receiver. For a given user, i.e. a given modulation scheme, packet length and distance to the access point, the service time probability mass function can thus be calculated as

$$f_{T_b(L,m,d)}(\tau) = \sum_{k=1}^{5} p_k(L,m,d) \cdot f_{T_b(k,L,m)}(\tau). \tag{2.23}$$

Now, the service time distribution for the scenario, where users are located in a circular area around the access point with radius R is derived. As said above, inside this area the users are randomly distributed according to a spatial uniform distribution. We want to investigate the effect that the locations of the users have on the overall performance first and not yet the effect that the number of users have. Hence, the users are placed with the density $\rho = N/\pi R^2$ in location areas with the radius R. The density is scaled with the radius of the placement area, in order to have the same amount of users connected to the access point.

All sources are assumed to send packets with a fixed and equal mean packet rate. Further, the queue employs a simple First-Come-First-Serve (FIFO) queuing discipline. This means that all sources have on the average the same number of packets in the queue and the same number of packets being served. This is independent of the service time, which can vary significantly between packets sent to different users. From this, it follows that the probability that the next packet to be sent is destined to a certain mobile station, is equal for all stations, regardless of their location.

In order to determine the overall service time distribution, the area is divided into rings around the access point. The first ring has an inner radius of 0 and outer radius Δr. The ring i has an inner radius of $\sqrt{i-1} \cdot \Delta r$ and an outer radius of $\sqrt{i} \cdot \Delta r$. This gives us, that all rings have the same area $A = \pi \cdot \Delta r^2$ and because the distribution of users is homogenous, the expected number of users in each ring is also equal.

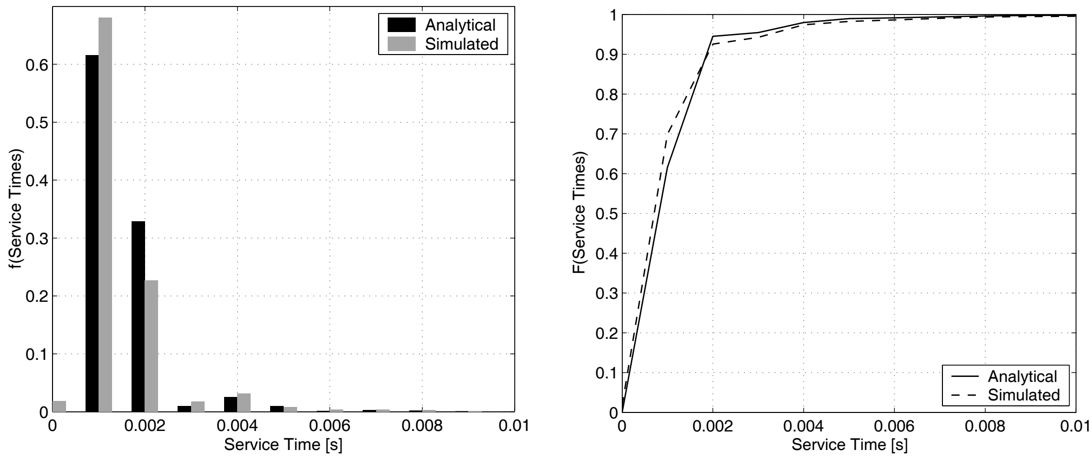
Iterating over all of these rings, their contribution to the service time distribution is calculated. If we assume Δr to be small enough, all locations in area i have roughly the same distance $d = \sqrt{i} \cdot \Delta r$ to the access point. The probability mass function of the service time of one user at distance d is given in Eqn. 2.23. Now the probability that a packet is destined to a user located in area i is required. The expected number of users is the same in all of these areas. Further, as all CBR sources send with the same rate, the number of packets in the queue is the same for all users. Thus, the probability of a packet being destined to a user in one area, is the same for all of these areas and given as $1/i_{max}$. The probability mass function of the service time can then be calculated as:

$$f_{T_b(L)} = \sum_{i=1}^{i_{max}} f_{T_b(L,m,\sqrt{i}\Delta r)}(\tau) \cdot \frac{1}{i_{max}} \tag{2.24}$$

It is important to note, that the shown service time distribution is only valid for a given user distribution. If the area, where users are placed is modified or a different process to place the users in this area is used, the service time distribution will be different.

The service time distribution allows to easily calculate the throughput of each user. Again, we benefit from the fact that all sources have on the average the same number of packets in the queue and thus being served. Hence, each mobile station of that scenario, regardless whether it is placed directly next to the access point or at the edge of the location area, will experience the same throughput. We only have to find the average service rate of the queuing system and divide this by the number of users.

The service rate distribution can be easily found from the probability mass function of the service time, where the probabilities are given for all possible service times. Each service time can be converted into a service rate by calculating $\mu = \frac{1}{T_b}$. From the service rate distribution, the expected service rate is calculated. This is divided by the number of users to get the expected throughput per user.



(a) (b)

Figure 2.9: Histogram (a) and Cumulative Distribution Function (b) of the access point's service time, where users are randomly distributed in a circular area around the access point with radius $R = 100$m

2.5.3 Performance Evaluation and Verification

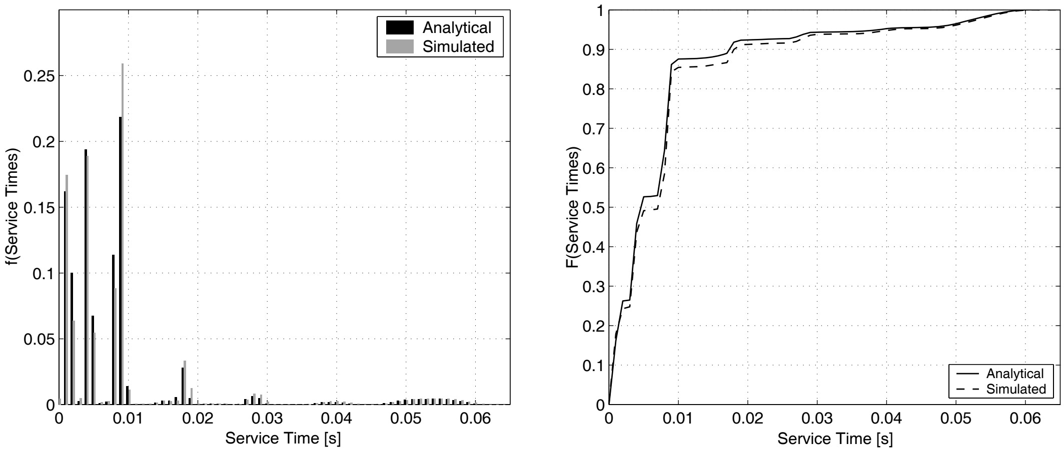
By means of a queuing model formulas for the service time distribution and the throughput per user were determined. In this section the service time distribution and throughput per user are calculated for some sample values and the analytical findings are verified by means of simulations.

Therefore, the scenarios are implemented in the network simulator (see [92] for details about the simulator). I used the existing IEEE 802.11 MAC implementation of the simulator in its version ns-2.26, but additionally implemented the rate adaptation functionality. The rate adaptation is based on the received signal strength as follows: Each node maintains for each other node, where communication is possible, an average received signal strength value. The average is continuously calculated over the last 20 samples. From this value the modulation scheme is selected. The modulation scheme influences the error rate performance as we have described in previous sections.

2.5.3.1 Service Time

First all users are placed in a circular area around the access point with radius $R = 100$ m. This means that most of the users should have very good channel conditions and thus low service times. The measured and calculated service times are shown in Fig. 2.9(a). The cumulative distribution function (CDF) of the service times is depicted in Fig. 2.9(b).

We see that the simulation matches the analytical model fairly well, which can be seen especially well in the CDF plot. As one could expect, the service time is low, i.e. most of the packets are served within one ore two milliseconds. These are the packets that need



(a) (b)

Figure 2.10: Histogram (a) and Cumulative Distribution Function (b) of the access point's service time, where users are randomly distributed in a circular area around the access point with radius $R = 200\text{m}$

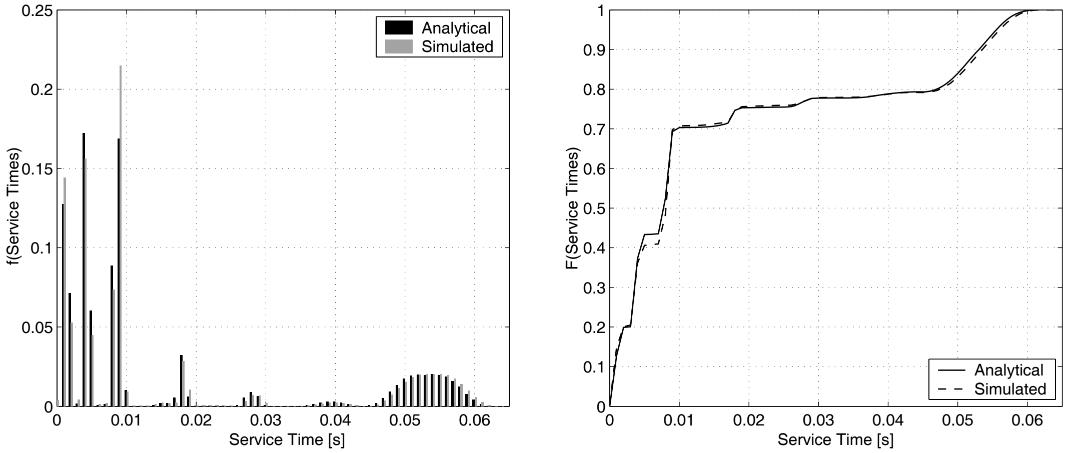
one transmission attempt only and are transmitted with a high bit rate modulation scheme. Because of the close distance between users and access points, the received signal strength is high and these packets dominate this scenario.

However, we see a second smaller peak around four milliseconds. These are the packets, which are transmitted with a lower modulation scheme. Still, the absolute value of the delay is small, because the modulation scheme influences only the MAC PDU transmission, which is only one part of the overall service time, as was shown in the previous sections. It can be seen in the figures, where more users are placed in greater distance to the access point, that the retransmissions attribute more additional delay.

The important thing to note for the verification of the analytical model is that the two peaks can be reproduced reasonably well.

The radius of the location area is now increased to 200m and the results are shown in Fig. 2.10(a) and Fig. 2.10(b). The simulations still match the calculations very well. All peaks in the histogram can be reproduced. Each of these peaks represents a certain modulation / number of retransmissions combination. We see that some of these peaks are now centered at much higher delays, which is attributed to the retransmissions. A retransmission requires a new channel access delay, which is also increased compared to the previous delay, and additionally another transmission delay, which is constant. This delay is more significant than a lower bit rate modulation scheme.

We see especially, that the peaks at higher service times get broader. This is because the number of retransmissions increases and thus the probability mass function is generated using the convolution of the contention window distributions. If two distributions have x_1 and x_2 different values, where the probability is non-zero, the convolution of these distributions has $x_1 + x_2$. Hence, the convolution broadens the peaks.



(a) (b)

Figure 2.11: Histogram (a) and Cumulative Distribution Function (b) of the access point's service time, where users are randomly distributed in a circular area around the access point with radius $R = 220\text{m}$

In Fig. 2.11(b) and Fig. 2.11(a) the same results are shown for users located in an area of 220m radius. Again, the simulation matches the calculations. Here we see that the last peak, which corresponds to the case of five transmission attempts at the lowest modulation scheme, is increased at the cost of the other peaks. Users at distances of more than 200m only have a very low probability of receiving any packet correctly.

As the user distribution is homogeneous over the area and the area grows by r^2 with the radius, the higher distances to the access point dominate in the user distribution. That is why we can see the whole service time distribution being shifted significantly to higher delay values, by increasing the location area radius only by 10%.

2.5.3.2 Throughput

The service time distribution shows already a strong dependence on the user locations. Now, we want to see the effect of the user locations on the throughput that each user receives. Therefore 25 scenarios are considered, where the users are placed in a location area with radius between 10m and 250m. The throughput of each of these is measured and the mean and variance is calculated. Together with the theoretical throughput the result is depicted in Fig. 2.12.

Again, we see that the analytical model is verified by the simulations: the measured average throughput matches the analytically received values.

We see that the variance of the measured throughput is very low, which means that all users are getting roughly the same throughput. As explained before, this is because packets to all destinations have equal probability to be served by the access point and it is here verified by the simulations.

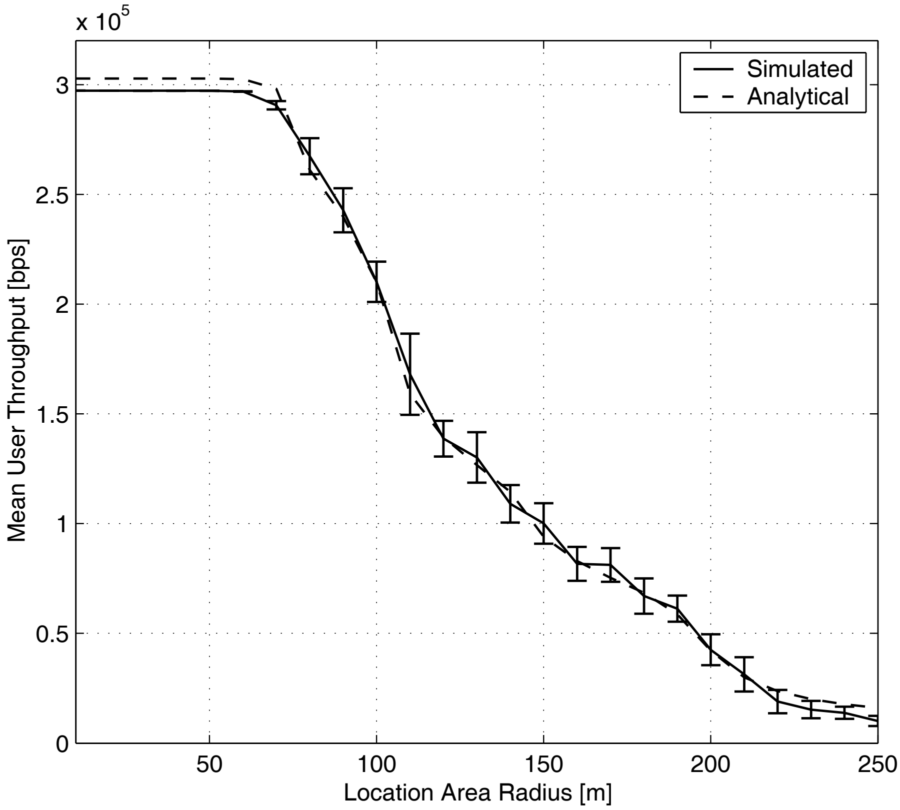


Figure 2.12: Throughput per user (for the simulated curve the mean and variance over all users is shown)

This means the throughput of a user is determined mainly by the location of other users. We see that in this case the throughput of all users may degrade from 300 kbps to 10 kbps only because users are spaced further apart from the access point. The reason for this behavior is that all users access the channel with equal probability, but stations with lower modulation schemes and/or higher transmission attempts dominate in time. So, much more time is spent on serving the users that are located at greater distance, while the users next to the access point will get the same packet rate as those. What is more, because the area where stations can be located increases with r^2, more stations are located in greater distance. This results in the drastic throughput degradation in scenarios, where users are located in larger areas.

We have seen that the throughput of any user is determined by the locations of all users. Now, we want to see the effect that one or few users which change their location, have on the performance of a group of users, that stays directly next to the access point. This is a common scenario that one could see when a user is connected to an access point and downloading some data as some others do as well. Then, this one user starts moving away from the access point, but does not terminate the connection. The access point will still try to send all packets for this user. The transmission will require a long time, because of a low bit rate modulation scheme being used and possibly many retransmissions. During that time all other packets wait in the queue of the access point. Hence, it is clear that the throughput of all other users decreases. However, the extend of this throughput degradation would be interesting to know.

We show this in Fig 2.13. In case of one user moving away from the access point, we see already a huge performance degradation of all users: The throughput drops by roughly 75%. In case of two or three moving users the throughput further decreases, but the most impressive result is the huge influence that the location of one user has on the performance of all others.

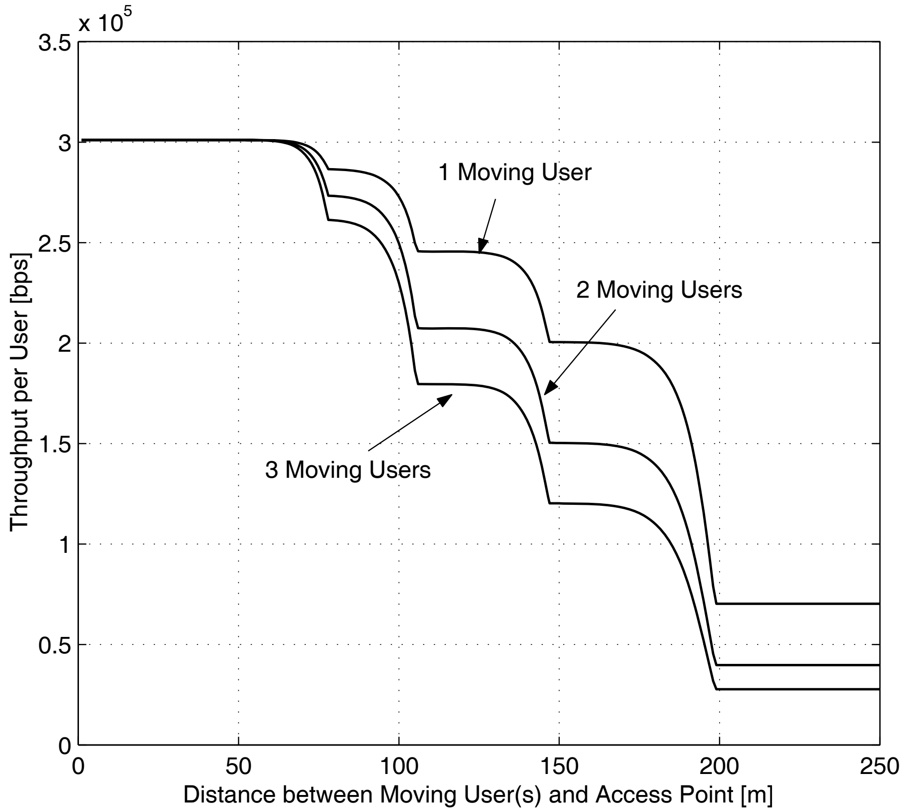


Figure 2.13: Throughput per user for the case of one or few user(s) moving away from the access point

This scenario shows the problem of an open WLAN: As no admission control or dropping policy is implemented, few users in large distances can degrade the throughput of all others and lead to a highly inefficiently run access point.

2.6 TCP Traffic over WLAN

A strong dependency of the user performance on all users' locations was shown for CBR traffic in the previous section. In this section a different application is considered and it is investigated whether the performance dependence is persistent. Most of the applications that use the Internet today are still based on the Transmission Control Protocol (TCP). Hence, the traffic seen in the Internet and also in WLAN Hot Spots is mainly consisting of TCP packets. In this section the performance of TCP connections over WLAN is evaluated.

The scenario is depicted in Fig. 2.14. Users are assumed to download data from Internet servers via the access point. The users are distributed around the access point, located at distances d_n. The TCP connections have a basic round trip time of $RTT_0(n)$. This round trip time is attributed to the Internet link and does not include the queuing delay at the access point and the transmission time of data over the wireless link. The Internet link is further assumed to contribute a packet drop rate of p_I, where packet drops occur independently from each other. The TCP packets of a user n arrive at the access point with a packet rate of λ_n. The TCP sender in the Internet hosts adapt this packet rate according to the seen overall loss rate and the overall round trip time.

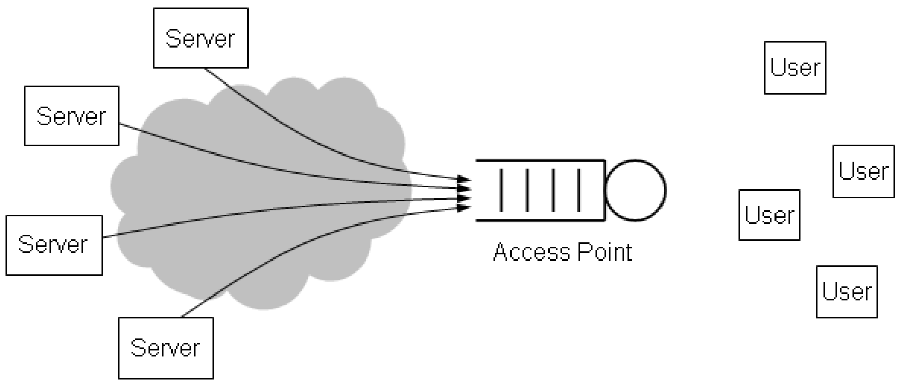


Figure 2.14: Considered scenario of several TCP users downloading data from Internet servers via a single WLAN access point.

2.6.1 TCP Model

According to the well known TCP model of Padhye et al. [94] the packet rate of a TCP flow depends on the flows round trip time and the packet loss rate. In our case the arrival rate of one flow can be described as:

$$\lambda_n = \frac{1}{RTT(n)}\sqrt{\frac{3}{2 \cdot p(n)}} \tag{2.25}$$

where $RTT(n)$ denotes the round trip time of the user n and is given as the sum of the basic round trip time $RTT_0(n)$, the queueing delay in the access point, T_q, the transmission time of the data packet, $T_{td}(n)$, and the transmission time of the TCP Acknowledgement, $T_{ta}(n)$:

$$RTT(n) = RTT_0(n) + T_q + T_{td}(n) + T_{ta}(n) \tag{2.26}$$

The loss probability $p(n)$ is composed of the losses in the queue of the access point, due to queue overflow, p_q, the losses over the wireless link of the data packet, when the last MAC retransmission attempt has failed, $p_w(n)$ and the loss probability of the Internet link, p_I:

$$p(n) = 1 - [(1-p_q)(1-p_w(n))(1-p_a(n))(1-p_I)] \tag{2.27}$$

Losses of TCP acknowledgements are not considered to influence TCP's throughput, because of the cumulative acknowledgements used by TCP.

2.6.2 Access Point Model

The access point is again modeled as a queuing system, having one queue and one server. The queue has a length of B, in which packets from all down-link TCP connections arrive. The server works according to the DCF protocol.

2.6.2.1 Arrival Process

All TCP sources send their packets via the access point to the users. Hence all flows multiplex into a single packet stream at the access point. As it was done for example in [80] and [6], we assume that the packets of the multiplexed stream arrive according to a Poisson process, having an arrival rate which is the sum of all users' packet arrival rates:

$$\lambda = \sum_{n=1}^{N} \lambda_n \tag{2.28}$$

2.6.2.2 Service Time Distribution

The service time is depending on the destinations of the packets that are transmitted, because of the channel busy times dependence on the location of the user. In Section 2.5 the service time distribution for fixed user locations and a fixed rate of each source was calculated and verified.

The service time is a discrete random variable and hence can only adopt a limited number of values. However, the number of different service times that can occur is large as can be seen from e.g. Fig. 2.10(a). For the investigations here, we approximate the service time distribution to fewer discrete values and their probabilities. This allows solving the queuing system for its packet loss rate and packet sojourn time numerically in reasonable computation time.

In Fig. 2.10(a) and Fig. 2.11(a) we have seen that there are certain peaks in the distributions. These peaks correspond to certain number-of-retransmission and modulation-scheme combinations, e.g. the first peak represents the packets that are transmitted in the first transmission attempts using the highest modulation scheme. The values around the middle of the peak stem from the fact that the contention time is randomly chosen from a uniform distribution, depending on the maximum contention window.

For the approximation, only the expectation values of the service times are used for each of these number-of-retransmission and modulation-scheme combinations, i.e. all peaks are approximated by their centre value. The expected service time of all combinations is already given in Eqn. 2.17. The probability that this service time happens for a particular packet depends on the destination node's location. The location determines the used modulation scheme, which is chosen to optimize the throughput according to Fig. 2.8. Further, the location and modulation scheme determine the probability that exactly k transmission attempts are required.

In Eqn. 2.19 the probability that exactly k transmission attempts are required, p_k, was calculated considering only transmission errors due to corruption. This was correct for the case, where we considered only downlink flows. Now, that TCP as the transport protocol is assumed TCP acknowledgements are transmitted uplink. Collisions are possible between downlink data packets and uplink TCP acknowledgements. This effect of a flow's self-collisions was already described in [120].

In the special case of all TCP flows going downlink and all TCP acknowledgements going uplink, the probability that a TCP packet collides with a TCP acknowledgement can be easily calculated as follows: A TCP data packet is transmitted from the access point to a terminal. The MAC layer PDU is received by the terminal and given to the TCP socket, which generates a TCP acknowledgement and schedules this for transmission at its MAC layer. During this time the MAC acknowledgement was already received by the access point, which has chosen a backoff time for the next packet to be transmitted. The backoff time was chosen from an interval of CW_{min} possible values, where one of these values leads to a collision with the TCP acknowledgement that is to be transmitted from the terminal to the access point. Which backoff time would lead to a collision depends on the time that the TCP socket needs to produce an acknowledgement and pass this to the MAC layer. The important thing is that it is exaclty one value and thus the probability to chose that one value out of CW_{min} values is $p_c = 1/CW_{min}$.

In the calculation of the collision probability, it was assumed that only one TCP acknowledgement and one data packet are contending for channel access. It is possible that more than one TCP acknowledgement is waiting to be transmitted. However, this is a rare event as it can only happen, when the data packet wins the channel contention against a TCP acknowledgement. The access point has to backoff for the channel access, as it has just transmitted a packet. The terminal senses the channel idle at the time the TCP acknowledgement is to be transmitted and thus only has to wait the time T_{difs}. If a collision had happened, the probability that the access point is able to access the channel first is increased, but the collision itself is a rare event. Given these reasons, the case that several acknowledgements are waiting to be transmitted is neglected. It will be shown in the performance evaluation and verification that this assumption can be made.

Now, the probability that exactly k transmission attempts are required is calculated as:

$$p_k = (1-p_e)(1-p_c) \cdot (1-(1-p_e)(1-p_c))^{(k-1)} \tag{2.29}$$

Considering the uplink traffic, the expected transmission time also has to be adjusted. During a TCP acknowledgement is sent, the access point can neither transmit itself a packet, nor reduce its backoff counter, because the channel is not sensed idle. As reasoned above, it is assumed that in most cases the data packets have to wait for the TCP acknowledgements to be transmitted. During the transmission of the TCP acknowledgement, the backoff timer for the data packet is paused. Hence, we have to add the transaction time of a TCP acknowledgement to the expected service time. The TCP acknowledgements are transmitted with a modulation scheme that is determined from the distance between the access point and the acknowledging station. This station is the receiver of the previous TCP data packet. Hence, we can add it to the transaction time of that TCP data packets transmission, where the transaction time of the acknowledgement is calculated using the same modulation scheme as the TCP data packet. The service time then still is depending only on the number of transmission attempts and the used modulation scheme and the probability for a certain service time is then given as:

$$P(\overline{T_b} = X \mid d = r) = \begin{cases} 0 & \forall\, T_b(k,m,L),\ m \neq m(r) \\ p_k & \forall\, T_b(k,m,L),\ m = m(r) \end{cases} \tag{2.30}$$

Thereby, $m(r)$ is the modulation scheme that optimizes the MAC layer throughput at distance r according to Fig. 2.8. p_k is the probability that exactly k transmission attempts are required considering corruption and collisions as given in Eqn. 2.29.

The fraction of packets in the queue for user n is proportional to the arrival rate of his flow in relation to the sum of all arrival rates. The probability that the packet, which is currently being serviced, is for user n is thus given as λ_n/λ. Now, we can calculate the service time distribution as:

$$f_{T_b} = \sum_{n=1}^{N} P(T_b = X \mid d = r) \cdot P(d = r) \tag{2.31}$$

$$= \sum_{n=1}^{N} P(T_b = X \mid d = r) \cdot \frac{\lambda_n}{\lambda} \tag{2.32}$$

2.6.2.3 Markov Chain Analysis of the Queueing System

The access point can now be described as an M/G/1/B queueing system, where the packets arrive according to a Poisson process with an arrival rate λ. The service time can only adopt discrete values and the distribution thereof was calculated as shown above. The service time distribution depends thereby on the number and locations of the users and their number of packets in the queue in relation to the total number of packets in the queue.

Having the arrival rate and service time distribution, now the loss probability p_q and the mean time spent in the system T_q can be calculated. The queueing system is thereby described as a Markov chain as shown in Fig. 2.15. The states of the chain are named with tuples (s, t), where s denotes the number of packets in the queue and t is the index of the service time T_t of the packet that is currently being serviced. An exception is state 0, where no packet is in the queue.

In each of the states a packet is currently being serviced with service time T_t. At the time the service of this packet has started s packets were in the queue. The transitions between states happen only, if a new packet is being serviced. The probability that a service time T_t is chosen is denoted by $P(T_b = T_t) = P_{T_t}$. The probability of x new packets arriving in time T_t follows from the arrivals according to a Poisson process as

$$P_x(T_t) = \frac{(\lambda \cdot T_t)^x}{x!} e^{-\lambda T_t}. \tag{2.33}$$

The transition probability from state (s_1, t_1) to state (s_2, t_2) is denoted by $p_{s_1 t_1; s_2 t_2}$ and the stationary probability of being in the state (s, t) is given as $\pi_{s,t}$.

The transition probabilities are then given as:

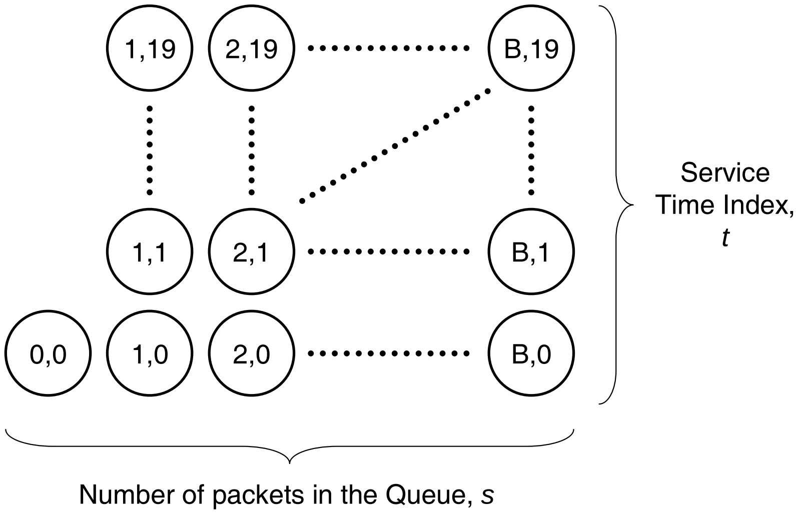


Figure 2.15: Markov chain of the M/G/1/k queueing model of the access point.

$$p_{s_1t_1;s_2t_2} = \begin{cases} P_{T_{t_2}} & \text{if } s_1 = 0 \ \wedge s_2 = 1 \\ P_{s_2-s_1+1}(T_{t_1}) \cdot P_{T_{t_2}} & \text{if } s_2 \geq s_1 - 1 \wedge \\ & \quad s_2 < B-1 \ \wedge s_2 > 0 \\ \left(1 - \sum_{i=0}^{B-s_1-1} P_i(T_{t_1})\right) P_{T_{t_2}} & \text{if } s_2 = B-1 \ \wedge s_1 > 0 \\ P_0(T_{t_1}) & \text{if } s_1 = 1 \ \wedge s_2 = 0 \\ 0 & \text{else} \end{cases} \tag{2.34}$$

The first case describes the transition probabilities from the state with an empty queue to any state with one packet in the queue. The probability is given by the distribution of the service time, as calculated above. Other transitions from this state are not possible. Transitions two or more steps to the left cannot happen, because this would mean that at least two packets were serviced, which would be against the state transition definition. In the second case, $s_2 - s_1 + 1$ packets have arrived during the service of the current packet. Together with the departing packet the transition goes $s_2 - s_1$ steps to the right. However, this is only true for states with less then $(B-1)$ packets in the queue. Transitions to a state with exactly $(B-1)$ packets in the queue are described with the third case. Here the probability that more than $B - s_2$ packets arrived during the service time is summed up. Then packets would have been lost, and the queue would have been filled at the time just before the transition happens. The transition will move one packet from the queue into the server, leaving $B-1$ packets in the queue. The last case describes transitions from states with one packet in the queue into the state with an empty queue, which happens with the probability that no packet arrives during the current service time.

The stationary distribution $\vec{\pi}$ is found by solving the equation P $\vec{\pi} = \vec{\pi}$, where P denotes the matrix of all transition probabilities.

To calculate the loss probability, the expectation value of the number of lost packets during one packet being serviced is calculated. Iterating over all states, the probability that more packets arrive during the service time than would fit into the queue is calculated. The sum over these probabilities multiplied with the number of lost packets results in the expectation value of lost packets:

$$\mathrm{E}\left\{\text{\# losses}\right\} = \sum_{s=1}^{B}\sum_{t=1}^{20}\pi_{s,t}\left[\sum_{i=B-s+1}^{\infty}(i-1)\cdot P_i(T_t)\right]. \tag{2.35}$$

The expected number of arrivals per serviced packet can be calculated similarly as:

$$\mathrm{E}\left\{\text{\# arrivals}\right\} = \sum_{s=1}^{B}\sum_{t=1}^{20}\pi_{s,t}\cdot\lambda\cdot T_t. \tag{2.36}$$

The loss rate is then calculated as:

$$p_q = \frac{\mathrm{E}\left\{\text{\# losses}\right\}}{\mathrm{E}\left\{\text{\# arrivals}\right\}}. \tag{2.37}$$

To calculate the expected time spent in the system, the mean service time, $\overline{T_b}$, is first calculated from the service rate distribution function. The expected time spent in the system is then given as:

$$T_q = \sum_{s=1}^{B}\sum_{t=1}^{20}\pi_{s,t}\cdot s\cdot\overline{T_b}. \tag{2.38}$$

We can see here already, that the time spent in the system and the packet loss rate in the queue are both depending on the locations of the users, because of the service time distributions depending on the locations. Further, they are influenced by the arrival rate of all users, because again these influence the service time distribution and additionally the expected number of arrivals as calculated above. This dependency will be shown in more detail in the following throughput calculation and performance evaluation.

2.6.3 TCP Throughput Calculation

Now, the loss probability and sojourn time of the queueing system are shown to be depending on the packet arrival rate and service time distribution. On the other hand, we have the packet arrival rate depending on the user distribution, the queueing delay and the loss probability. This results in a fix-point problem, which can be described by a system of non-linear equations. For each set of user locations these equations can be solved to get the TCP transmission rates.

2.6.4 Performance Evaluation and Verification

The model, described in the previous sections, should be used to evaluate the performance of TCP users and again identify the parameters, on which this mostly depends. Therefore, first two scenarios are investigated, where the density and number of users is different, but the area in which they are located is the same. Then, two scenarios are considered, where the same number of users is located either in near proximity of the access point, or the users are divided into two groups. One of these being located near the access point, the other at a greater distance. For these scenarios the TCP throughput of each user is calculated. To verify our results, the scenarios are also simulated using the same simulation environment as described in Section 2.5.3. For all following investigations it is assumed that the Internet link contributes $RTT_0 = 200$ms fixed delay and $p_I = 1$% loss rate.

2.6.4.1 Uniformly Distributed Users

For this investigation the users are assumed to be located at equidistant distances from the access point. In the first scenario, $N = 10$ users are located at 20m, 40m, ... , 200m distance from the access point. In a second scenario the density is doubled. Thus we have 20 users placed at 10m, 20m, ... , 200m distance from the access point. Thereby, we placed the users on a straight line originating at the access point. However, placing the users anywhere on the cirlce with the same distance to the access point would give the same results. The results are depicted in Fig. 2.16.

First thing to note is that the analytical model fits the simulation results very well. The analytical values are always within the shown 95% confidence intervals of the simulations.

Further we see that within each scenario, nearly all users get the same TCP throughput, regardless of their location. This is because they all share the same access point queue and thus suffer from the same packet drop rate at entering the queue and the same delay before leaving the queue. This is in contrast to a first intuition, suggesting that the users which are located in proximity to the access point should get a higher throughput.

However, there are few users that achieve less throughput. As said before, the TCP throughput depends on the Round Trip Time (RTT) and the overall loss rate. The loss rate on the wireless link (at the TCP layer losses occur only, if all transmission attempts at the MAC layer have not been successful) influences the overall loss rate as well. Even though this loss rate is negligible for most of the users, the user located at $d = 100$m suffers from losses on the wireless link. This is, because the rate of the wireless link is adapted to maximize the MAC layer throughput. The MAC layer throughput is maximized, by tolerating few losses, but using a faster modulation scheme. These few losses indicate congestion to the TCP sender, which in turn reduces the sending rate. The next user located at $d = 120$m does not suffer from these losses, because the MAC layer uses a lower modulation scheme already. The loss rates are then again very high at the border of coverage, where only few packets are received correctly, even though the lowest modulation scheme is used. Here, the TCP throughput drops to zero.

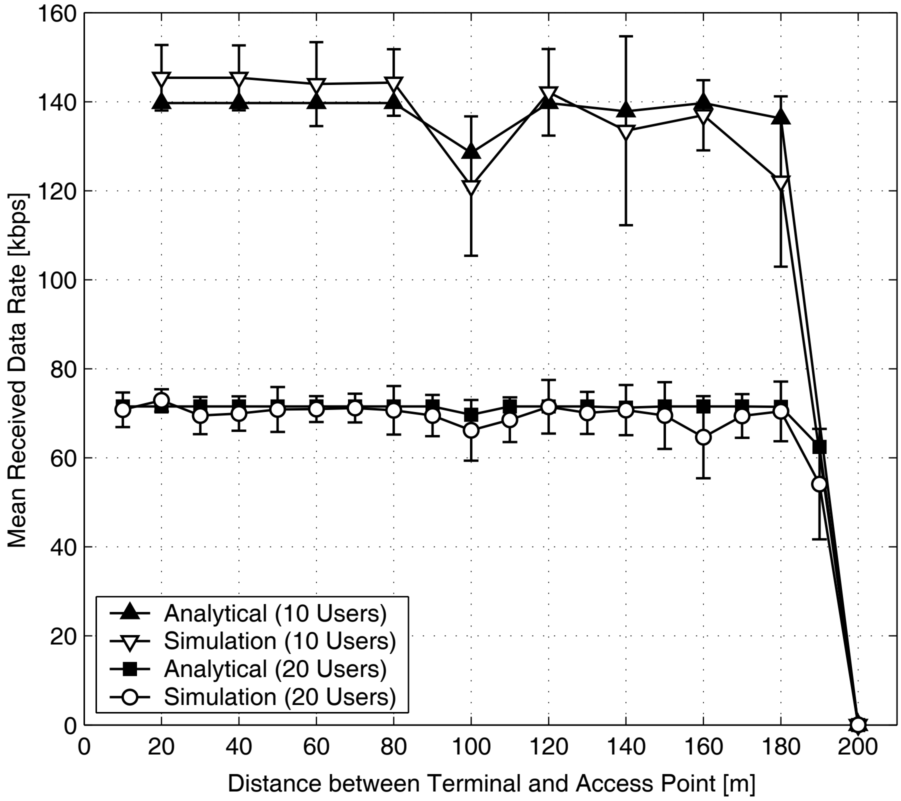


Figure 2.16: TCP throughput for users distributed uniformly along a line from the access point.

We can see that the throughput is in general higher in the scenario with a lower user density. This is clear from common sense, because less users share the scarce resource access point. The throughput is roughly halved, because the user distribution is not changed, but only the number of user is doubled. Hence, the access point can still transmit roughly with the same packet rate.

2.6.4.2 Different User Locations

In this section two scenarios are investigated, where the number of users is the same, but the locations change. In the first scenario, all users are located in near proximity to the access point (maximum distance 50 m). In the second scenario, half of the users are still near the access point, while the other half is located hundred meters further away. The received TCP throughput in presented in Fig. 2.17.

We see that, again, the simulations match the theoretical findings very well. Further we see again, that within a scenario all users get a similar data rate share. However, the rate each user gets, is highly different between these scenarios. Users in the first scenario

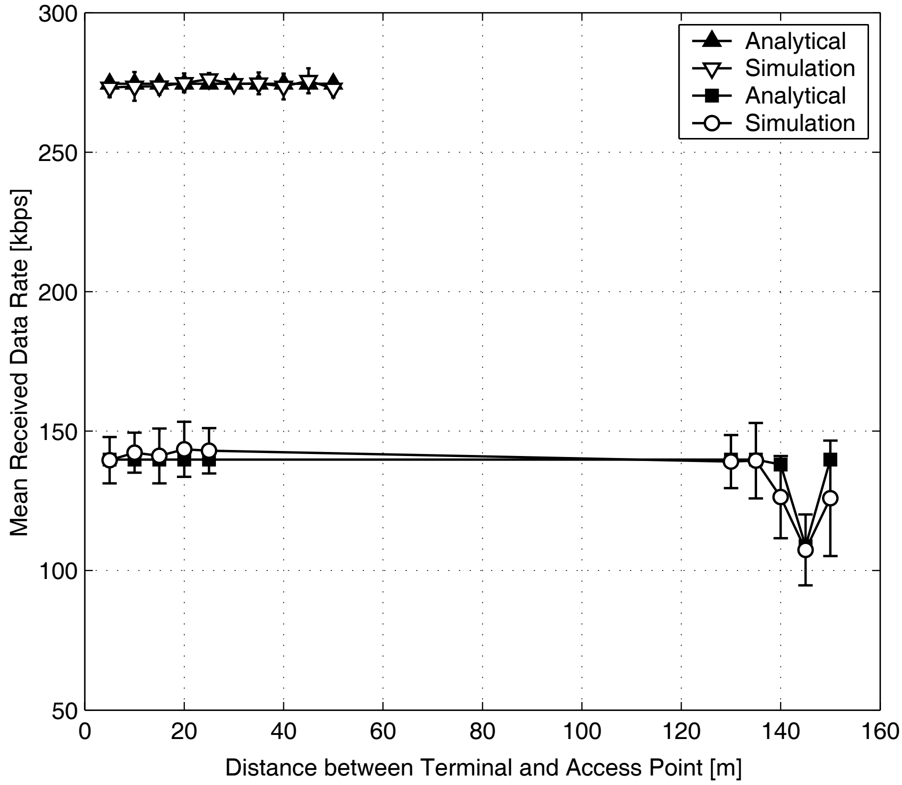


Figure 2.17: TCP throughput for two different user location distributions. In the upper curve all users are located in near proximity to the access point, while in the lower curve half of the users moved 100m further away.

nearly double the data rate, even though the number of users has not changed. Hence, the location of five users reduces the throughput significantly for all users.

The confidence intervals in the first scenario are very small, because all users are very near to the access point. Thus, the packet loss probabilities are very low, even though the highest modulation scheme is used. The low packet loss rates reduce the probability of retransmissions and the contention window is kept at small values. This reduces the randomness and thus the variability of the throughput significantly.

2.6.4.3 TCP Throughput Dependence on User Locations

We have seen that also for TCP throughput the locations of users are one of the most influencing factors. Similar to the evaluation of CBR traffic over WLAN, it is interesting to see the TCP throughput of one user, right next to the access point, changes if the location distribution of the other users had changed.

Again, one user is placed next to the access point and others randomly in a circular area around the access point with a radius R_{max}. Within this area the users are placed randomly with a user density set to $\rho = N/\pi R_{max}^2$. This ensures that the same number of users are placed, regardless of the chosen R_{max}. The maximum radius is varied from $R_{max} = 50$m to $R_{max} = 200$m and the TCP throughput of the one user next to the access point is calculated and measured. The result is shown in Fig. 2.18.

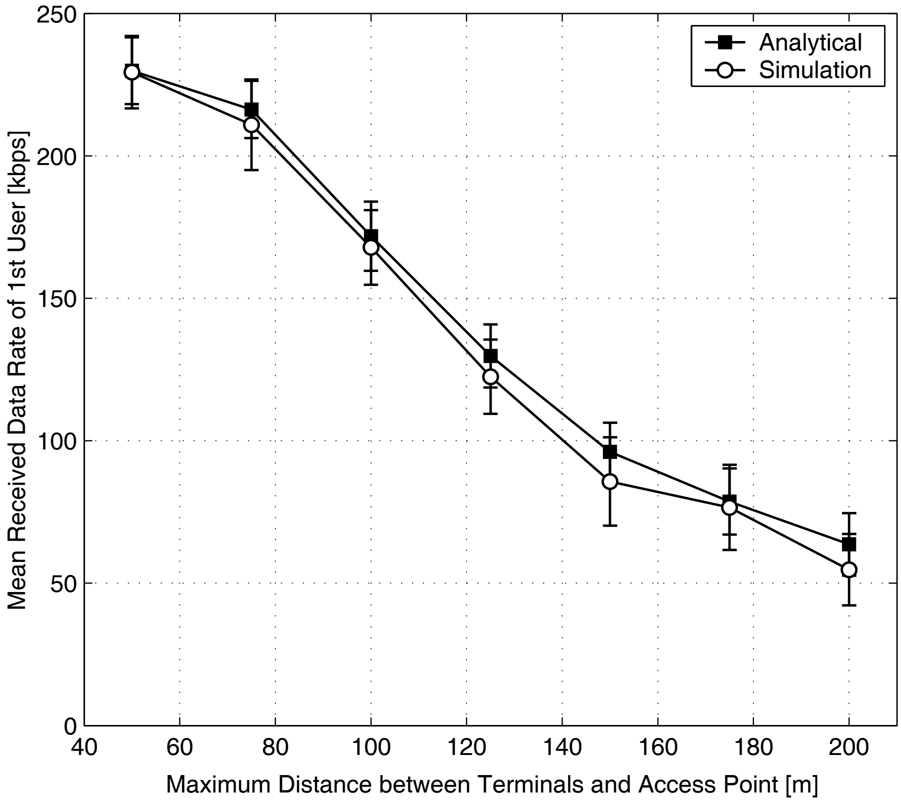


Figure 2.18: TCP throughput of one user next to the access point, while other users are placed randomly with in a circular area around the access point with radius R_{max}.

The most striking result is the extent to which the throughput of a user depends on the locations of other users at the same access point. While the tested user has not moved, the throughput was degraded from about 230 kbps to 60 kbps. This shows also the importance of throughput estimation techniques for WLAN access. Handover decisions should consider user locations as well as quality of service or admission control systems. A user having the choice between different access networks, should not only choose the network according to signal strength or bit error rates, as it is done common handover algorithms or admission control systems today. Instead the quality of service should be estimated, and it is shown with this work, that all users and their locations have to be considered for that.

2.7 Bi-Directional Voice Traffic over WLAN

In this section the performance of bi-directional voice applications over WLAN is evaluated. In the previous section, it was already shown that a station can be modeled with an M/G/1/B queuing system, which was solved for the delay and loss probability. Here, additionally the channel utilization and collision probability will be calculated.

A similar scenario to what was considered in the previous section is assumed here: An access point serves N users. The users are distributed around the access point, located at distances d_n. Each user is running a voice application and is communicating with another voice application which is assumed to be connected to the access point via a lossless line with zero delay. This enables to derive the quality of service parameters of the WLAN without the effects from other (possibly long delay and lossy Internet) links. The voice applications are non adaptive and generate packets according to a Poisson process with a constant packet source rate λ_n.

2.7.1 Modifications of the Queueing Model

In the previous section, the access point was already modeled as an M/G/1/B queuing system. Here, this procedure is extended to model each station as a similar queuing system. However, now that significant bi-directional traffic is considered, modifications to that model are required.

As given above the applications generate packets according to a Poisson process, and hence the arrival process at the terminals is Poissonian with arrival rate λ_n. At the access point N voice applications generate packets and the resulting arrival process is the sum of the single processes and is again Poissonian this time with the arrival rate $\sum \lambda_n$.

The collision probability was already included in the calculation of the probability that a certain amount of transmission attempts are required (see Eqn. 2.29). However, the collision probability is depending on the number of stations that want to access the channel, the channel utilization and the transmission attempt rate of the stations. It will be shown later how to calculate the collision probability from these values.

Considering the uplink traffic, the expected transmission time also has to be adjusted. A station that wants to transmit a packet draws a random backoff timer value and decrements that only when the channel is sensed idle. If another station is transmitting, the backoff timer is paused. This means that the backoff time needs to be divided by $(1-U)$, where U is the utilization of the channel. The service time is then depending on the number of transmission attempts, the used modulation scheme and additionally the channel utilization.

The analysis of the queuing system proceeds similar to the presented analysis in the previous section, which results in the calculation of the loss probability and the sojourn time.

2.7.2 System of Equations

Now, a system of equations is described, which is used to calculate the voice over rate-adaptive WLAN performance. First the collision probability is calculated and the probability that the channel is idle. From that the utilization of the channel is derived. Then, we can calculate the service time distributions of all stations and solve the corresponding queuing systems for loss probability and delay. The loss probability is then used to calculate the rate with which transmission attempts are started.

Each station may need several MAC layer transmission attempts for each successful IP packet transmission. Similar to many related work articles, e.g. [15], it is assumed that each station starts transmission attempts with exponentially distributed interarrival times. It will be shown in the performance evaluation, that this assumption is valid. In our case we have processes with rate $\lambda_{tx,n}$ for the terminals and $\lambda_{tx,0}$ for the access point (the access point is denoted as station 0 in variables and equations). It will be shown later how to calculate these values.

If the channel is idle for at least the time T_{DIFS}, we can calculate the probability that station n does not start a transmission in a time slot as

$$p_{s,n} = e^{-\lambda_{tx,n} \cdot T_{slot}}. \tag{2.39}$$

The probability that station n starts a transmission in a time slot is thus

$$p_{tx,n} = 1 - p_{s,n}. \tag{2.40}$$

The probability that all stations do not start a transmission in a time slot and hence the channel is idle for this time slot is given as

$$p_i = \prod_{n=0}^{N} p_{s,n}. \tag{2.41}$$

The channel idle time, i.e. time until the next transmission attempt happens is geometrically distributed and so the mean value is given as

$$\overline{T_i} = \frac{1}{1 - p_i}. \tag{2.42}$$

The probability that exactly one station starts sending in a time slot is given as:

$$p_{tx=1} = \sum_{i=0}^{N} \left[p_{tx,i} \cdot \prod_{j=0;j\neq i}^{N} p_{s,j} \right] \tag{2.43}$$

The probability that more than one station starts a transmission in a time slot is given as:

$$p_{tx>1} = 1 - p_i - p_{tx=1}. \tag{2.44}$$

Now, the probability is required that the next transmission attempt, which is started after the idle period, occurs in a collision, which is given as the conditional probability of a collision under the condition that a time slot is not idle. This probability is given as:

$$p_c = \frac{p_{tx>1}}{p_{tx=1} + p_{tx>1}} = \frac{p_{tx>1}}{1 - p_i} \tag{2.45}$$

The utilization of the channel is defined as the fraction of time the stations are not allowed to decrement their backoff timer. It is given as

$$U = \frac{\overline{T}_{tx}}{(\overline{T}_{tx} + \overline{T}_i)}. \tag{2.46}$$

The expected channel idle time until the next transmission attempt is started, $\overline{T}_i$, was already calculated. The transmission time for station n is simply given as the time the station needs to transmit the packet and wait for the acknowledgement, which depends on the used modulation scheme. Additionally the time T_{DIFS} needs to be added, because stations need to sense the channel idle for at least this time before they are allowed to decrement their back-off timer. The average transmission time of the access point is the average transmission time over all other stations weighted with their packet arrival rates. The overall mean transmission time is calculated by averaging over all transmission times of the stations and the access points and weighting these again by their arrival rates.

With the collision probability and the utilization we can calculate the service time distributions of all M/G/1/B queueing systems. The arrival rate is given by the application. Hence, these queuing systems can be solved as it was described in the Section 2.6.2.3 and as results the packet loss rate, p_q and the queuing delay T_q is received.

The queuing system needs to be solved here for all stations in the network, which increases the computational effort significantly. The effort of solving the queuing system of the access point is the same as it was for the TCP traffic analysis. However the effort of solving the queuing systems of the stations is less: The stations use a fixed modulation scheme and thus the number of different service times that can happen is reduced by a factor of four. Thus the number of states in the Markov chain is reduced by a factor of four as well and thus the calculation of the stationary state probabilities is done much faster. This enables to estimate the quality of service for voice communications over WLAN in reasonable computation time.

Having solved the queuing systems for their sojourn time and loss probability, the mean transmission attempt rate can be calculated. The arrival rate at the queue is given as λ_n. However, some packets are dropped, because of the limited queue size and thus the rate is reduced by the dropping rate, p_q. Each packet, however, might need several transmission attempts. The expected number of transmission attempts per queued packet can be calculated using the probabilities p_k as given in Eqn. 2.29:

$$\overline{N}_{tx,n} = \sum_{k=1}^{5} k \cdot p_k. \tag{2.47}$$

The transmission attempt rate is thus given as:

$$\lambda_{tx,n} = \lambda_n \cdot (1 - p_q) \cdot \overline{N}_{tx,n}. \tag{2.48}$$

The transmission attempt rate was used to calculate the collision probability as described above. Thus, we now have a set of non-linear equations, which are solved numerically. In the next section some interesting results from the calculations as well as simulation results that support the findings are described.

2.7.3 Performance Evaluation and Verification

The model, described in the previous sections, is used to evaluate the performance of voice applications and especially identify the parameters, on which this mostly depends. First, the dependence of the voice performance regarding loss rate and delay on the number of users and their locations is presented. Scenarios are set-up with the number of users varied from $N = 10$ to $N = 20$. Further the scenarios are varied by the locations of the users. Again, all users are placed randomly in a circular area around the access point with a radius R. This maximum distance to the access point is varied from $R = 25$ m to $R = 250$ m.

To validate the analytical findings simulations of these scenarios are performed as well. The same rate-adaptive wireless LAN implementation in the network simulator is used as described above.

2.7.3.1 Packet Loss Rate

The first investigated performance parameter is the packet loss rate.

The uplink loss rates are negligible in all considered scenarios. All flows go downlink via the access point and thus the arrival rate at the access point, compared to the stations' arrival rate is multiplied by the number of active flows. The transmission opportunities are however fairly distributed between all stations. This results in the access point having a much higher load than the other stations. In scenarios where the access point is loaded as such that losses are below the threshold for reasonable quality of service, the stations

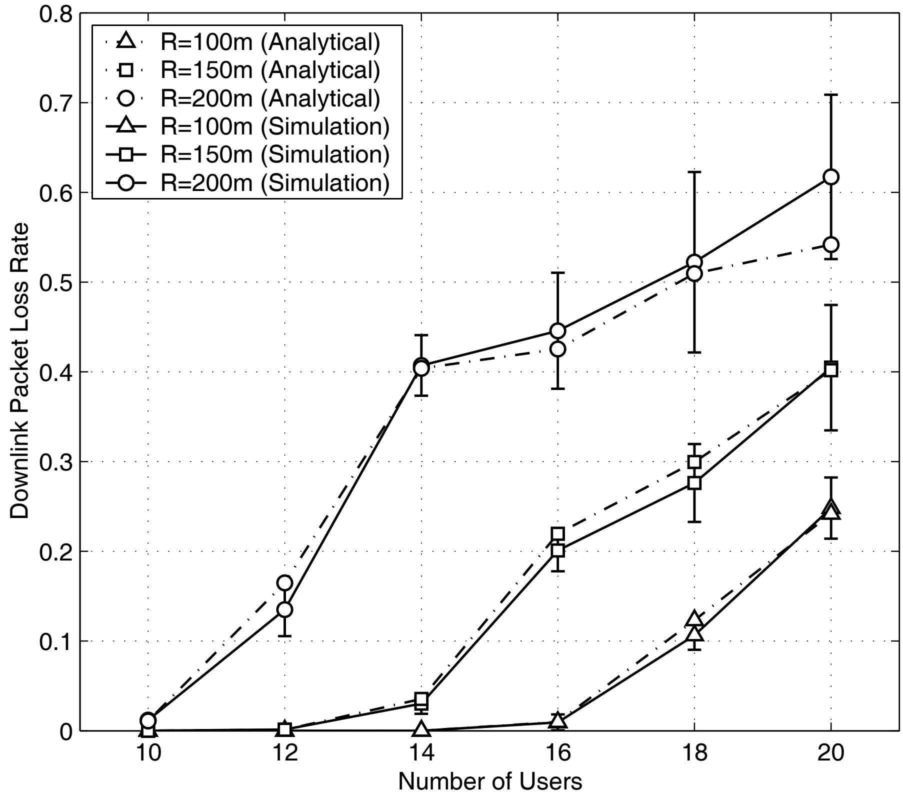


Figure 2.19: Downlink packet loss rates vs. the number of users

only have a very small load. This is why the queues in the stations are most of the time empty and queuing losses do not happen at all. The losses due to transmission errors and collisions are negligible. This will be verified by showing the low collision rate later. Hence the uplink packet loss rates are not shown here.

The mean downlink packet loss rate vs. the number of users is depicted in Fig. 2.19 and vs. the maximum distance to the access point in Fig. 2.20.

The first thing to note is again that the simulations match the calculated results well. All calculated values are within the 95% confidence intervals of the simulation results.

Further, it can be seen that the loss rate increases with the number of users and with the radius of the area in which the users are placed. As said above, the losses happen mainly at the queue in the access point (the low collision probability is shown below). Hence, it is clear that a higher number of users induce a higher loss rate: The arrival rate at the access point queue increases while the service rate stays constant.

Similar reasons can be given, why the packet loss rate increases with the radius of the area in which the users are located. The distance between access point and the user determines the modulation scheme with which the packet is sent and the error and retransmission probability. Hence, a user near to the access point can communicate faster. As the average

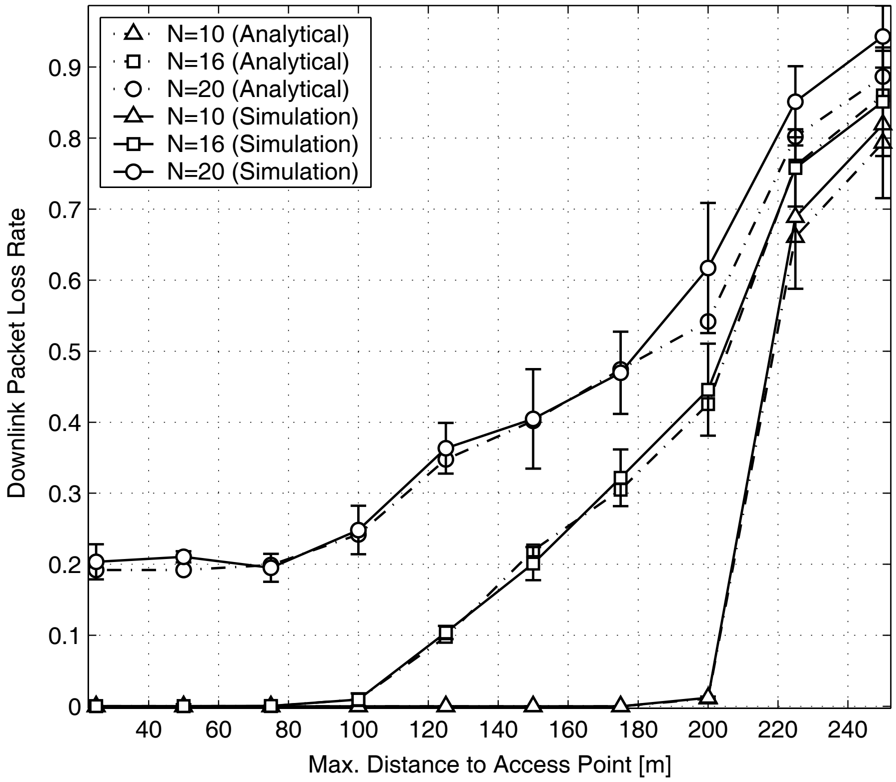


Figure 2.20: Downlink packet loss rates vs. the maximum distance to the access point

distance to the access point increases with the radius of the location area, the packet transmissions need on the average more time. This means the service rate decreases, while the arrival rate stays constant. Thus, the packet dropping rate at the queue increases.

It is assumed that using current voice codecs the voice quality would be still acceptable for packet loss rates up to 5%. Looking at Fig. 2.19 and Fig. 2.20 again, we see that up to $N = 16$ users can use the voice over WLAN service well, if they are all located within a maximum of $R = 100$ m to the access point. If some users move out of this area the loss rate increases rapidly to values, which would render the voice service useless. The loss rate behaves similarly, if more stations use the voice service: We see that $N = 20$ users will not be able to use the service, regardless of their locations.

2.7.3.2 Delay

The packet losses happen, because more packets enter the queue at the access point than leave it. The queue fills up until its capacity is reached and packets need to be dropped. The increasing number of packets waiting in the queue should be visible from the transmission delay as well. The downlink delay is measured as the time between entering the queue at the access point and being received correctly at the station. The uplink delays

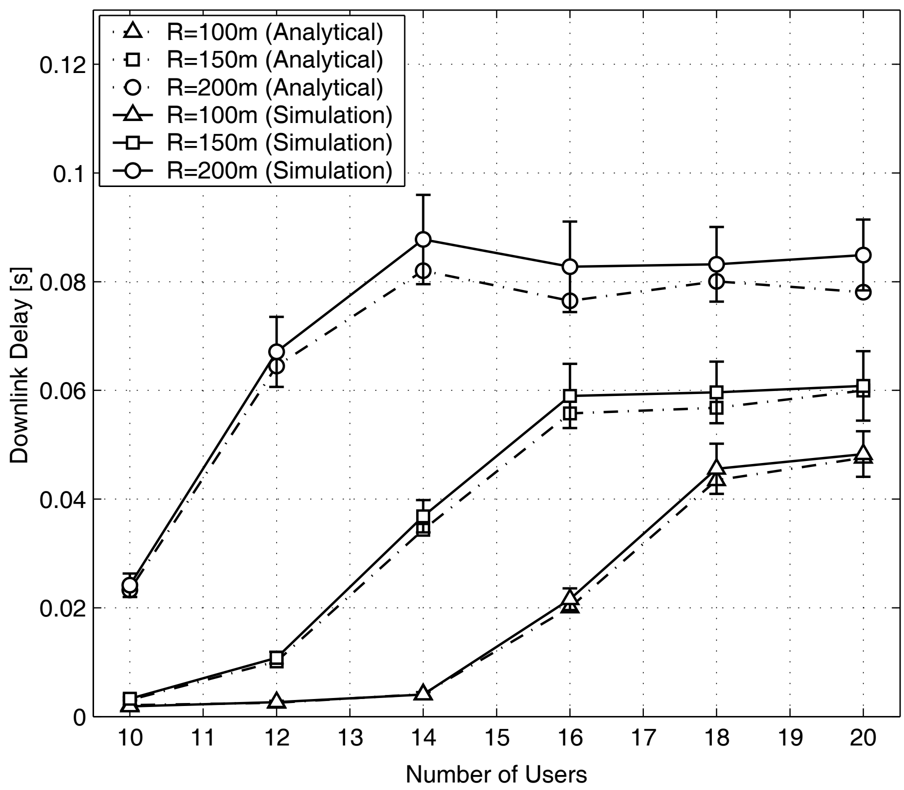


Figure 2.21: Downlink delay vs. the number of users

are all small in the considered scenarios and thus not shown here. The reason is again the significantly higher load of the access point compared to the stations' load and as the delay is dominated by the queuing delay in the access point, the stations' delay is small. The mean downlink delays are shown in Fig. 2.21 and Fig. 2.22.

The delay follows the same trend as the packet loss rate. However, the delay starts to increase already with fewer users and reaches saturation. This is because the queue fills, but does not need to drop packets with e.g. $N = 14$ users. In scenarios where a significant amount of packets need to be dropped, the queue is always full and the delay reaches a maximum, even if the number of users is increased further.

However, increasing the location area of the users does not lead to a saturation point of the delay. The increased location area does not only fill up the queue, but also decreases the service rate, which increases the delay further.

Comparing Fig. 2.19 and Fig. 2.21 we see that for scenarios with reasonable packet loss rates (again, assuming that typically voice codecs could accept loss rates of up to 5%), the delay is always below 20 ms. Hence, one could concentrate on the loss rate as performance criterion.

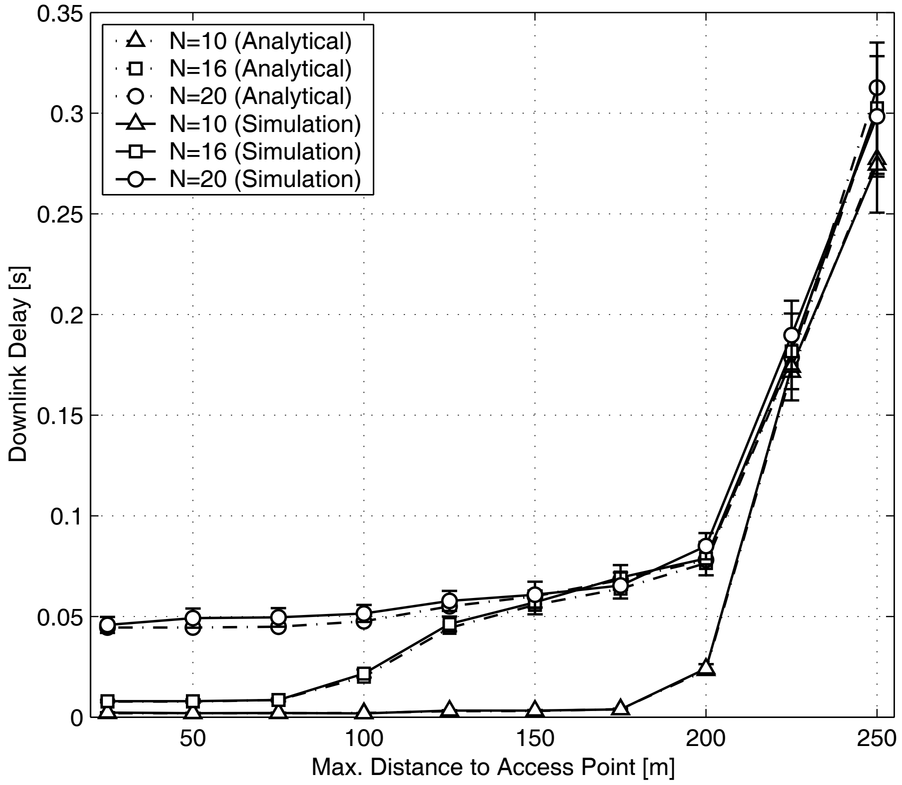


Figure 2.22: Downlink delay vs. the maximum distance to the access point

2.7.3.3 Collision Probability

It was stated above, that the main reason for downlink packet loss is queue overflow in the access point. Here, the collision probability is shown and evidence is given that collisions are negligible compared to packet drops at the queue. The collision probability for the considered scenarios is given in Fig. 2.23. We see that the collision probability increases significantly with an increasing number of users. The more users want to access the channel simultaneously, the higher the collision probability. In most of the cases the access point is saturated and thus always has a packet to send. The stations' packets could therefore collide with the access point's packets. Further, increasing the number of active users increases the probability that multiple users want to access the channel at the same time as well, leading to a collision between packets for two stations.

The collision probability increases only slightly with the radius of the location area of the users. The slight increase can be reasoned by a higher transmission rate, because of retransmissions. This is also visible from Fig. 2.24, where the channel utilization is plotted for the considered scenarios. Here, we see that the channel utilization increases with the number of users and with the radius of the location area. A higher channel utilization increases the probability that multiple stations try to access the channel simultaneously.

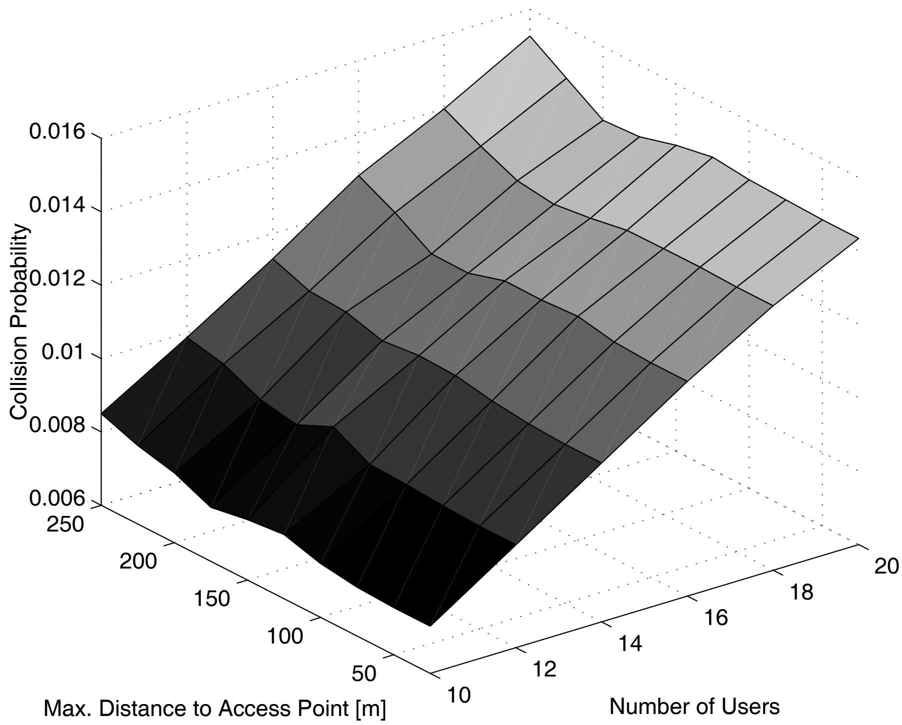


Figure 2.23: MAC layer collision probability vs. the number of attached users and the radius of the location area of these.

However, the effect that more users have on the collision probability is much more significant.

The collision probabilities are in general very low, because we consider non saturated sources. The access point is in most cases the only station that has always data to send. The other stations are fed by a Poisson source with a significnatly smaller arrival rate. The service rate could be as high as the access point's service rate, because of the DCF distributing the transmission opportunities fairly between all stations including the access point. Hence in the stations the service rate is much higher than the arrival rate. Thus, the other stations are most of the time idle. Low collision probabilities for non saturated sources were already reported in [9] and this finding is reproduced with the model presented in this section.

2.7.3.4 Idle Time Distribution

In the course of deriving the collision probability, it was assumed that MAC packets are transmitted according to a Poisson process. This is reasonable, because the stations are not saturated, i.e. they do not always have a packet to be sent. The arrival process at the stations is Poisson and the sum over all stations is then Poisson. However, the access point is for most scenarios saturated and thus the packet transmission process is not Poissonian,

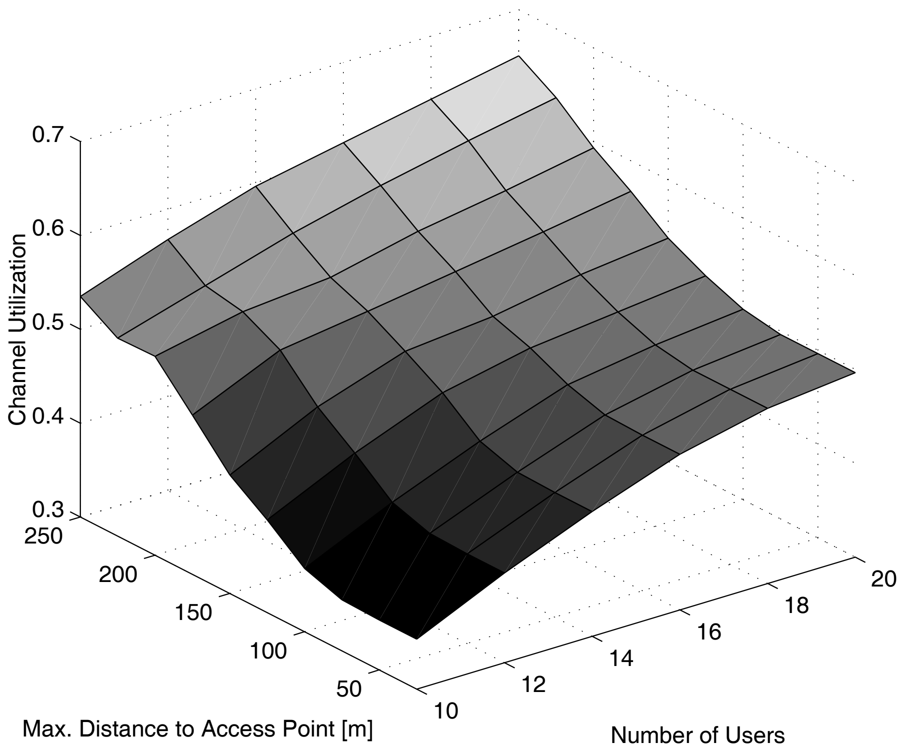


Figure 2.24: Channel utilisation vs. the number of attached users and the radius of the location area of these.

but follows the service time distribution, which we have calculated above as well. Thus the overall process is not Poissonian, but we show that it can be approximated by such a process.

The channel idle time was measured in the simulations for a specific scenario and compared to an exponential distribution with the calculated mean idle time. The cumulative distribution functions of both are depicted in Fig. 2.25. We see that the assumption was valid, as the distributions differ not significantly. Further, the loss rate and delay from the calculations were compared with the simulations and have shown a good match.

2.8 Conclusions

An analytical model of CBR flows, TCP flows and voice traffics over rate adaptive wireless LAN was developed and verified with simulations. With help of this model performance parameters of the WLAN service can easily be calculated for arbitrary user distribution. The level of detail at bit, packet and flow level distinguishes the model from related work in this research field and allows to draw interesting conclusions on the performance influencing parameters in rate adaptive WLAN.

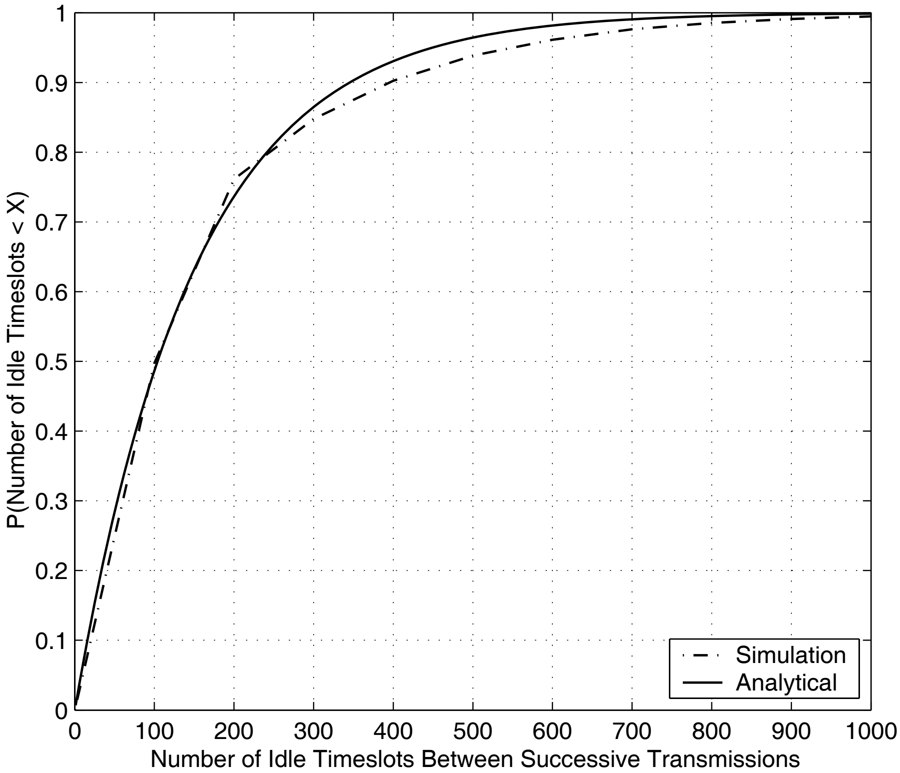


Figure 2.25: Cumulative distribution function of the simulated channel idle time compared to an exponential distribution with the calculated mean channel idle time. Channel idle time is given in time slots T_{slot}.

The most striking result is the extent to which the quality of service of one user depends on the locations of other connected users. This shows the importance of quality of service estimation for inter-system handover decisions, which should consider user locations.

One could figure out clear thresholds regarding user locations and number of users for which the quality of service would still be in reasonable bounds. This helps in designing quality of service control algorithms that consider the locations (or signal strength) of users to derive user admission or user drop decisions. This will be exploited in the next chapter about a location based quality of service control system.

Another important finding is that the quality of service is highly asymmetric: All users have nearly no packet loss and very low delay uplink. However, all users share possibly high packet loss and high delay downlink, because they all communicate via the access point, which is the bottleneck in the considered scenarios. It was shown that in scenarios with reasonable quality of service for the users, the channel utilization is around 50% only.

Chapter 3

Location Based Quality of Service Control

3.1 Introduction

In Chapter 2 it was shown that the throughput of rate adaptive WLAN as it is used today, is mainly depending on the number of users that are connected to the access point and on the location of these. The throughput of a user can be decreased by orders of magnitude only because other users have changed their locations. As the users are usually mobile, this leads to high Quality of Service (QoS) variations of all communication participants at an access point.

Many applications that are in use today require a certain minimum QoS. E.g. voice communication requires a minimum throughput and maximum delay. If these applications are used in a WLAN Hot Spot, as evaluated in Chapter 2, it is likely that a user will receive the minimum QoS at some times and will fall short of that QoS at other times. In cases of heavily loaded access points it is also possible that the minimum QoS can never be reached. The quite low and further highly variable quality of the WLAN service is a problem for using these applications in WLAN Hot Spots.

To overcome this problem, the QoS needs to be increased in general and especially be less variable. This could be achieved by introducing admission or QoS control. By QoS control it is meant here that new connections of some users might have to be rejected and even ongoing communication of a user could be dropped. This is done for the sake of many other users, which would then still be able to communicate with a given minimum QoS. To reject or drop the right users at the right points in time in order to use the given resources most efficiently is an ambitious aim and in this section different QoS control systems are proposed that fulfill this aim more or less well.

A QoS control system should consider the users' locations, as was shown in the previous section[1]. Users which are close to the access point, receive the data with high signal strength and can achieve a high throughput. These should not suffer from users which are far away from the access point. In preferring the users in proximity to the access point, one could further increase the overall performance of the WLAN. The QoS control algorithms are defined in detail below and are evaluated regarding QoS guarantees and overall throughput of the access point via simulations.

In the next section first the requirements of a QoS control system for WLAN are analyzed. Thereby the differences of QoS control in wireline and wireless communications are highlighted. In Section 3.3 the related work on QoS for WLAN is described and it is explained, why the related work is not suitable. Then two QoS control systems are defined based on user locations in Section 3.4. The first system being a straightforward approach of allowing users to communicate only, if they are within a certain admission area around the access point. The second system is more sophisticated in sizing this admission area dynamically depending on the load of the access point. These two QoS control systems are evaluated by means of simulations in Section 3.5.

3.2 Requirements on QoS Control in WLAN

QoS control for WLAN is different from QoS control in fixed networks. Fixed throughput guarantees can not be given, because of the wireless channel. If users depart from the access point, their throughput will degrade regardless of what a QoS control system could do. This is also the difference to cellular systems, where moving users would handover to other base stations. WLAN is characterized by a limited coverage and handover to another access point is usually not possible. Hence, first some requirements on a QoS control system for the WLAN scenario are defined and described.

First, one requires from a QoS controlled WLAN system that users receive a certain QoS with a certain probability. This probability is a good measure of the QoS control system's performance: The higher it is the better.

Besides a QoS communication probability, a QoS control systems should increase the offered QoS in general or the amount of users to be accepted. Basically, the given resources of the WLAN system should be used as efficiently as possible.

As said in the previous section, users in proximity to the access point require less resources and thus should receive a preferred service. These users should have a higher probability to reach their requested QoS.

Summarizing, QoS control in WLAN needs to fulfill the following requirements:

[1]The user's location is considered as the performance indicating parameter. One could of course use the received signal strength at the access point as well and in a real life implementation, this is how it should be done. According to the analytical model the location of a user is mapped to the signal strength. The location of the user is more descriptive and helps in understanding the correlations between users and WLAN performance and hence it is used here.

- Enable QoS communication with a certain probability.
- Increase the overall throughput of the WLAN.
- Minimize the QoS violations for users within a certain maximum distance to the access point.

3.3 Related Work

The need for mechanisms to support QoS communications in WLAN has been identified by the research community for a long time. The DCF is robust and needs little co-ordination between nodes, but offers a probabilistic service only. QoS cannot be guaranteed. Not even relative QoS, i.e. giving high priority stations a better QoS than low priority ones, can be provided by the basic DCF.

The IEEE realized this lack of QoS support and defines a QoS extension to DCF in its amendment IEEE 802.11e [59]. An Enhanced DCF (EDCF) is defined, which allows for differential QoS. According to the standard, a station implements separated queues for different traffic service classes. The parameters of the DCF, such as minimum and maximum contention window, are adjusted according to the priority class of the current packet being sent. By means of this configuration, prioritized channel access is realized. There has much work been done on EDCF and similar technologies. Analytically and by means of simulations, the performance of different parameter settings have been evaluated and a variety of these have been proposed in the literature.

The EDCF functionality accomplishes the aim of having relative QoS. However, looking at the requirements that were defined in Section 3.2, a different approach is needed. According to the requirements we have to differentiate users according to their locations and reject or drop users that do not use the resources efficiently. The EDCF approach is orthogonal to the needed solution in that it differentiates traffic classes. It can well be used additionally to the QoS control system that will be defined.

In [35] Deng et al. first modified the backoff parameters of DCF to give priority channel access to selected traffic flows. Qiao et al. propose backoff parameter settings to give weighted fairness to flows and maximize the network efficiency in [100]. Aad et al. propose different mechanisms to differentiate the service in [1]: Modifying the inter frame spacing, modifying the contention window parameters or allowing different packet sizes for different service classes. In [8] the Banchs et al. modify the backoff parameters of service classes dynamically, based on measured QoS parameters. Kanodia et al. propose a slightly different scheme in [68]: The channel access is also differentiated by means of different contention window parameters, but the coordination between stations is different: The stations piggyback their head-of-line packet's priority value on their transmitted MAC packets (RTS/CTS or Data/Ack packets). Hence, all stations in the neighborhood are able to generate a priority table and adopt their backoff parameters accordingly.

Vaidya et al. describe a distributed fair scheduling mechanism for WLAN in [111]. Thereby, stations are assigned a weight corresponding to their throughput service class. The backoff values are then chosen depending on the weight and additionally the size of the packet to be sent. Flows with smaller sized packets would get a lower throughput when having the same number of transmission opportunities. By considering the packet size in the backoff parameter calculation this is compensated for. In [95] Pattara et al. propose a distributed deficit round robin scheduling. According to the throughput class a node gets a service quantum rate assigned. The backoff value is decremented by this rate, meaning that the probability of a transmission opportunity increases with this rate.

Xiao et al. describe a distributed admission control in [123]. However, it still relies on a central QoS aware access point. The access point calculates a transmission budged, i.e. an additional transmission time that can be used in the next beacon period, and communicates this budget in the broadcasted beacon. The stations might add part of the budget to their currently used time share. In case of zero transmission budget, new flows are rejected. In [124] Xiao et al. extend this scheme by accepting new flows tentatively. The performance is then measured during the next beacon interval and according to the results the flow might continue or be rejected.

In [104] Shah et al. describe another centralized approach. A bandwidth manager is used to assign flows a time share to access the channel. The time share is calculated from the bandwidth requirements and the currently available bandwidth.

In [97] Pong et al. want to give service guarantees in addition to service differentiation. Hence, the authors use EDCF and identify the need to have additionally admission control. Therefore, they measure the collision probability and use the Bianchi model [15] to calculate the throughput of a certain service class with the corresponding backoff parameters. The admission control entity compares the achievable throughput to the required QoS of all flows and determines whether the new flow is to be admitted. Yang et al. use a similar admission control mechanism in [125]. It is based in a slightly different throughput model. Additionally to admission control for real time flows, best effort flows are accepted. Rate control in all nodes is used for best effort services, i.e. it is controlled that best effort service data is only admitted into the network in such an intensity that the real-time flows are not disturbed. Similarly, Zhai et al. proceed in [127]: First a channel busyness ratio is calculated. Thereby the stations measure continuously for each time slot the probability that a successful transmission is going on, an unsuccessful or the slot is idle. From that the busyness ratio is calculated. The authors note that for networks with a business factor below a threshold delay and loss is low and the network is efficiently run. In contrast, if the busyness ratio increases beyond a threshold, the network is overloaded and QoS cannot be guaranteed. Hence, admission control and rate control for best effort traffic is done to keep the busyness ratio below the threshold.

Similarly, Zhang et al. propose an admission control system in [128]: Traffic is categorized into service classes. If a new flow wants to be admitted, it is checked whether bandwidth from the best effort service class can be converted to the requested service class. The conversion is done, because different service classes have different DCF parameters according to the IEEE 802.11e amendment. If the best effort bandwidth can be converted and further enough residual bandwidth would remain in the best effort class, the flow is accepted.

Gu et al. propose in [43] another measurement based admission control algorithm. Nodes measure either the relative occupied bandwidth or the average collision probability. Based on thresholds of these measures, new flows are accepted or rejected. In [112] and [11] the Virtual MAC and Virtual Source are proposed, which are basically measurement based admission control algorithms as well. The Virtual MAC resides in parallel to the MAC implementation. It behaves like the real MAC in that it contents for channel access by means of the DCF backoff algorithm. However, packets are not really sent, but the Virtual MAC only checks whether a collision would have had happened, if the packet had been sent. This enables the Virtual MAC to estimate the channel access delay and collision probability. In the Virtual Source this is further extended by simulating a real application. Virtual packets are generated, time stamped and passed through the Virtual MAC. This enables a more detailed delay and application layer loss rate estimation. A similar approach is followed by Zhang et al. in [81]. However, the authors substitute the simulation of the Virtual MAC with a model based approach. They build an analytical model of WLAN under non-saturation conditions from which the performance parameters are calculated. The model is however based on exponentially distributed service times, which leads, as shown in the previous chapter, to wrong performance results.

While these mechanisms enable QoS differentiation and partially give also QoS guarantees for some users, the focus of this work is rather on the WLAN system itself. As stated in the requirements in section 3.2, the given resources should be used as efficiently as possible. This is not achieved with the solutions given above. User locations have a significant influence on the overall performance of the WLAN, but are not considered in any of the given solutions above. As users that are nearer to the access point require less resources, these users should be handled prioritized, as stated in the requirements. This is also not possible with any of the above mentioned QoS control systems.

Some of the related work mentioned above uses a model-based admission control, e.g. [81] or [97], which is a useful approach for this work as well. The WLAN model as described in the previous chapter, can be used to predict the QoS of users and the performance of the whole system. Based on the predicted performance users can be admitted, rejected or even dropped from the system. This approach will allow to guarantee QoS to some users on the one hand and optimize the overall performance on the other hand. Inspired by the related work as cited above, I will follow that approach in the following sections and show its benefits by means of a detailed performance analysis.

3.4 Location Based QoS Control

In the previous chapter I have shown that the throughput depends on the locations of the users. Further, in the requirements it is stated that users at certain locations should have a preferred QoS. Hence, a QoS control system for rate adaptive WLAN should consider the locations of users.

In this section two different location based QoS control mechanisms are described. While the first one is rather simple and straightforward, the second one reacts dynamically to the connected users and their locations.

3.4.1 Simple Location Based QoS Control (SLBQC)

A straight forward approach to location based QoS control would now be, to accept connections only for users within a certain maximum distance to the access point. The area in which connections are accepted is denoted as the admission area. This algorithm is referred to as Simple Location Based QoS Control (SLBQC). In this algorithm all users are accepted within a maximum distance to the access point of d_{max}. If users move further away, their communication is interrupted and the call is dropped. If users try to connect from a distance larger than d_{max}, the call is rejected right from the beginning. If these users move into the admission area, the call is established.

SLBQC requires a fixed maximum distance, which is set for the performance evaluation to $d_{max} = [60, 100, 150]$ m. Looking at Fig. 2.8 again, we see that for $d_{max} = 60$ m connections with the highest modulation scheme are accepted only. For $d_{max} = 100$ m the two highest schemes are accepted and for $d_{max} = 150$ m we block connections with the lowest modulation scheme.

3.4.2 Location and Throughput Based QoS Control (LTBQC)

A more sophisticated admission control scheme would try to set the size of the admission area dynamically, depending on the load of the access point. This is what is done in the Location and Throughput Based QoS Control (LTBQC).

Here, the signal strength of all users is continuously monitored at the access point. Frequently, the expected throughput of each user is calculated from these signal strengths. In addition the calculation is done if a new user wants to connect to the access point. If the expected throughput falls short of the defined minimum throughput, users have to be dropped. The system will select the users which are farthest away from the access point, because these users have the highest influence on the performance degradation of the whole system. Further, in the requirements in Section 3.2 it was stated, that users in proximity to the access point should get a preferred QoS. After dropping a user, the expected throughput is again calculated and more users might have to be dropped. Eventually the expected throughput will be larger than the set minimum threshold. Due to moving users, the expected throughput changes rapidly over time and it might well be that a user is again accepted at the next throughput calculation.

3.5 Performance Evaluation

The performance of the QoS control systems is evaluated by means of simulations. I will first describe the considered scenario and then present a detailed performance evaluation of the QoS control systems.

3.5.1 Scenario

A scenario is assumed, where users want to download content from a server connected to the access point with a constant data rate, e.g. a voice streaming service. This corresponds well to the investigation on CBR traffic via WLAN, that was conducted in Section 2.5. The users move in a fixed area, which could be e.g. an area where passengers wait for their planes at an airport.

Scenarios are simulated in which N users move randomly in a circular area around the access point. The location area has a radius of $R = 200$ m. The users move according to the random waypoint mobility model: a user selects a random waypoint of the considered location area and moves there with a constant speed, randomly chosen from the interval [1m/s; 3m/s]. After reaching this location, a new waypoint is chosen and the user moves there with a new constant speed.

As said above, a voice streaming application is considered. Users download the voice stream via the access point with a fixed data rate of 48 kbps. The minimum throughput per user is thus also $TP_{min} = 48$ kbps.

The same simulation environment is used as was done in the evaluation of the analytical WLAN model in the previous chapter.

For each user, the received QoS in terms of throughput is monitored. This allows to calculate the fraction of time during which a user had on average not received the minimum throughput. This metric represents the QoS violation probability, i.e. the probability with which a user will receive a QoS less than the guarantees. Further, the overall throughput at the access point is measured continuously.

3.5.2 Quality of Service Violation Probability

First, in Fig. 3.1 the QoS violation probability is shown, i.e. the fraction of time that users did not get the minimum throughput. Thereby, the number of users in the scenario is varied from $N = 5$ to $N = 50$. The number of users determines the load of the network.

In very low load scenarios (5 or 10 users), the fraction of time that the users did not receive the required throughput is very low without using any QoS control. This is because the load of the network is low and all users can be satisfied in most cases, even if some users are at large distance from the access point. By blocking users that are farther away than 150 m the fraction increases slightly. This is due to the blocking of users that could have

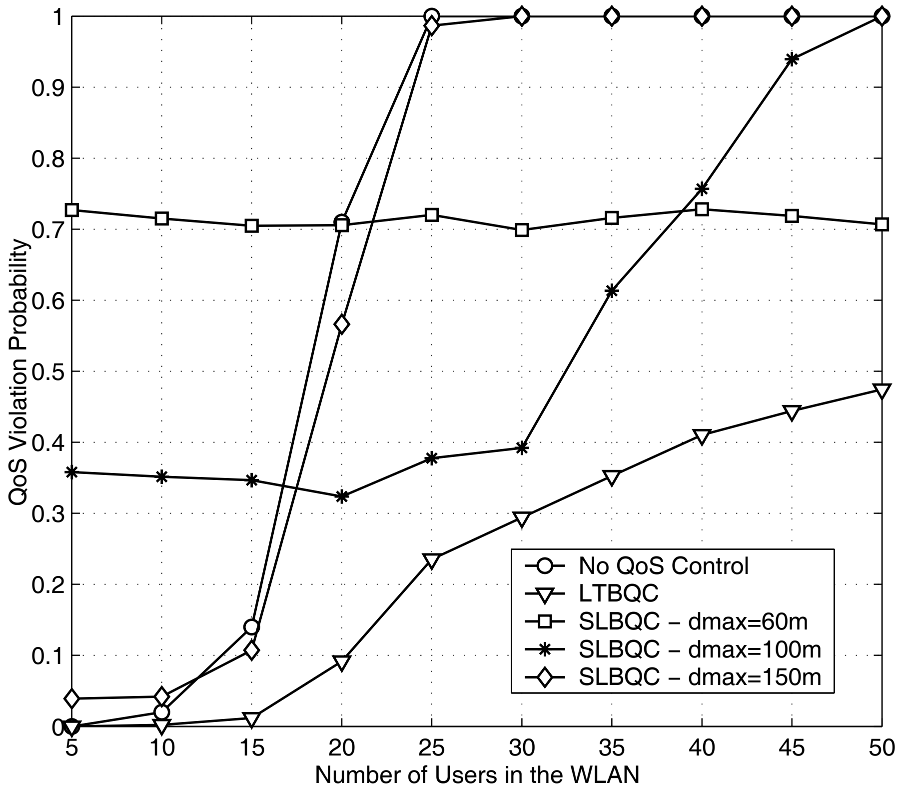


Figure 3.1: The QoS violation probability measured as the fraction of time that a user did not receive the minimum required throughput.

been allowed according to the access point load. We see this more clearly when looking at the SLBQC scheme with a maximum distance of $d_{max} = 100$ m. About one third of the time the users cannot receive the required QoS, because during this time they are at locations that are farther away than 100 m. Still, we see from the other schemes, that most of these users could have been accepted. The LTBQC performs always best, because it admits just as many users as the system capacity allows. Even for very low load scenarios, some users are blocked. E.g. for ten users, the average fraction of time that a user did not receive the required QoS is 0.2%. Even though this is negligible compared to the other schemes, where the time fraction is 2.3% for no admission control and at 4.2% for SLBQC with $d_{max} = 150$ m, we see that blocking some users is beneficial for all users on average, even for low load scenarios.

In higher load scenarios, the benefit of QoS control is even greater. We see that for 25 users in the system, the QoS can never be kept without control. However, with LTBQC in only about 20% of the time the users suffer from QoS degradations.

Comparing the different SLBQC results, we see that a certain admission area size is performing at a constant level for low loads in the network. This is because the performance is only determined by the number of users that are outside the admission area. All users inside the admission area are accepted and the QoS can be delivered. At some load, this changes, e.g. for a load of 30 users in the network, all users in the admission area of 100 m radius can be satisfied (even though only about two thirds of the users are inside this admission area). Increasing the number of users further, leads to a higher amount of users in the admission area and the throughput of all admitted users degrades. That is why the QoS violation probability increases that drastically from 30 users onwards. The same behavior is visible for an admission area of 150m radius and a threshold load of 15 users.

The LTBQC does not suffer from this effect. Here, the admission area is dynamically adjusted to fit the load of the access point. In high load scenarios the admission area would be quite small, but inside this area, all users receive a good quality of service.

Next, it is checked whether the overall throughput of the network could be increased by QoS control. The average throughput at the access point is shown in Fig. 3.2.

Further, the throughput at the access point over time is depicted for a short simulation period in Fig. 3.3.

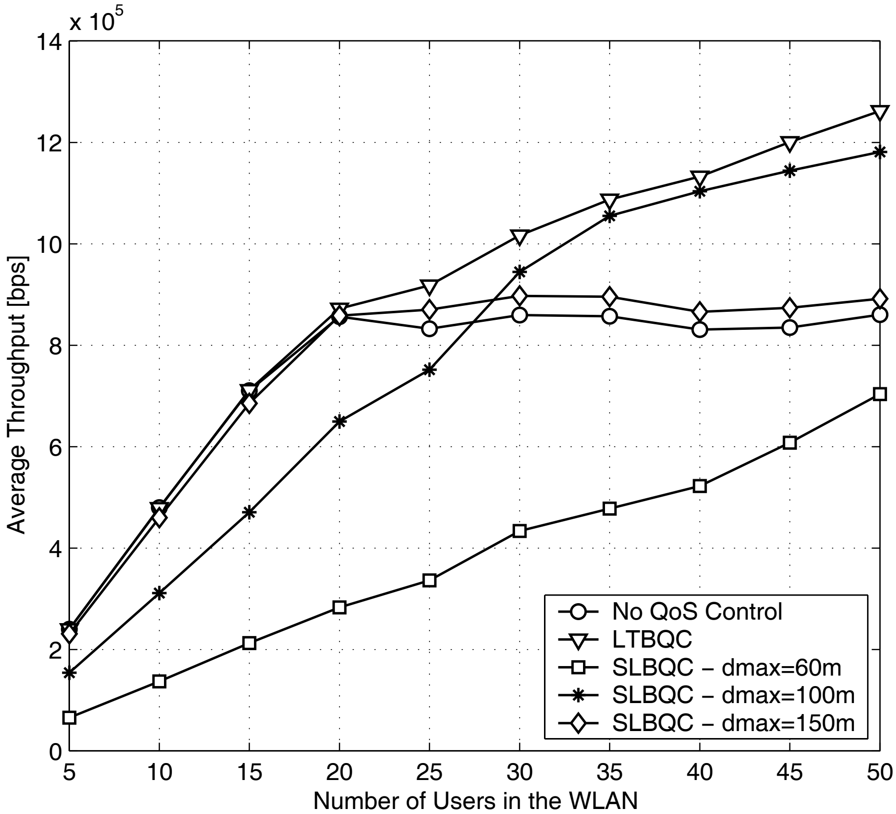


Figure 3.2: Average throughput at the access point.

From Fig. 3.2 we see that the average throughput at the access point increases nearly linearly for low load scenarios and without QoS control. This is because nearly all users can be satisfied and receive the required bit rate. Thus, the more users, the higher the overall throughput. However, this changes dramatically in scenarios of more than 20 users. The interesting point is that the average throughput still increases well from 15 to 20 users, while the QoS violation probability is for 20 users already as high as 70%. This is because without the QoS control all users are accepted and their received bit rate contributes to the overall throughput. However, the bit rate is in most cases below the required service bit rate and thus the users suffer from QoS violations.

The SLBQC mechanisms show a constantly increasing throughput with the load until a threshold is reached. This threshold is the maximum capacity of the access point for the user distribution inside the admission area. We see that for an area with radius 150m the maximum average throughput is shortly below 1 Mbps. In case of an admission area of 100m radius, the throughput is still increasing until up to 50 users in the network. Still, from the QoS violation probability, we see that for this amount of users, no one will receive an acceptable QoS, just all users will receive something.

Again, LTBQC performs best, because it adapts the admission area to the load. Hence, if many users are placed next to the access point, the admission area could be quite small, but the average throughput would be very high. We can expect the dynamic LTBQC to outperform all other schemes for all offered load values.

In Fig. 3.3 the average throughput of the access point in a high load scenario is shown over time. Interesting to see is that the throughput of the SLBQC with the smallest admission area is highly variable over time. This is because the number of users in the admission area is highly variable and all users inside the area get exactly the required bit rate. Hence, if a user enters or leaves the admission area, the throughput at the access point increases or degrades immediately by one time the required bit rate. That is why we also see only few discrete values in the curve, namely integer multiples of the required bit rate.

Now, that we have seen that LTBQC fulfils two of the requirements we set for an admission control system for WLAN, namely that the QoS communication is possible with a certain probability and that the overall throughput in the WLAN is increased, it is interesting to see whether the third requirement is fulfilled as well: Does the admission control minimize the probability of QoS violations for users close to the access point?

To answer this question, the scenarios as described above were simulated again, but this time one static user is placed next to the access point. The QoS of this user is measured in terms of received bit rate over time. Again, the QoS violation probability is evaluated by measuring the fraction of time the user did not receive the required minimum bit rate.

In Fig. 3.4 the QoS violation probability of the one user next to the access point is presented. Using no QoS control, even this user will not be able to use the voice service if the WLAN is loaded with more than 10 users. For a load of 15 users, the probability of service violation is already 10% rendering the voice service useless. The picture is only slightly better with a fixed admission area of 150m radius. The service can still not be used in cases of 15 users or more. Differently, the QoS guarantees are kept for reasonable network load values if the admission area size is reduced to 100m. In this case the minimum throughput is reached all the time if less than 30 users are in the system. However,

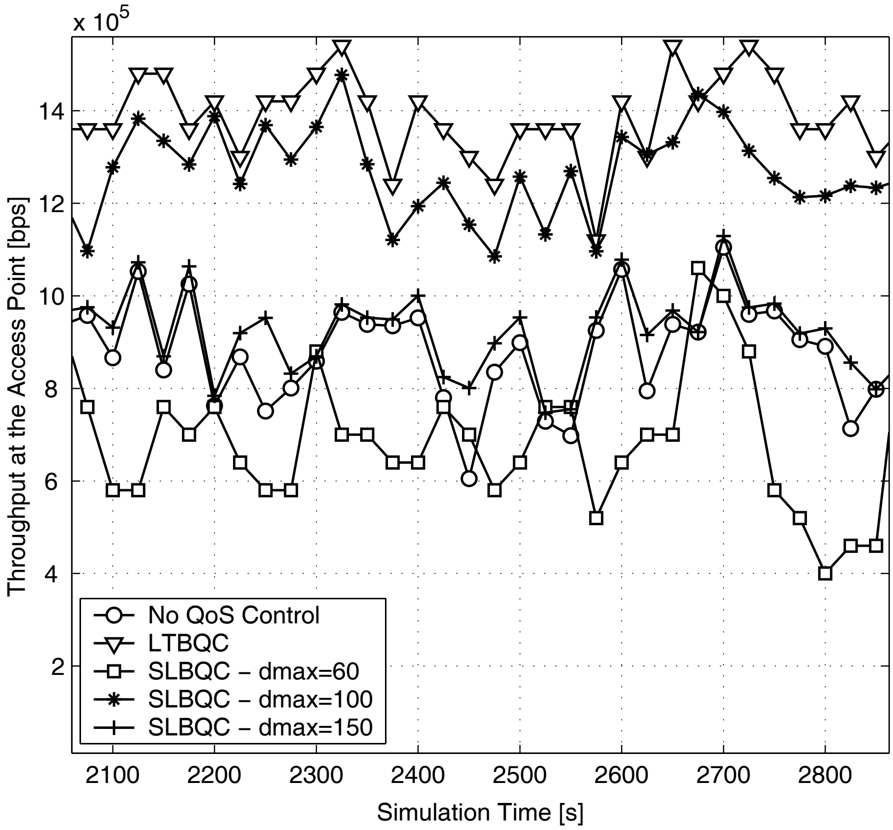


Figure 3.3: Throughput at the access point over time.

for higher load scenarios, the QoS violation probability increases rapidly. Basically, this could have been seen from Fig. 3.1 already, where the curve for this admission area size increases also rapidly for loads of 30 users and more.

The QoS violation probability is always zero if the admission area has a radius of 60m only, or if it is dynamically adjusted according to the LTBQC algorithm. Still, we have to keep in mind that for a fixed admission area size with a radius of 60m we accept only about 30% of the users in our scenarios and thus block 70% of the users right from the beginning.

3.6 Conclusion

QoS control in WLAN systems is different from what is done nowadays in cellular systems or even fixed networks. It is in general not possible to give fixed QoS guarantees to a user connected to an isolated access point. Still, some applications today require at least a minimum QoS, e.g. voice communications require a certain minimum throughput. To en-

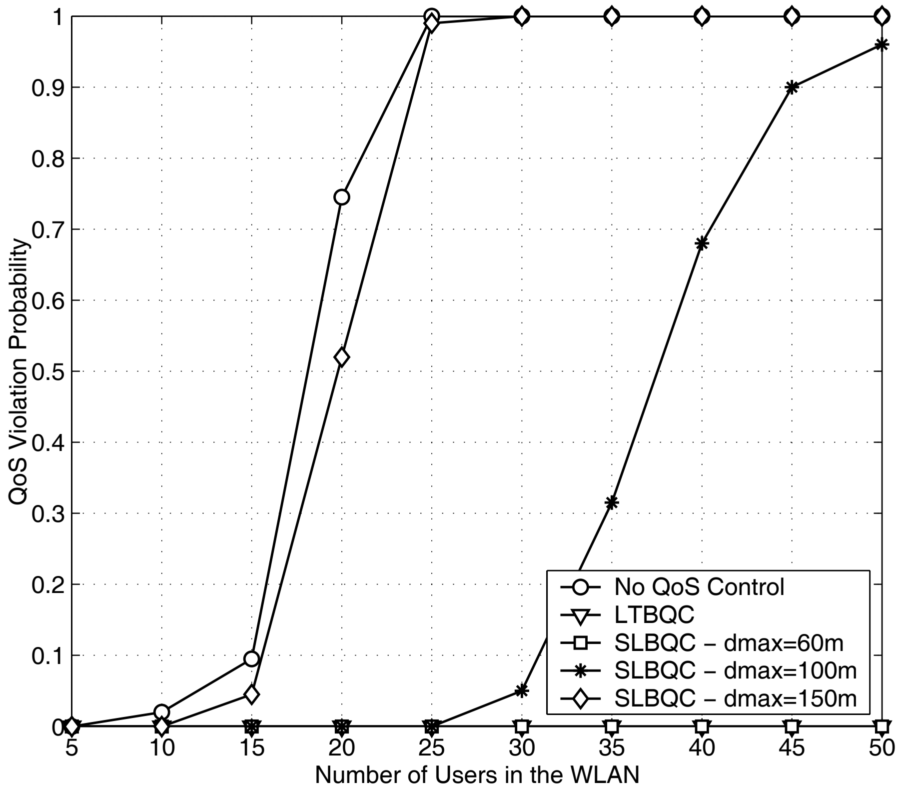


Figure 3.4: QoS violation probability of one user next to the access point.

able such applications to be used efficiently in WLAN systems, three basic requirements on an admission control system were defined.

Keeping these requirements in mind and the fact that the QoS in WLAN mainly depends in the location of users, two QoS control strategies are proposed. The SLBQC mechanism allows users to connect to the access point only if they are placed in an admission area of a certain fixed size. The LTBQC algorithm adjusts the size of the admission area dynamically depending on the load of the access point.

In a performance evaluation it was shown that the minimum QoS requirements of a voice streaming application can only be met, if the network is loaded with very few users. For reasonable network loads, the probability that the QoS guarantees are violated increases rapidly and renders the voice service useless.

Defining a fixed admission area, it was possible in some scenarios to allow using the voice streaming service. Especially a user in near proximity to the access point will have a very high probability to receive the required throughput. However, this comes to the cost of efficiency. Many users are blocked, even if the access point is not heavily loaded. The overall throughput at the access point is reduced in most scenarios.

Adjusting the admission area size dynamically, as is done with the LTBQC algorithm performs very well: The number of users that have to be blocked is minimized; the access point is loaded to its maximum, but not beyond. A user which is located close to the access point will be able to use the voice service in all considered scenarios. The overall throughput at the access point is drastically increased compared to all other investigated schemes, including the case of no QoS control.

Summarizing it can be said that QoS control in WLAN can allow the use of voice services and by means of the LTBQC algorithm the efficiency is maximized in terms of number of service users and throughput at the access point.

Chapter 4

Wireless Multi-Hop Internet Access

Deployment of WLANs in Hot Spots allows serving some users with sufficient QoS, as we have seen in the previous chapter. However, if the number of users increases, the capacity limit of the WLAN will soon be reached and either users have to be blocked or do not receive the envisioned QoS.

To increase the capacity of the networks multi-hop communication is seen as a possible solution. In wireless multi-hop networks terminals act as routers by forwarding flows for other stations. As it was shown in Chapter 2, the distance between sender and receiver influences the quality of the transmission. Low distance communication is more reliable and faster. Hence, communication via multiple small hops could improve the communication quality compared to transmitting the data directly from the source to the destination. However, the wireless medium is shared between all stations and each transmission generates interference, which means that communication via multiple hops generates more interference than communicating directly.

In this chapter, the first task is to find out what we can gain from multi-hop communication when providing Internet access at Hot Spots. Secondly, it is investigated how to gain from it. The contribution thereby is twofold: Firstly, theoretical upper bounds of the performance in wireless multi-hop access networks are calculated. These are used to see the effect of multi-hop communication in general and further to judge whether one should use rate adaptation or power control in these networks. Secondly, routing algorithms and metrics are investigated in order to see to which extend they reach the performance bounds. Two new routing metrics are then proposed, which should use the existing capacity of the network more efficiently.

I will first describe an analytical model of a wireless multi-hop access network, based on the WLAN model described in Chapter 2. The analytical model allows formulating an optimization problem to find the best routes from sources to destinations. These routes are used to calculate an upper bound of the performance that can be achieved in wireless multi-hop access networks in Section 4.2. Thereby, the effect of rate adaptation and power control is also investigated to see, whether these technologies lead to significant performance improvements. Suggestions are given, whether to use these technologies in wireless Internet access networks.

In Section 4.3 routing algorithms and metrics are investigated. Therefore, the related work is first reviewed and candidates from the variability of existing routing algorithms and metrics are selected to be investigated in greater detail. Two new routing metrics that try to use the existing network capacity most efficiently are described in Section 4.3.3.1. The selected prior art algorithms and the newly defined routing metrics are evaluated in Section 4.3.4. An analytical performance evaluation method is used to judge the efficiency of the different routes that were found using the different algorithms and metrics. Thereby voice traffic is considered only. The evaluation method is then verified by means of simulations. In the simulations the performance of the routes under TCP traffic is shown as well.

4.1 A Model of the Multi-Hop Access Scenario

The first task of multi-hop communication investigations is to describe the scenario and assumptions in form of a network model.

The wireless multi-hop Internet access scenario has many similarities with pure ad-hoc networks as being researched for years now: Nodes can act as source, destination or router simultaneously. Flows are routed through the network from source to destination, possibly via multiple intermediate nodes. The main difference to pure ad-hoc networks is that in the multi-hop Internet access scenario a set of access points exist, which are not mobile. The access points start or terminate all flows, because all users of this network want to communicate with the Internet via any of these access points.

Usually, wireless users are mobile. Moving users require that a routing protocol cares about mobility management and information exchange between nodes to get node and link status information. Routing protocols for wireless multi-hop networks are a separated research field, which is not considered in this work. This field is orthogonal to the capacity analysis that I will describe here. Routing algorithms are investigated and it is shown how many flows can be routed in a certain scenario, considering a perfect routing protocol. Any routing protocol that delivers the required information to the routing algorithm could be used then. It is the goal to clearly distinguish the routing algorithm's performance and the capacity of the network from the routing protocol's performance. Therefore, all influence of the routing protocol is neglected and only the routes that certain routing algorithms and metrics would find in a wireless multi-hop access scenario are evaluated.

In the model N users are randomly distributed in a circular area with a radius of r, according to a spatial uniform distribution. Either one or three access points (APs) act as gateways to the Internet. In case of one AP, it is located in the centre of the area. In case of three APs, they are located at the corners of a triangle having a distance of $\frac{5}{4}r$ between each others.

Two different services in the network, namely Voice over IP (VoIP) and TCP are evaluated. In the first case K users want to use the VoIP service. The communication peer is assumed to be connected to the Internet and hence all users communicate via the access point. Since the VoIP service is bi-directional, $F = 2K$ flows have to be routed, where K flows start and K flows are terminated at the access point. The communicating users are selected randomly from all users. I denote by $s(f)$ the source and by $d(f)$ the destination node of flow f. A VoIP source generates packets with a constant packet rate and packet size. The packet rate is defined as $R_f = 50$ packets/s, the packet size is 72 Bytes including IP, UDP and RTP headers. The maximum packet loss rate is 5%.

In the case of TCP service, users are assumed to be downloading data from the Internet. Again, K users are actively downloading data using TCP. All users download from the Internet and thus in the network the access point is the source and the node the destination of the flow within the access network. The TCP bit rate is not fixed. TCP adapts its sending rate according to the seen round-trip time and the loss rate and thus the flows are called elastic flows. All senders use a TCP NewReno implementation.

The access points and nodes use WLAN according to IEEE 802.11x as communication means. Rate adaptation is used with M different modulation schemes. All of these schemes have different transmission rates and different error properties for a given received signal strength. For the numerical analysis, the four modulation schemes of IEEE 802.11b with 1 Mbps, 2 Mbps, 5.5 Mbps and 11 Mbps are assumed. Additionally to what I have assumed in the previous chapters, a node transmits with one of P different power levels.

In Chapter 2 an analytical model of rate adaptive WLAN was described and the MAC layer throughput in dependence of the received signal strength and the modulation scheme was derived. The modulation scheme that maximizes the MAC layer throughput is here always used to transmit the packet. Together with the propagation model and the locations of the users, for each two users i and j the modulation scheme $mmax^p_{i,j}$ can be determined, that is used when transmitting with transmission power level p. Further, in Chapter 2 the packet transmission rate was calculated, which depends on the received signal strength, the packet size and the modulation scheme. The possible packet transmission rate between nodes i and j using modulation scheme m and power level p is denoted as $R^{m,p}_{i,j}$. Similarly $D^{m,p}_{i,j}$ denotes the transmission delay of a packet from node i to node j with modulation scheme m and power level p.

If the received signal strength is above a signal detection threshold, the nodes can hear each other, even though they might not be able to communicate. If nodes hear each other they also disturb each others communication and this is denoted by the interference indicator $a^p_{i,j}$ as

$$a^p_{i,j} = \begin{cases} 1 & \text{if nodes } i \text{ and } j \text{ disturb each others transmission at power level } p \\ 0 & \text{else} \end{cases} \tag{4.1}$$

In IEEE 802.11 the DCF controls the medium access to ensure that nodes that can hear each other do not transmit at the same time. Communication is always bi-directional, because acknowledgements are transmitted from the receiver to the sender. Thus, a link interference indicator, that describes if communication via link (i,j) disturbs the communication on link (k,l), can be calculated as follows:

$$b^p_{(i,j),(k,l)} = \min\left(a^p_{i,k} + a^p_{j,k} + a^p_{i,l} + a^p_{j,l}, 1\right) \tag{4.2}$$

The link interference indicator is limited to a maximum of one, because it should only be indicated, if any node of a first link is disturbed by any node of the second link. The amount of disturbance between links is not relevant for the following considerations.

Further, I define a route indicator $r^{f,m,p}_{i,j}$ as

$$r^{f,m,p}_{i,j} = \begin{cases} 1 & \text{if flow } f \text{ is routed with modulation scheme } m \text{ and transmission} \\ & \quad \text{power } p \text{ between nodes } i \text{ and } j \\ 0 & \text{else} \end{cases} \tag{4.3}$$

In case of constant bit rate flows, where a flow has the bit rate R_f, the load of link (i, j), considering all interfering links and flows routed via these links, is calculated as:

$$L_{i,j} = \sum_{k=1}^{N}\sum_{l=1}^{N}\sum_{p=1}^{P}\sum_{m=1}^{M}\sum_{f=1}^{F} r^{f,m,p}_{k,l} \cdot \frac{R_f}{R^{m,p}_{k,l}} \cdot b^p_{(k,l),(i,j)} \tag{4.4}$$

4.2 Theoretical Performance Limits

Power control and rate adaptation are two mechanisms of wireless communication techniques which are successfully deployed today. Rate adaptation uses the existing bandwidth more efficiently, by choosing a modulation scheme that maximizes the throughput for a given signal to noise ratio. Using rate adaptation in multi-hop networks could enhance its performance similarly. As it is already implemented in WLAN cards and access points today, it can be used without additional costs.

Using power control, a station adopts the transmission power. In individual power control every station individually sets the transmission power to the lowest possible value that still enables communication to the intended receiver. Individual power control in WLAN environments according to the IEEE 802.11x standards is problematic. The IEEE 802.11 channel access mechanisms are not designed for power control. Work arounds and solutions to the problem have been proposed, but all of these have some drawbacks and performance issues. Still, it was shown in [5] and [91] that due to the spatial reuse of the transmission frequency, the capacity of the network can be increased.

It should be shown here, to which extend one could theoretically gain from power control and rate adaptation in wireless multi-hop Internet access networks. It is important to know these performance limits, to judge whether one should use these technologies in the considered networks at all. As said above, rate adaptation can be used without additional cost in general, but the routing protocol might need to propagate information about used modulation schemes between nodes, leading to an increased overhead. The use of individual power control is problematic in WLAN anyway, as described above. Hence, deploying these technologies results in an increased effort or cost and by calculating performance limits, it can be judged, whether it is worth the effort.

Using global knowledge, routes from mobile stations to the gateway are calculated such, that a certain throughput per flow can be guaranteed. The number of flows that can be routed with these guarantees serves as a theoretical upper performance limit. It is an upper limit, because the route calculation will find a route with these guarantees if one exists. It is a theoretical result, because the routes are calculated using global knowledge, which is usually not available to the routing entity. However, comparing the performance of these routes using power control and/or different modulation schemes lets us see to which extend one can gain from these mechanisms.

The involved technologies are first described in the next section in greater detail. Then, an optimization problem is formulated in Section 4.2.2, which is solved to calculate the optimum routes. The performance gains of using different technologies are presented and evaluated afterwards.

4.2.1 Power Control and Rate Adaptation in Wireless LANs

4.2.1.1 Power Control

Power control in ad-hoc networks leads to mainly two benefits. On the one hand the energy consumption of the terminals will be reduced, if they transmit with less power while still preserving connectivity. On the other hand the capacity of the network could be increased because of increased spatial reuse of the medium: Stations transmitting with less power generate less interference and possibly enable more stations to transmit simultaneously. Here, I am concerned with the latter benefit only, i.e. I want to investigate the possible effect that power control has on the network capacity.

In general power control mechanisms in ad-hoc networks can be divided into two categories, namely common power control and individual power control. In common power control, all stations of a network agree on the same transmission power level. Usually a power level is selected that assures a certain degree of connectivity between nodes that optimizes some performance parameters. In [45] and [91] it was shown that common power control could increase the throughput in ad-hoc networks.

In individual power control, each station might use a different transmission power level. The transmission power is usually selected as such that the next hop can be reached with a certain error probability or that the station has a certain number of neighbors. Its use in WLAN stations which use the IEEE 802.11 channel access mechanisms is however problematic. The medium access is based on the listen before talk principle, which means that stations have to sense the channel idle for certain periods of time until they can be reasonable sure that no other station is transmitting at the same time. This mechanism is based on symmetric communication, which means that a station is assumed to be able to hear other stations which are disturbed by its own transmission. This assumption is not necessarily valid, if stations transmit with different power levels. Hence, collisions could occur, because some stations might not be aware of ongoing transmissions.

In [66] Jung et al. describe extensions to the WLAN MAC layer that enables individual power control. Thereby the handshake messages, which are exchanged before the data, are transmitted with the maximum power level. This enables stations that overhear these reservation messages to reserve the following medium time for the transmitting station. The data packet is then transmitted with lower power. During the data transmission the power is periodically increased shortly to signal the ongoing transmission to stations that were not able to receive the handshake messages correctly. The authors show that the energy consumption can be reduced significantly, while the stations do not suffer from throughput degradations or an increased number of colliding data transmissions. However, spatial reuse of the medium is not possible with this kind of mechanism. There are other publications that use similar mechanisms, e.g. [5], [42], [99]. All of these have in common that spatial reuse is not possible.

In [90] the Monks et al. follow a different approach by introducing busy tones. Busy tones are sent on a different frequency and indicate to other stations ongoing transmissions. While the authors show for this mechanism the possibility to reduce energy consumption and additionally gain from spatial reuse of the medium, the costs are additional frequencies and non standard conform hardware.

Summarizing it can be said that using individual power control in ad-hoc networks to gain from spatial reuse of the channel is possible with additional costs. It will be investigated here, whether individual power control in multi-hop access network scenarios is worth the effort.

4.2.1.2 Rate Adaptation

Rate adaptive WLAN stations select for unicast data transmissions a modulation scheme that maximizes the throughput for a given signal to noise ratio. The higher the received signal strength is, the higher can be the data rate of the modulation scheme that can be used while still keeping a certain bit or packet error probability. Rate adaptation is widely used in WLAN stations today. In [67] Kamerman et al. describe the auto rate fall-back mechanism that follows a trial and error approach to select the optimum modulation scheme. The data transmission is first tried with a high bit rate modulation scheme and reduced in cases of transmission errors. In [52] Holland et al. describe a rate adaptation mechanism that measures the signal strength of the handshake messages, which are sent with the low bit rate modulation scheme before the data is transmitted. The signal strength is then used to select the modulation scheme that provides a certain maximum error probability.

Even though rate adaptation is possible to be used in general without additional cost, as it is implemented in WLAN cards today, there are issues which need to be considered, when using it in multi-hop communications. The routing algorithms can gain from using rate adaptation, by finding more suitable routes using possibly higher bit rate modulation schemes. However, the information about the possible modulation schemes on the links needs to be transported to the entity that calculated the route. This is the task of the routing protocol. Even though the routing protocols' performance is not explicitly considered in this work, it is worth noting that the protocols need to gather and transport more information. This increases the overhead and resource consumption and thus leads to some costs as well.

4.2.2 Optimization Problem

In this section an optimization problem is described that helps in finding routes for the flows using global knowledge of the network. A practically relevant routing algorithm would usually not have all this information available and thus might not find these routes. Thus the result is more of a theoretical nature, but still provides useful insights.

The performance is measured by whether a routing for a certain scenario was found that fulfils the QoS requirements of all flows. A theoretical upper limit of the performance is then calculated by means of the optimization. Doing so with and without certain technologies enables to judge a maximum performance gain that one could expect from using the corresponding technology.

With the definitions from Section 4.1 the optimization problem is formulated as follows:

Minimize

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{p=1}^{P}\sum_{m=1}^{M}\sum_{f=1}^{F} r_{i,j}^{f,m,p} \cdot \frac{1}{R_{i,j}^{m,p}} \tag{4.5}$$

subject to

$$\forall f: \ \sum_{i=1}^{N}\sum_{p=1}^{P}\sum_{m=1}^{M} r_{s(f),i}^{f,m,p} = 1 \tag{4.6}$$

$$\forall f: \ \sum_{i=1}^{N}\sum_{p=1}^{P}\sum_{m=1}^{M} r_{i,d(f)}^{f,m,p} = 1 \tag{4.7}$$

$$\forall (f,i), \text{with } i \neq s \wedge i \neq d: \ \sum_{j=1}^{N}\sum_{p=1}^{P}\sum_{m=1}^{M} r_{i,j}^{f,m,p} = \sum_{j=1}^{N}\sum_{p=1}^{P}\sum_{m=1}^{M} r_{j,i}^{f,m,p} \tag{4.8}$$

$$\forall (i,j): \ L_{i,j} \leq 1 \tag{4.9}$$

$$\forall(i,j,m,p), \text{with } m > mmax^p_{i,j}: \quad r^{f,m,p}_{i,j} = 0 \tag{4.10}$$

In the objective function (Eqn. 4.5) all transmission delays of all routes are summed up. By minimizing this objective function, routes are selected, which have a shorter transmission delay, if that is possible. The objective function is not important for the performance metric: The performance is measured by identifying whether a routing can be found that fulfills the QoS requirements of all flows. This is determined by the contraints only and hence the objective function enables to select the best routing from multiple possible routings, but the performance is measured only whether at least one feasible routing was found.

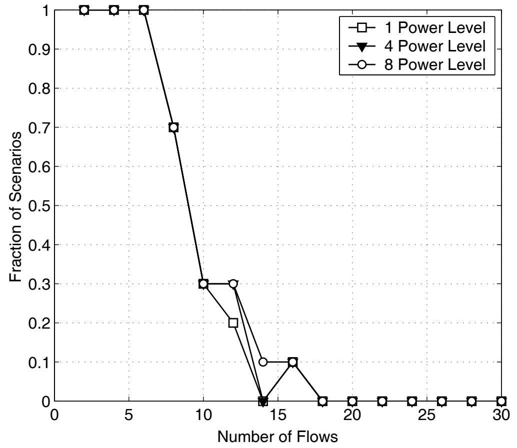
In Eqn. 4.6 it is ensured that each node that is a source node for a certain flow has exactly one route of this flow leaving the node. Similarly in Eqn. 4.7 it is ensured that for a destination node of a certain flow exactly one route of that flow enters the node. Eqn. 4.8 considers all nodes that are neither source nor destination and restricts the solutions that for each of these nodes the net flow is zero, which means that the same number of routes enter and leave the node for a specific flow.

The physical constraints from the wireless medium are put into the last two constraints. In Eqn. 4.9 the load of each link is limited to a value smaller or equal to one. The load is defined in Eqn. 4.4 and considers all routes in the interference range of the link. Another restriction from the wireless medium is that some modulation schemes are not possible between specific nodes. It is described above already, that each link has a maximum possible modulation scheme assigned, which depends on the received signal strength. By Eqn. 4.10 it is ensured that routes are used only with the maximum possible modulation scheme.

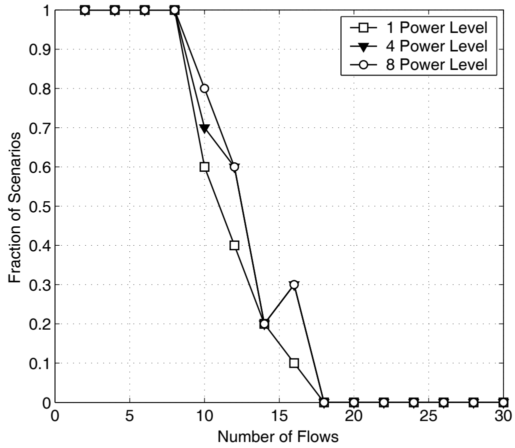
The Mixed-Integer Linear Program (MIP) as described above can be solved for reasonable sized networks with the commercial solver CPLEX [61], which employs the Branch-and-Bound method (see [76] for details). The result is a complete routing of all flows under the constraints given above. If several solutions are possible the routing with the least transmission delay would be chosen. However, it is possible that no solution exists. This could be because too many flows need to be routed via a specific interference region and links in that region would have a link load larger than one. Another reason could be that a node is located too far from any other node and communication is simply not possible. If the optimization problem does not find a feasible solution, the QoS requirements of flows cannot be guaranteed.

4.2.3 Performance Evaluation

Several multi-hop Internet access scenarios are evaluated regarding the possibility to find a feasible routing. It is assumed that only one access point is present in the center of a circular area with a radius of $R = 200m$. Different other parameters of the scenario are varied: The node density is varied by selecting the number of nodes as



(a) $N = 20$



(b) $N = 40$

Figure 4.1: Fraction of scenarios in which a feasible routing was found for one modulation scheme and different power levels

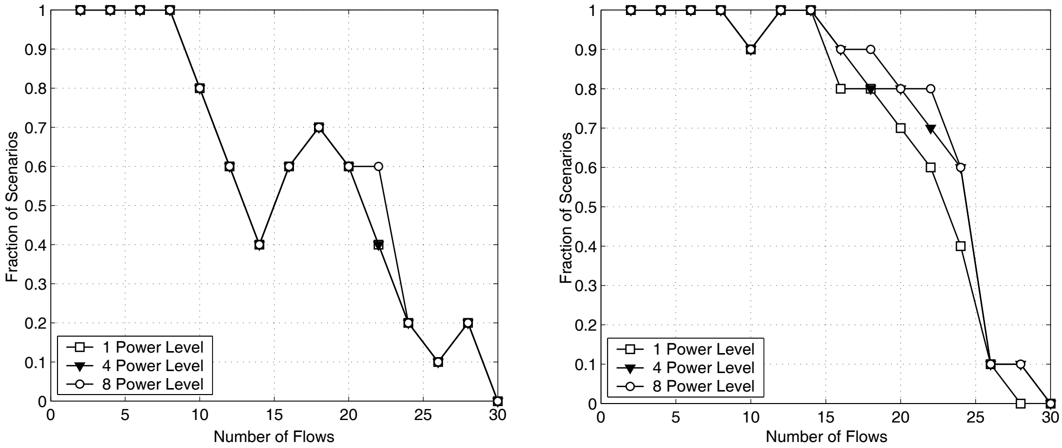
$N = 20$ or $N = 40$. The offered load to the networks is varied by selecting the number of flows in the network between $F = 2..30$. Further, the scenarios with power control are evaluated, where the station can select one out of eight different power levels ($P = \{0.1, 0.5, 1, 5, 20, 50, 100\}$mW), with power control of four different power levels ($P = \{0.5, 5, 20, 100\}$mW) and without power control, where every packet is sent with a transmission power of $P = 100$ mW. The rate adaptation is evaluated by allowing either four modulation schemes with $\{1, 2, 5.5, 11\}$Mbps, only one modulation scheme of 1 Mbps or only one modulation scheme of 5.5 Mbps.

For each parameter combination ten random scenarios are generated. Thereby, the users are located randomly across the area according to a uniform distribution. The nodes that use voice communication are selected randomly from all mobile stations. The corresponding node is assumed to be located in the internet and hence all routes end at the access point. To judge the performance of power control and rate adaptation, each of these scenarios is optimized and the fraction of scenarios per parameter combination in which a feasible routing was found is counted. Further mean link utilization in the network or at the access point is calculated.

4.2.3.1 Power Control Performance

First, the effect of power control is evaluated, if rate adaptation is not used. Therefore, the nodes use only one modulation scheme with 1 Mbps. Fig. 4.1 shows the fraction of scenarios where a feasible routing was found for $N = 20$ and $N = 40$ nodes. The number of flows was increased from 2 to 30.

The first thing to note is that the number of used power levels has hardly any effect on the performance. In the scenarios with $N = 20$ nodes, routes are found for all scenarios with



(a) $N = 20$ (b) $N = 40$

Figure 4.2: Fraction of scenarios in which a routing was found for all modulation scheme and different power levels

up to six flows. In higher loaded networks, a routing was found for only a small fraction of scenarios. With one power level only, it was not possible to route 14 flows or more. With four or eight power levels eventually a routing is found for some higher load scenarios, but the gain is not significant.

The overall performance is increased with an increasing number of nodes. For $N = 40$ nodes, a feasible routing is found for roughly two more flows in the networks. The small performance gain for eight power levels is due to the higher node density. Routes are found with small distance hops, where a reduced transmission power leads to less interference. However, the relative performance between the different power level configurations is not significant again.

The non significant performance gain of power control in the above mentioned scenarios could be, because only one modulation scheme with 1 Mbps is used. To exclude that reason, I now take a look at the effect of power control in rate adaptive scenarios. Therefore, the nodes were allowed to use any of the four modulation schemes. The results are shown in Fig. 4.2.

Basically, the curves follow the same trend as in the previous scenarios in that the difference between different numbers of power levels is hardly visible. For $N = 40$ nodes power control could lead to small performance benefits. Hence, there must be a different reason for the non-significant influence of power control on the number of flows that can be routed. With power control the interference can be reduced and an increased spatial reuse of the wireless channel should be possible. However, in the considered multi-hop Internet access scenario it is not possible to benefit from spatial reuse.

To explain this the mean link utilization of the network is calculated. The average link load (as defined in Eqn. 4.4) is calculated over all links, where communication is possible. One would expect that the utilization is lower when using power control, because sending

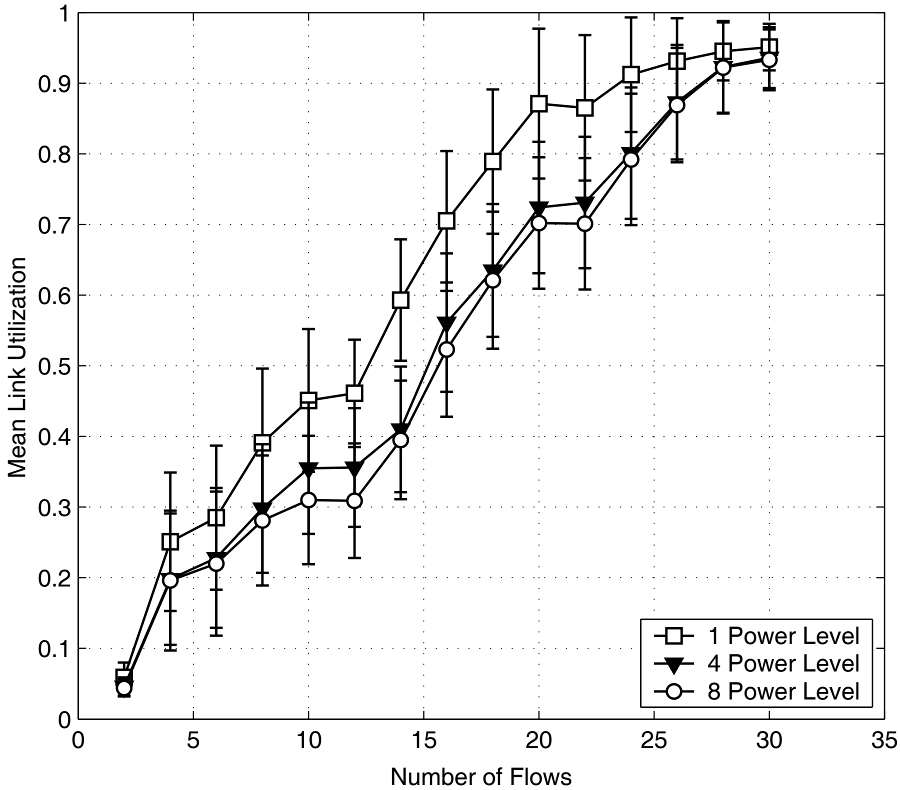


Figure 4.3: Mean link utilization for different power control configurations

with lower transmission power interferes less links. The result is depicted in Fig. 4.3: The mean link utilization in the case of $N = 40$ nodes using all modulation schemes is shown. The three curves show the mean link utilization for different numbers of power control level and the confidence intervals (95%) of these. We see that the differences are not that high; even if the mean link utilizations over all scenarios is lower with power control, the confidence intervals overlap in most of the cases.

The small differences in the mean link load mean on the one hand, that spatial reuse of the wireless channel is not that efficient in the considered multi-hop Internet access scenario. On the other hand, even these small differences do not increase the number of scenarios without overload situations significantly. This is because of the access point being a common bottle neck. Even though the link load in the whole network is decreased slightly, the load at the access point is not. All flows terminate at the access point and hence must have a link to the access point. Regardless of the power being used, the wireless channel at the access point is used by all flows. This can be seen from Fig. 4.4, where the mean link utilization is shown of all links of the access point only. The use of power control makes no significant difference in the link utilization at the access point. That is, why there is no significant difference in the number of routed flows: The access point is the common bottle neck and using power control does not help in removing this.

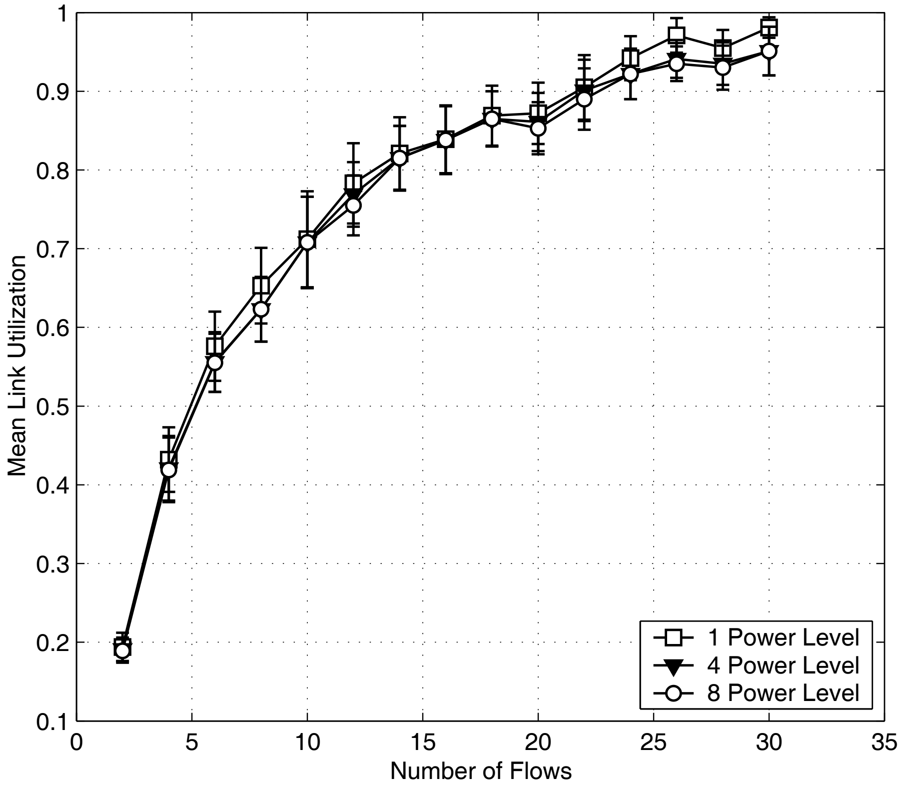


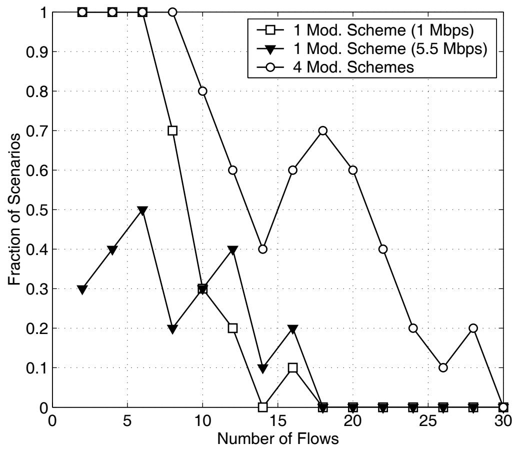
Figure 4.4: Mean link utilization at the access point for different power control configurations

4.2.3.2 Rate Adaptation Performance

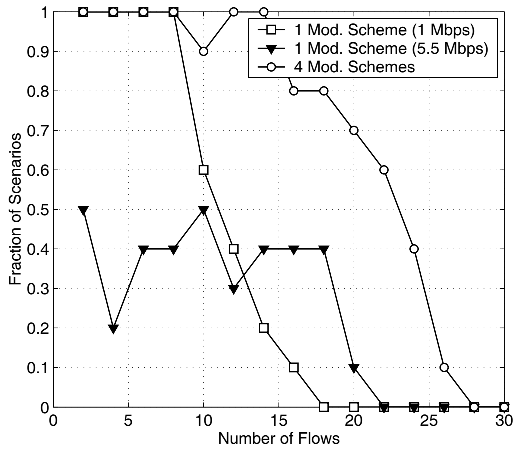
We have seen that we cannot increase the performance significantly by using power control. By reducing the interference on a spatial basis, the bottle neck at the access point cannot be removed. Rate adaptation reduces the interference on a temporal basis: Packets can be sent with higher modulation schemes and consume less time on the channel. Hence, the channel can be reused by other stations more quickly. From this interference reduction, the access point might benefit as well.

In this section the results of the evaluation of rate adaptation in multi-hop Internet access scenarios is shown. Some of the results are already included in Fig. 4.1 and Fig. 4.2, where the power control results with and without rate adaptation are depicted. In order to see the effect of rate adaptation more clearly these figures are summarized in Fig. 4.5 and additionally the graph of using only the 5.5 Mbps modulation scheme is added.

We see from Fig. 4.5 in the 20 and 40 node scenarios that for a low load of the network a correct routing could be found by either using all modulation scheme or using only the lowest modulation scheme. However, if the nodes use only the 5.5 Mbps scheme a correct routing is found only in some scenarios. This is because the network is not connected. There are nodes that are at large distance to all other nodes. These nodes can be reached



(a) $N = 20$



(b) $N = 40$

Figure 4.5: Fraction of scenarios in which a routing was found for different rate adaptation configurations

with the low modulation scheme, but not with a higher one. Hence, we see that the low modulation schemes are a must, when it comes to connect low node density networks.

Further, we see that for an increasing load on the network, the higher modulation schemes pay off. If in the 40 node scenarios in Fig. 4.5(b) more than 18 flows should be routed, it is not possible in any scenario to find a feasible routing with only the lowest modulation scheme. This is because a transmission with the lowest modulation scheme blocks the shared channel at each packet transmission for a long time; time in which another station could have used the channel. It can be said that the low rate modulation scheme produces a high temporal interference. Using only the one higher bit rate scheme, a routing in some higher load cases can be found. Using all modulation schemes a feasible routing is found, also in one scenario with 26 routed flows.

The strong effect that rate adaptation has can be reasoned by the significantly reduced transmission times if higher modulation schemes are used as it was already guessed above. In this way the interference is reduced as it is done using power control, but on a temporal basis in contrary to the spatial basis. This can be seen from Fig. 4.6, where the mean link load in the network is shown. Similarly to the power control investigations, we see a small difference in the mean link load in the network when using rate adaptation. Even though the difference is larger than for the power control investigations, there must be another reason for the performance gain of rate adaptation that we have seen above. The reason is given in Fig. 4.7, where the mean link utilization at the access point only is shown. Here, the difference is not less than in the whole network, but even larger. This is because small hops with high bit rate modulation schemes are preferred in proximity to the access point and these hops block the channel for less time, because of the higher transmission speed. We see the reduction of temporal interference in this figure.

In contrast to the interference reduction on a spatial basis, the temporal reduction has a

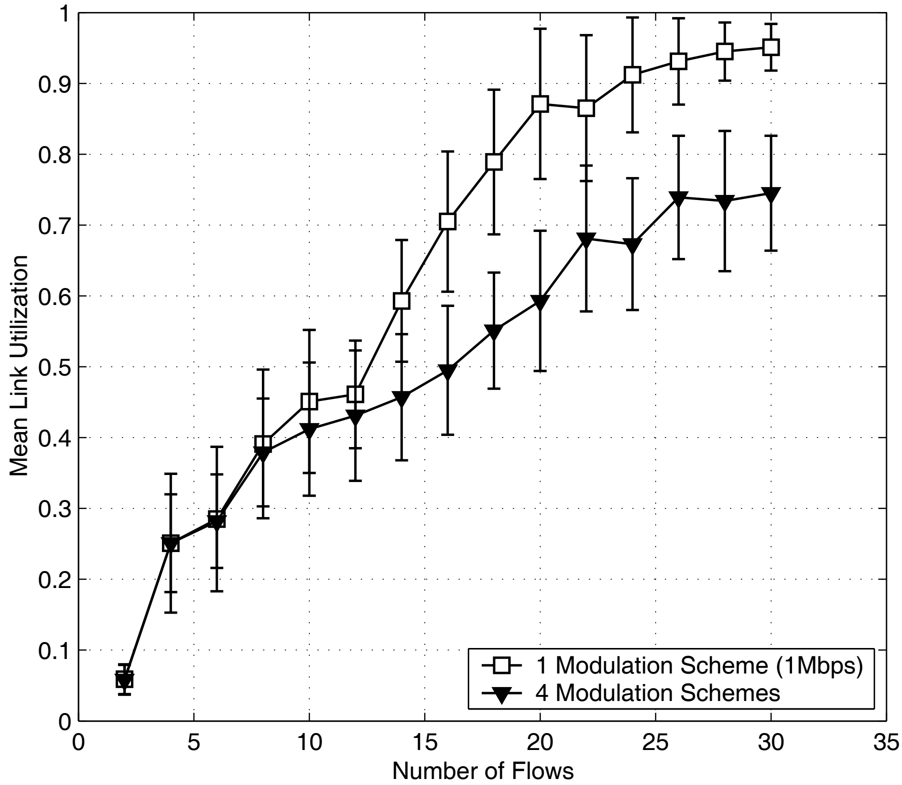


Figure 4.6: Mean link utilization in the whole network

great effect on the performance. The access point is still the bottle neck, because all flows still have to be routed via the access point. But the flows are routed towards the access point with higher bit rate modulation schemes and hence the access point is blocked by a single flow for less time. This allows accepting more flows to be routed.

4.2.4 Conclusion

The performance of rate adaptation and power control in the special wireless multi-hop access network scenario was evaluated. The scenarios consisted of one access point and a number of mobile stations. The performance was thereby measured as the probability that for a specificly loaded network a routing could be found that fulfills the QoS contraints of all flows.

The most striking result is that power control has nearly no influence in the considered access scenarios. The mean link utilization is reduced marginally by using power control, which shows that spatial reuse of the radio channel is facilitated, but the results show that in the considered scenarios, this can hardly be exploited at all. I gave evidence that this is because of the access point being the bottle neck. Hence a spatial bottle neck exists, which cannot be relieved by reducing the spatial interference.

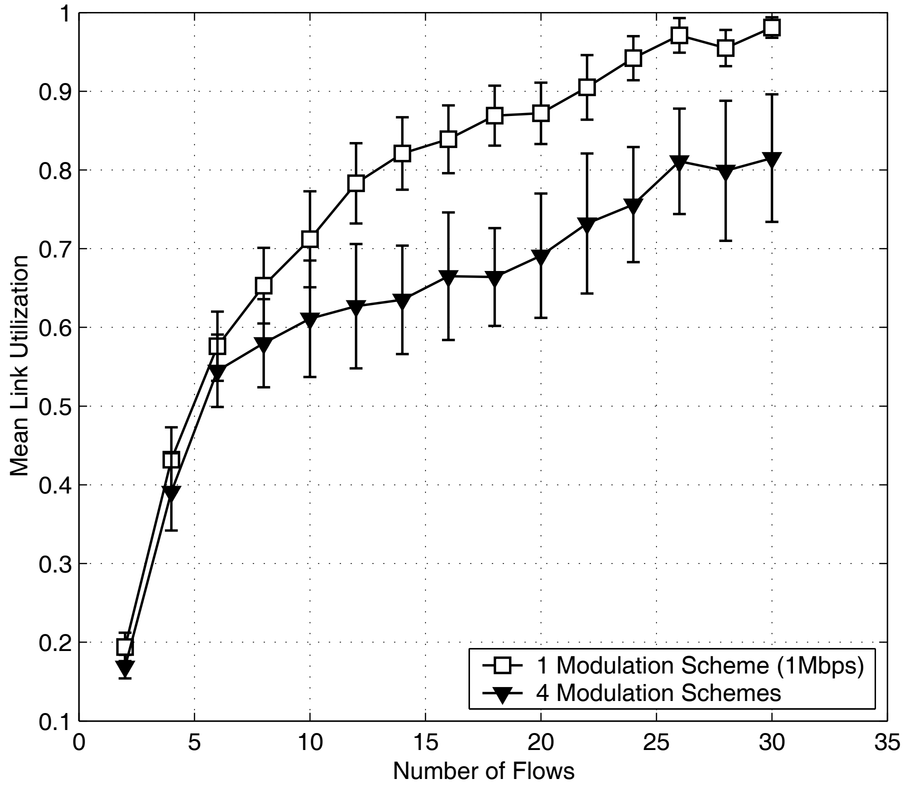


Figure 4.7: Mean link utilization at the access point

Rate adaptation however, which reduces the interference in the network on a temporal basis by reducing the transmission time when using higher modulation schemes, has a great effect on the performance. In some scenarios about double the number of flows could be routed by adding three higher modulation schemes to the basic one.

Considering the deployment issues of power control and further the existing implementations of rate adaptation in WLAN cards today, it can be said that in the Internet access scenario power control is not worth the effort, while rate adaptation surely is.

4.3 Routing Algorithms and Metrics

4.3.1 Introduction

The capacity of wireless multi-hop access networks is limited, as it was shown above. The maximum number of flows that can be routed in a scenario were calculated, by finding a routing that fulfills the QoS constraints of all flows. These routings were found using an optimization problem, which was based on global knowledge: Information about all flows and all nodes were present.

In real networks, this information is usually not present in the routing entity. Usually a routing protocol is run that distributes the information about the nodes, e.g. communication possibilities, possible modulation schemes etc. A routing algorithm uses this information to find a route[1].

In this work, the performance of routing algorithms and metrics in wireless multi-hop access networks is evaluated. Therefore, related work done in the area of ad hoc routing algorithms is reviewed first and some of the proposed algorithms and metrics are selected for the performance evaluation. Further, two new routing metrics are defined that have the aim of using the existing network capacity more efficiently. To evaluate the new and existing algorithms and metrics the theoretical upper bound of the performance is again calculated to judge not only the relative performance, but to compare the work also to an absolute performance limit. Further, a new evaluation procedure, based on an analytical network model, is presented, that lets us calculate the performance efficiently. It is proven by means of simulations that the analytical evaluation procedure fulfills its aim of judging the performance of the routing algorithms.

Again, it is the goal to clearly distinguish the routing algorithm's performance from the routing protocol's performance. Therefore, all influence of the routing protocol is neglected and the routes that certain routing algorithms and metrics would find in a wireless multi-hop access scenario are evaluated only.

Related work on routing algorithms in wireless multi-hop networks is described in detail in the next section. Then, new routing metrics are proposed in Sections 4.3.3.1 and 4.3.3.2. A set of existing algorithms and metrics plus the new metrics are evaluated in Section 4.3.4. Thereby, the performance is investigated in scenarios with Voice over IP and TCP traffic. Further, the performance is compared to a theoretical upper bound of the performance.

4.3.2 Existing Routing Algorithms and Metrics

As described above, several research fields are related to wireless multi-hop Internet access networks and hence a variety of routing algorithms and metrics has been proposed. In this section the main related work is reviewed. It is tried to categorize the existing routing algorithms and metrics. Therefore, I define the following categories, according to the main metric that the algorithms consider:

- Hop Count: These are the most basic routing algorithms that were designed mainly for classical pure ad-hoc networks. They enable communication between peers but do not optimize the routes regarding any performance parameter.

[1] In many implementations and proposals a routing metric is specified. In this case it is assumed that a routing algorithm finds the least cost route according to the given metric. Both, routing algorithms and metrics are interesting and relevant for usage in the wireless multi-hop Internet access scenario

- Stability: Algorithms and metrics from this category aim at using paths that have less probability to suffer from link breaks or outages.
- Link Load: Here, the algorithms and metrics considered try to minimize the load of the network and improve directly user performance parameters, such as delay or throughput.
- Interference: The user performance parameters are here improved by minimizing the interference, e.g. by blocking the shared channel for less time or selecting paths that disturb less other nodes.
- Energy Consumption: Algorithms from this category try to minimize the energy consumption of stations and maximize the network life time.

Putting the existing algorithms in one of these categories is only an attempt to structure the huge amount of related work. Of course there are algorithms that fit into multiple categories, where I either chose the category which fits best to the original aim of the algorithm's designer or mention them in multiple categories.

4.3.2.1 Hop Count

In the early years of ad-hoc networking, researchers first had to solve the problem of enabling basic communication. Performance was rather an issue in terms of protocol overhead than throughput, reliability or delay. As the research concentrated on the ad-hoc routing protocol, the routes that these protocols found were mainly minimum hop count routes. The Minimum Hop Count (MHC) routing metric finds routes that minimize the number of hops between source and destination. I will consider this metric in the performance evaluation, as it represents the most basic way of finding a path and the metric is still used in nearly all routing protocols, today.

4.3.2.2 Stability

Reliability of links and stability of paths is a severe issue in mobile ad-hoc networks. Moving nodes lead to frequent link breaks, which in turn lead to packet loss and throughput degradation. Selecting a more stable path could be obtained with various mechanisms.

One approach is that nodes send beacons to their neighbors with a fixed frequency. From the number of beacons that a neighbor has received in a certain interval, the link quality can be judged. This approach is followed e.g. in [110], [85] and [34]. After judging the quality of the link, different use can be made of it: In [110] links with higher quality are always preferred. This leads to an increased hop count, but as more reliable links are used the path stability is increased. Lundgren et al. use the link quality measure in [85] to avoid Gray Zones. Gray Zones are links, where small routing protocol packets might be transmitted correctly, but larger data packets could not be transmitted reliably. Avoiding these links leads again to a slightly increased hop count, but higher reliability and path stability. In [34] De Couto et al. use beacons to estimate an error probability of links. By means of the error probability an Expected Transmission Count (ETX) is calculated, which defines the expected number of link layer transmission attempts to transmit a packet via that link. Finally a path is chosen, that minimizes the cumulative number of expected link layer transmissions. Here again, more reliable links are chosen. This approach further aims at increasing the path throughput.

A different approach to judge the link quality is to measure the received signal strength. This approach is followed e.g. in [38], [98], [41], [32], [54], [85] and [116]. In general all these algorithms use signal to noise thresholds to avoid highly unreliable links. The problem of this method is that the signal to noise ratio is usually fluctuating heavily, making it difficult to estimate the real mean value. Filters could be used, but suffer always from a tradeoff between agility and stability. Having a fast responsive filter will give unreliable signal to noise estimations and vice versa. However, some of these algorithms have proven to work well in certain scenarios in avoiding highly unreliable links.

Yet, a slightly different approach is followed in [4], where the authors do not use the signal strength directly, but the changes of the signal strength, to estimate the remaining link lifetime. Again, the problem of efficiently filtering the sampled data occurs, but the authors show that with this algorithm a link break can be foreseen in many cases. In [40] Gerharz et al. also estimate the link lifetime. However, it is not done based on signal strength, but on the existing link duration. The mechanism is based on the observation that the remaining lifetime of a link depends on its history. For some mobility models the authors show that their assumption is true. It can be doubted however, that this mechanism is usable in real world scenarios, where mobility occurs more randomly.

Other routing algorithms use the physical distance between nodes to judge the link quality, e.g. [63]. A drawback of this method is that nodes need to be equipped with Global Positioning System (GPS) receivers or other means to know their exact location. Further, the distance is not necessarily directly related to the link quality: Obstacles in the line of sight could render a small distance link useless.

4.3.2.3 Link Load

In ad-hoc networks communication often suffers from low throughput and high delay. This is due to the wireless channel being shared by many nodes. Links are loaded by the flows that are routed via the link itself, but also because other transmissions on the same wireless channel interfere and load the link additionally. There are different approaches in the literature to find paths that reduce the link load and hence increase the throughput and decrease the delay.

Most of the proposed algorithms try to avoid links that are already highly loaded. Therefore, measurements have to be conducted that reveal the current load of nodes and links. The measurements can be active, i.e. introducing additional load in the network, or passive. The active mechanisms could face two problems: The measurements could generate a significant amount of overhead and further lead to route flapping, i.e. instabilities because of the load dependence of the route. The passive mechanisms still could face the route flapping problem, but do not introduce additional load to the network.

One way to measure the load of a node is to count the number of flows that this node is forwarding. This is done e.g. in [48]. However, the number of routed flows is only roughly linked to the load of a node, since flows can have different sizes, forwarding packets via certain links can be more resource consuming, etc.

To consider the link load, the number of packets in the transmission queue can be explicitly used. This is done in e.g. [119], [55], [79] and [31]. Such cross-layer approaches require tampering with the protocol stack, but are shown to reveal the load of a node quite well. In contrast, implicit measurements of the transmission queue state is done in [83], [105] and [3] by measuring packet delays. These delays include queuing, channel access, transmission and possibly retransmission delays and thus reflect not only the load of the link, but a more general quality metric. The wireless channel is usually assumed to be the bottleneck resource and thus the measurement of channel utilization will give information about the state of the transmission and interference area rather than of a single node. This is exploited in [28] and in addition to the queue state also considered in [55], [79] and [31].

Another possibility to avoid highly loaded links or areas is to estimate the currently available bit rate of a link. In fixed networks this could be done by the packet pair approach [69]. Thereby, two packets are sent directly one after each other to ensure that they enter the probed queue back to back. The time difference, between receiving the first and the second packet lets the receiver directly estimate the available bit rate of the bottleneck link. The packet pair approach is used by Draves et al. in [36], however the overhead is significant and the estimation unreliable. In [129] Zhu et al. estimate the available bit rate by assuming a perfectly scheduled TDMA medium access.

For the following performance evaluation, I want to compare a routing algorithm that considers the link load and tries to balance the load. The algorithms that try to avoid highly loaded links or areas are all based on measurements of certain link or node characteristics as can be seen from the description above. Hence, these algorithms performance highly depends on the way this information is gathered, propagated and used, i.e. it depends on the routing protocol. As I want to evaluate the routing algorithms performance only, a way needs to be found to evaluate the effect of load balancing without considering the protocol influence. Therefore, a simple load balancing algorithm is defined, which relies on the link load directly. This metric is usually not easily obtained. Collecting the data is already problematic and filtering it suffers again from the tradeoff between agility and stability. However, for these investigations, it is assumed that the link load is known and Dijkstra's algorithm is used to calculate the routes with the link cost defined as:

$$c_{i,j} = \min\left(\frac{1}{1 - L_{i,j}}, c_{max}\right), \tag{4.11}$$

where $c_{i,j}$ denotes the cost of a link between nodes i and j, $L_{i,j}$ is the current load of a link calculated from all routes in the interference range of the sender i and receiver j and c_{max} is a maximum cost that should not be exceeded. This is theoretical routing algorithm is called the Load Balancing Algorithm (LBA).

4.3.2.4 Interference

An independent approach to load measurements and load balancing is to exploit certain properties of the wireless links. As described above, in [34] De Couto et al. define the expected transmission count (ETX) as the expected number of link layer transmissions that are required to transmit a packet via a certain link. This metric includes a reliability measure and by minimizing the cumulative number of expected transmission counts for a path, the medium usage is minimized. Hence, delay and interference is minimized and the authors show an improved performance in an experimental ad hoc network.

Another property of the wireless link is that most networks are based on rate-adaptive Wireless LAN. Rate adaptivity is the ability to select a modulation scheme with a certain transmission rate for a packet transmission considering signal to noise ratio, packet size, error properties of the modulation scheme etc. such, that the overall throughput is maximized. Basically that means that packets are sent with higher modulation schemes and bit rates via links with a higher signal to noise ratio. The difference in bit rate is significant, e.g. the channel bit rate can be varied from 1 Mbps to 11 Mbps in IEEE 802.11b and up to 54 Mbps in IEEE 802.11a/g. While delay based algorithms implicitly consider the used modulation scheme, Awerbuch et al. consider a multi-rate MAC layer explicitly in [7]. They define the Medium Time Metric (MTM) as the time a transmission will block the wireless medium for other transmissions. Thereby, packet header transmission with a lower modulation scheme, data transmission with a higher modulation scheme, channel access overhead and possible retransmissions are considered, by calculating the cost of a link as:

$$c_{i,j} = \frac{overhead + \frac{size}{rate_{i,j}}}{reliability_{i,j}}, \tag{4.12}$$

where $overhead$ denotes the time required for sending the packet headers with the lowest modulation scheme, accessing the channel and waiting for and receiving the acknowledgement. $size$ is the size of the part of the packet that is transmitted with the higher modulation scheme and $rate_{i,j}$ the data rate of that modulation scheme. $reliability_{i,j}$ is the expected probability that a link layer transmission is successful. Putting that together, we have the cost as the expected time a packet transmission via that link blocks the channel against other transmissions. This algorithm is also selected for the performance evaluation to represent the class of algorithms that does not depend on the load in the network. It is similar to the expected transmission count algorithm, but considers different modulation schemes as well and hence a performance benefit can be expected. In [106] Stevens describes an algorithm that reduces the interference, by minimizing the number of interfered nodes of a path. The algorithm is called Least Interference Routing

(LIR). For each two nodes in combination with every possible transmission power level, an interference indicator, I, is defined. Different ways are proposed to set the interference indicator: Firstly, $I_{i,j} = 1$ if nodes i and j could communicate with good signal strength with each others, else it is set to zero. This means that the indicator is set to one if the nodes are neighbors. Secondly, the Signal to Noise Ratio (SNR) is measured between nodes and according to thresholds the indicator is set to 1, 0.5 or 0. The route with the minimum cumulative interference indicator is chosen. For the performance evaluation I have selected the first possibility to calculate the interference indicator only.

4.3.2.5 Energy Consumption

Energy consumption is a major issue in ad-hoc networks. Many algorithms have been proposed to conserve the energy and increase the network lifetime, e.g. [71]. In [102] the cost of a link is set to the speed of its energy drain, in [103] nodes with low energy do not forward route requests or in [2] the transmission power, load of the node and its battery power is combined to a new routing metric. As most of these algorithms are independent from the throughput increasing algorithms described above, I will not consider them for the evaluation, which concentrates on finding routes that use network capacity most efficiently.

4.3.2.6 Other Algorithms

In [53], [84], [117] and [22] architectures are proposed that combine cellular and ad-hoc networks. Usually the nodes are equipped with a cellular and a WLAN interface. If the quality via the cellular link decreases, proxy clients are searched, that could forward the own connection more efficiently to the base station. Similarly in [72], [86] and [87] two hop relays are considered. It is shown that the performance is good, especially if the relaying node uses a different frequency, e.g. borrowed from a neighboring cell.

4.3.3 Routing Metrics for More Efficient Use of the Network Capacity

In this section new routing metrics are described that try use the existing capacity of the network most efficiently. Therefore, the cumulative interference that a route induces in other nodes is minimized. I propose two different metrics: Firstly, all interfered nodes of a link are considered; secondly, only information about the neighboring nodes of the receiver is used.

4.3.3.1 Interfering Load Metric

In a wireless multi-hop Internet access scenario the node density is usually quite high, because these networks are usually deployed in Hot Spots, where Internet access should be provided for many users. Further, a rate-adaptive WLAN is considered, which is already exploited with the MTM algorithm. The high density of nodes results in many nodes disturbing each others communication; only one node can communicate per time in an interference range. Hence, a routing algorithm should minimize the interference as much as possible to use the existing network capacity most efficiently.

The MTM algorithm makes a first step in minimizing the medium busy time, which effectively minimizes the interference. However, the algorithm does not consider the number of interfered stations. In a region with a high density of nodes, the same medium time would generate more interference than in low density regions. Hence, the idea of the Interfering Load Metric (ILM) is to sum up the link load that is generated in all interfered links, where a communication is possible (i.e. the maximum possible modulation scheme is greater than zero) and use this as the cost of the link. Dijkstra's algorithm is used to calculate the routes with the link cost defined as the number of interfered links times the medium time (medium time is defined in MTM, see Section 4.3.2.4 for the definition).

4.3.3.2 Receiving Load Metric

The problem of the ILM is the kind of information that is required. The knowledge about interfered links is usually not available to a node. Still, I will evaluate this algorithm, but further propose a slightly different one here, which is called Receiving Load Metric (RLM). The number of interfered links is thereby approximated with the number of neighbors of the receiver. This is based on the fact that the density of nodes in the direct neighborhood is strongly related to the number of possible links in the interference region. The number of neighbors is an information that is usually available in nodes, because common routing protocols require that information anyway. It will be shown in the performance evaluation, that one could indeed approximate as such, and achieve a similar performance. Hence, the cost of a link is defined as the number of neighbors of the receiver times the medium time (as defined in MTM in Section 4.3.2.4).

4.3.4 Performance Evaluation

In this section some routing algorithms and metrics are evaluated regarding their performance in a wireless Internet access scenario. A scenario is assumed as described in Section 4.1. The number of users, of flows and also of access points is varied. The location area of the nodes and access points is has a radius of $R = 400$ m.

Users of voice applications are interested in receiving a constant bit rate with a limited error rate and limited delay. Hence, the performance in scenarios with voice users is measured as how many voice flows can be routed without generating an overload situation in the network. This is similar to what has been done in the previous sections to evaluate the performance of rate adaptation and power control. TCP users see their performance rather in terms of throughput and hence this is used as performance measure in the TCP scenarios.

The optimization problem that was formulated in Section 4.2.2 lets us derive a routing without overloaded links, if such a routing exists. Here, this method is used again in the evaluation of the routing algorithms. As all information of the whole network is required for the optimization, which the routing algorithms do not require, this serves as a theoretical upper bound of the performance, i.e. I check with the optimization problem whether a feasible solution exists and see if the routing algorithms are able to find one as well.

In Section 4.1 it was shown how to calculate the load of a link. This allows to judge whether a certain routing of constant bit rate flows in a network leads to an overload situation, where an overload situation is defined to occur if any link has a load greater than one. This is used in the analytical performance evaluation in Section 4.3.4.1, where the routings of all algorithms and metrics are checked for overload situations. This evaluation procedure is fast and allows to quickly judge the performance of a new routing algorithm or metric.

The algorithms are also evaluated by means of simulations. Thereby, first the analytical evaluation procedure is verified by simulating the found routings and see whether a calculated overload situation corresponds to quality of service violation of the constant bit rate flows. Secondly, the routings are also compared regarding their performance with TCP flows.

4.3.4.1 Analytical Performance Evaluation

In this section the performance of the considered routing algorithms and metrics is evaluated analytically for the VoIP service. The analytical evaluation is computationally fast and allows a quick comparison of routing algorithms. However, some assumptions and abstractions have to be made. These assumptions are later verified in the simulations in Section 4.3.4.2.

The procedure of the analytical performance evaluation is simple: The load of each possible link in the network is calculated according to Eqn. 4.4. Then the link loads are checked for overload situations. The assumption is that an overload situation leads to increased packet loss (because of the queue overflow) and delay (because of excessive queuing) and renders the voice service useless. Hence, if an overload situation exists for a specific routing, this routing is counted as non-feasible. If no such overload situation is found the routing is counted as feasible. The optimization problem does basically the same: The load of the links is one of the constraints to be fulfilled.

To analyze the performance of the routing algorithms the following parameters of the scenario are varied:

- Number of access points: {1,3}
- Number of nodes: {40,50,60,70}
- Number of flows: {2,4,...,30} (1 access point) and {4,8,...,60} (3 access points)

For each parameter combination ten scenarios were generated with a random user and flow distribution. The main goal of the multi-hop access network is to route as many voice flows as possible to/from the access point. So again, the main performance parameter is defined as the fraction of scenarios for a specific parameter combination for which a feasible routing could be found. A feasible routing is found if the link load on any link in the network where a communication is possible is smaller or equal to one. Further the mean link load in the network is measured as the average link load over all links in the network, where communication is possible.

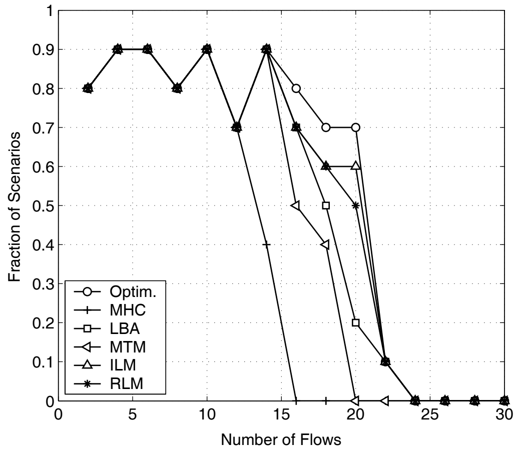
According to the procedure as described above, the performance of some algorithms from related work and the two newly defined routing metrics is evaluated and further the theoretical upper bound of the performance is shown. The algorithms and metrics from related work are the Minimum Hop Count (MHC) as the most basic routing metric, the Load Balancing Algorithm (LBA) that was defined to be a representative for the proposed algorithms that find paths which balance the load in the network, the Medium Time Metric (MTM) and the Least Interference Routing (LIR) algorithm. The Interfering Load Metric (ILM) and Receiving Load Metric (RLM) are evaluated, whether they fulfill their aim of using the network capacity in a Hot Spot scenario well. Further the result of the optimization problem is shown which provides an upper bound to the number of flows that can be supported.

In Fig. 4.8 the fraction of scenarios where a feasible routing was found and the mean link load for the scenarios with one access point and 40 - 60 nodes are shown. The graphs of the LIR algorithm are not shown, as this algorithm found in all scenarios exactly the same routes as the MHC algorithm. This can be reasoned by the fact that all nodes near the center of the location area have usually at least seven to ten neighbors, depending on the scenario. The nodes near to the center of the area are the important nodes, where the routing algorithm could lead to benefits, because all flows need to be routed to the center and hence this is the region of the most interference. The variability of the number of neighbors is however not that large. Hence, the routing metric is dominated by a constant offset for all nodes, which in fact lets the algorithm find the same routes as MHC does.

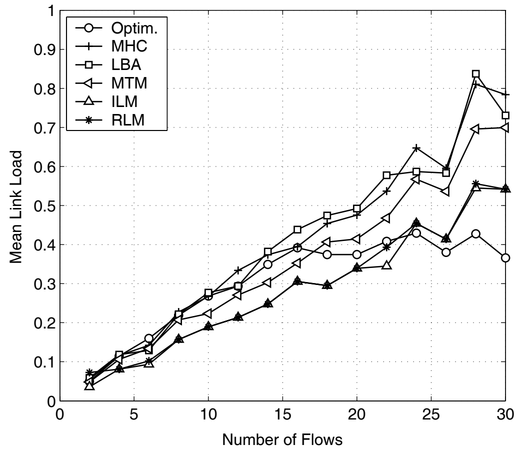
In general it can be seen that scenarios with few flows, i.e. up to 10 flows, all routing algorithms usually find feasible routings. Only in the scenarios with 40 nodes, all algorithms fail for few of these low load scenarios. This is because some nodes are simply not connected and so of course no routing algorithm is able to route packets to these nodes. In higher node density scenarios, this happens only once more, with 50 nodes and 8 flows.

If the load of the network increases, i.e. the number of flows to be routed increases, some algorithms are not able to find a feasible routing any more. In very high load scenarios it is never possible to find a routing without overload situations; in these cases the network is simply overloaded. This trend can be seen also from the mean link load, which increases rapidly with the number of flows.

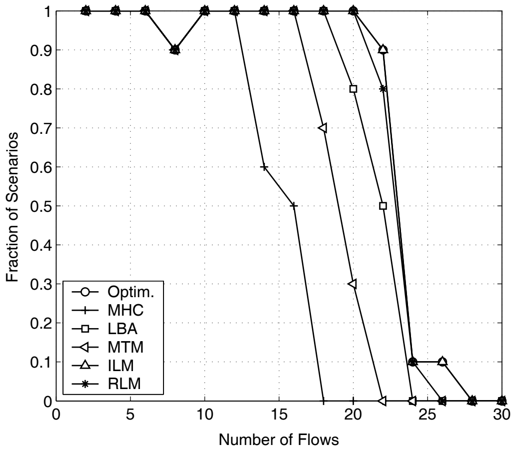
Looking at the differences between the routing algorithms, we see that the minimum hop-count (MHC) routing is always the worst performing algorithm, e.g. in 50 node scenarios, 12 flows could already lead to the MHC not finding a feasible solution; all other



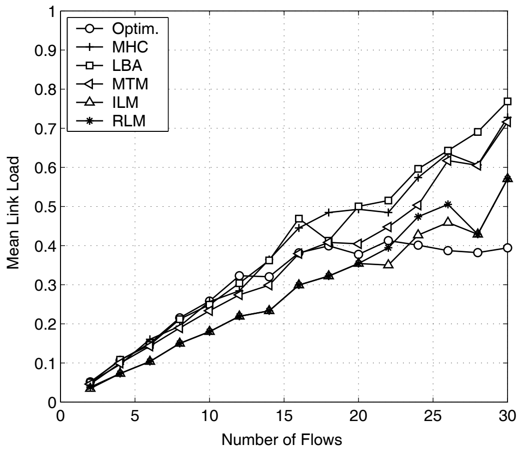
(a) 40 Nodes



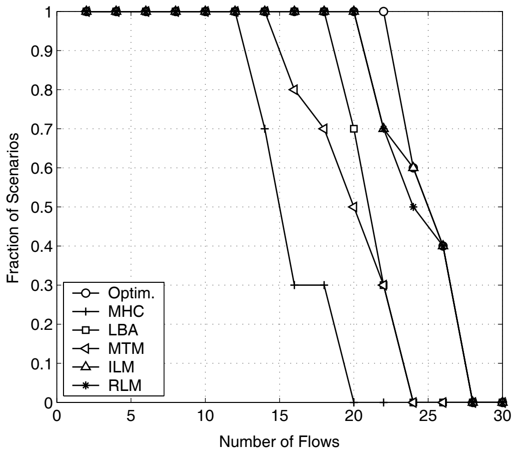
(b) 40 Nodes



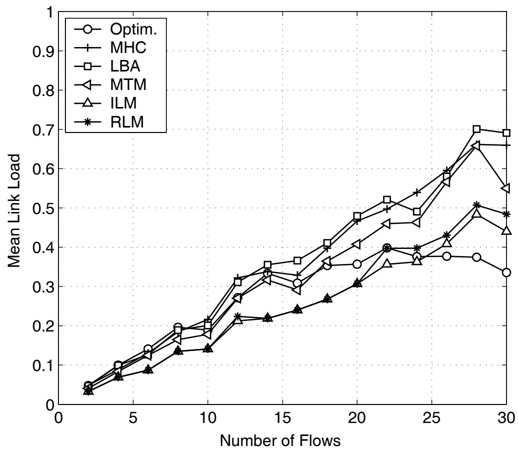
(c) 50 Nodes



(d) 50 Nodes



(e) 60 Nodes



(f) 60 Nodes

Figure 4.8: Fraction of scenarios where a feasible routing was found and the mean link load for the scenarios with one access point

algorithms find for all scenarios up to at least 16 flows a feasible routing. One could expect MHC to perform worst, as it is not designed to perform well in the Internet access scenarios, but to enable communication at all. Still, the significance of the room for improvements was surprising.

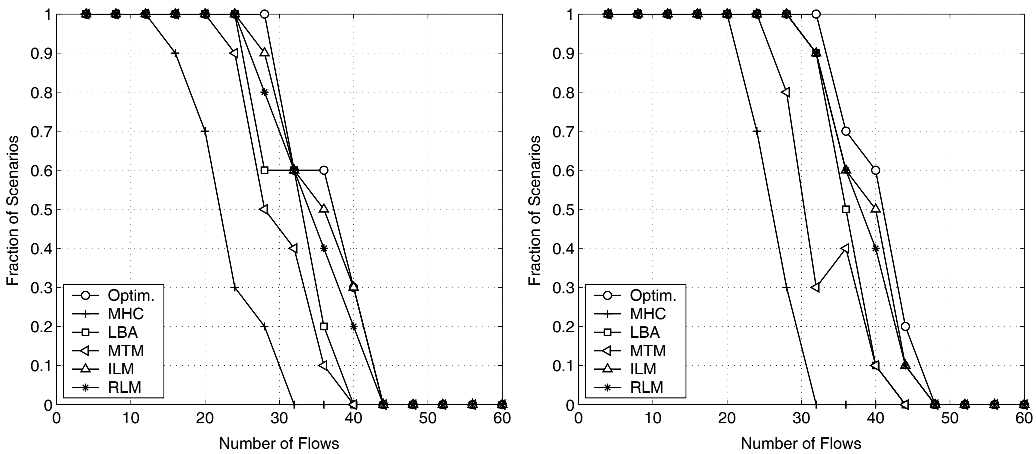
The optimal solution is of course always best performing, as it represents the theoretical upper bound. Looking at the mean link load figures, we can see how this is achieved. In low load scenarios, the mean link load is comparable to the other algorithms. This is because several solutions exist and the optimizer selects the solution with the least delay, which fulfills the constraints. In higher load scenarios this is different. Here, the number of possible solutions is restricted to the ones with a lower mean link load and one of these is selected. That is why we see a knee in the curves of the mean link load for the optimal solution.

The medium time metric (MTM) performs significantly better than MHC. However, there is still room for improvements. With MTM, links with higher modulation schemes are preferred, which leads to the performance gain. However, because of the special scenario, where the traffic density is high, especially around the access point, minimizing the medium usage is not enough. The mean link load is not significantly decreased, because even if the medium time of a flow is decreased, the interference is not necessarily decreased as well, because it also depends on the number of disturbed nodes.

That is, why the load balancing algorithm (LBA) performs nearly always better than MTM. Here, regions of high load are avoided and the network is more evenly loaded. We see that the mean link load is not significantly decreased compared to MTM or MHC. However, the performance in terms of fraction of correctly routed scenarios is. This is because the network is not less loaded, but rather more evenly. However, the performance for this algorithm will always depend on how the first routes are found. As these are basically routed according to the minimum hop count, problems arise after a few of these have been routed. The load balancing will now try to avoid the regions which are used already, but as long hops were preferred for the first routes, the next routes have to avoid large regions of high load.

This is what should be avoided with the newly defined algorithms: Generating much interference right at the beginning should be avoided. Each flow individually tries to minimize the interference it generates, which is beneficial for all flows. This can be clearly seen from the mean link load of the two algorithms (both induced load metric (ILM) and received load metric (RLM)): Right from the beginning the mean link load is low and significantly below the load of the other algorithms. In heavily loaded scenarios, the optimal solution reduces the link load as well, to values comparable to the two algorithms'. We see that this theoretical limit is reached impressively well with the ILM. It performs nearly as good as the optimal solution. Further, its approximation, i.e. RLM, performs nearly as good.

In Fig. 4.9 the results of the analytical evaluation for three access points are shown. Basically, the results are similar to what we have seen for one access point. Especially, the relative performances of the different routing schemes are the same as in the one access point scenarios. That is why only the results of the scenarios with 40 and 60 nodes are



(a) 40 Nodes (b) 60 Nodes

Figure 4.9: Fraction of scenarios where a feasible routing was found and the mean link load for the scenarios with three access points

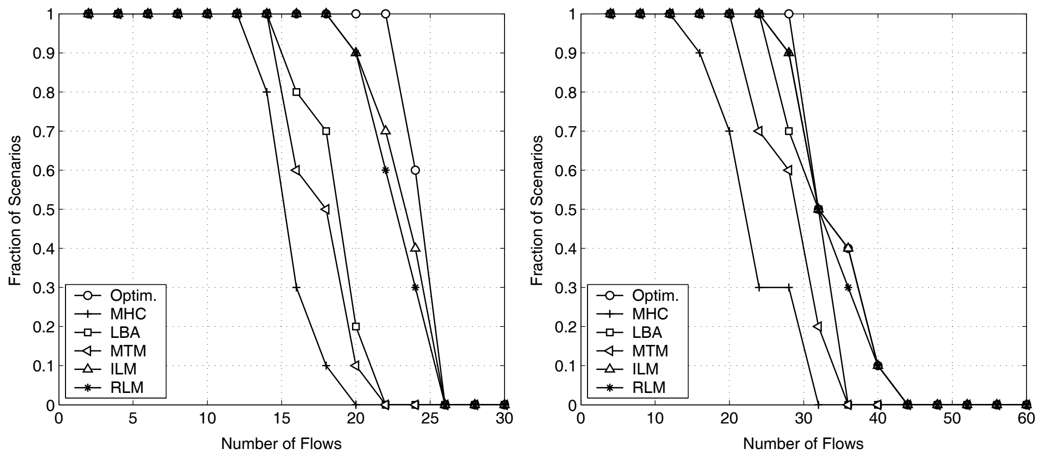
shown. However, the capacity of the network in the same area is increased. With three access points it is in all scenarios possible to route more than 30 flows with the optimal solution. Considering all schemes, it can be said that the capacity is increased by roughly 50%. This is achieved, by increasing the number of access points by 200%. The possible capacity gain from changing the routing scheme from MHC is nearly 100%. The realized capacity gain by using the RLM compared to MHC routing is more than 50%. Hence, we can conclude that we can gain much more from changing the routing algorithm than from using more access points in the same area. However, the two methods to increase the performance do not interfere with each other and changing the routing scheme and placing more access points in the area performs best.

4.3.4.2 Verification of the Analytical Evaluation Procedure

In the previous section the routing algorithms and metrics were evaluated by means of an analytical analysis: The link load was calculated and routings were counted as feasible if the load is smaller than one on every link in the network. To verify the proposed evaluation procedure as well as the performance results, I have also conducted simulations of the scenarios described above. The results are shown in the next section.

The analytical performance evaluation shown in the previous section has revealed interesting insights to the performance of the different routing algorithms and metrics. However, assumptions were made, which need to be verified in order to trust the results. Especially, it was assumed that an overload situation means that the QoS cannot be guaranteed. In this section this assumption is verified by means of simulations.

Again, the networks simulator and MAC layer enhancements are used as explained in Section 2.5. In order to evaluate the routing algorithms only, without influence from the



(a) 60 Nodes; 1 Access Point (b) 60 Nodes; 3 Access Points

Figure 4.10: Simulation results showing the fraction of scenarios where a feasible routing was found.

routing protocol, a slightly modified version of the Manual Routing Agent is used. With this routing agent, fixed routes are configured for all flows in all nodes off-line at simulation start-up. During the simulation, no routing messages are exchanged.

In the simulations a routing is said to be feasible if all flows' packet loss rate is below 5%. This is in contrast to the analytical performance evaluation, where it was looked for overload situations in the network. By means of the simulations it should be shown that these two criteria lead to similar performance results.

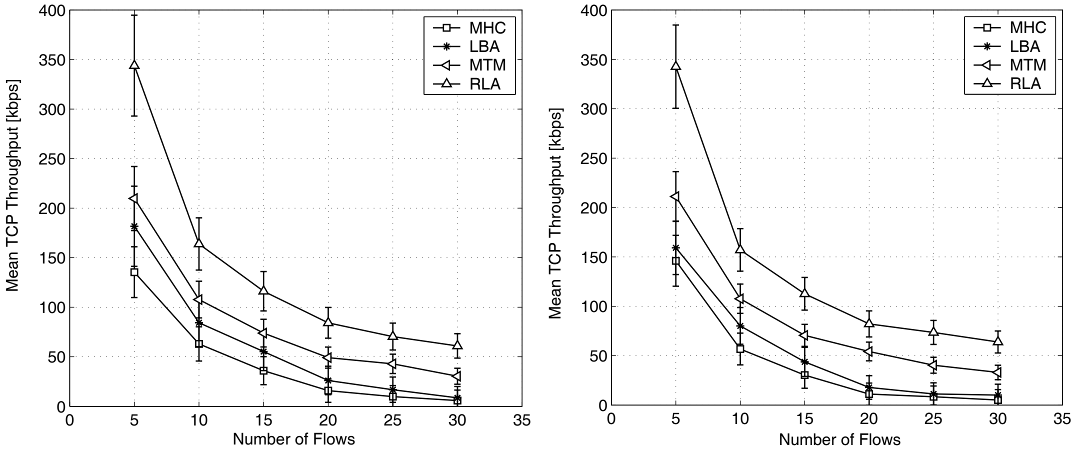
As the results are similar as in the analytical performance evaluation the results are shown only for the fraction of correctly routed scenarios and only for 60 nodes and one and three access points in Fig. 4.10. We see that the performance is slightly worse here than in the analytical evaluation. This is because in the analytical evaluation it was assumed that a link can be loaded up to 100%. However, because of the probabilistic channel access scheme, this cannot be reached in the simulations.

Still, the results match the simulated results quite well and especially the relative performance between the routing algorithms and metrics is the same. Hence, we can summarize that the evaluation procedure gives reliable results.

4.3.4.3 Simulation of TCP Traffic Scenarios

The analytical evaluation is used for a computational fast answer on whether a routing algorithm or metric is able to efficiently route VoIP flows. Additionally to the VoIP services, also scenarios where users download data from the Internet via TCP are simulated here. The TCP performance of the routing algorithms and metrics is presented below.

TCP downloads are still the most common application of WLAN Hot Spots today. As the data rate of the TCP connection is not fixed, but adapted by the TCP sender according to



(a) 40 Nodes (b) 60 Nodes

Figure 4.11: Simulation results showing the received TCP throughput.

the seen packet loss rate and round trip time, it is not possible to use the same performance criterion as for the VoIP evaluation. The quality of the routing algorithm does not influence the number of flows that can be routed, but rather the throughput that each flow would get. Hence, the received amount of TCP data is measured and divided by the time in which this data was received. This metric is called the TCP throughput.

The simulation results of the ILM are omitted in the figures for the sake of a better readability. Its performance is similar to the RLM's. In Fig. 4.11 we see again the same performance trend: MHC routing performs worst leading in some scenarios to less than half the throughput that is achieved with the routes from RLM. LBA performs worse than MTM in these scenarios. This is because the selection of high throughput links is more efficient than balancing the load with TCP. As TCP adapts its sending rate it can make use of high throughput paths in contrast to VoIP flows, which have a fixed data rate.

4.4 Conclusions

Multi-hop Internet access has characteristics that are similar to pure ad-hoc networks in some sense, but differ significantly in others. In order to judge the performance of routing algorithms and metrics, an analytical performance evaluation method was proposed first. Further a theoretical upper performance limit was calculated. The method allows a quick look on the performance of routing algorithms in case of fixed bit rate traffic. The methodology was verified by means of simulations.

A second contribution was to investigate the performance of ad-hoc routing algorithms in a wireless multi-hop Internet access scenario. We got very diverse results. Minimum Hop Count routing, as done by most ad-hoc routing protocols today, does not perform well and compared to the theoretical performance bound leaves much room for optimizations.

Some algorithms from the related work were investigated and have shown already an increased performance. However, there was still space left for optimization.

Thirdly, two new routing metrics were proposed, where the first, namely Induced Load Metric is more of a theoretical nature, as it requires information in the nodes that is usually not available. The second metric, namely Receiving Load Metric is an approximation of the first one, but requires only very little information and hence is expected to be easy to implement. The performance of these metrics is outstanding, as both nearly reach the theoretical upper bound in our scenario. This was shown by means of an analytical performance evaluation and further simulations.

Chapter 5

Conclusions

In this thesis wireless Internet access in Hot Spots was considered. Hot Spots are characterized by a high user density, which poses strong performance requirements on the wireless communication system. WLAN, as the mainly used access network in these scenarios provides high bit rate communication means, but offers an unreliable service with a highly variable quality.

To stabilize and increase the quality of service, the performance influencing parameters were identified first. Therefore, an analytical model of rate adaptive WLAN was developed. Although, a variety of analytical models exists, I have shown that none of these fits the requirements for the envisioned performance analysis. The new model is characterized by considering the bit level, packet level and different applications at flow level. It was shown that for all considered applications, namely CBR, TCP and bi-directional voice traffic, a strong dependence of the user performance on the locations of all other users exists. The throughput was shown to be degraded significantly, only because one or more other users have moved away from the access point. The model and all results were verified by simulations.

The strong user location dependence of the throughput inspired the development of a location based quality of service control system. Two such systems were proposed. Firstly, users within a certain admission area are allowed to communicate with the access point, and calls from users outside the admission area are blocked. In the performance evaluation it is shown that blocking few users can be beneficial for all other users and allows to give some quality of service guarantees. However, the static size of the admission area makes the system usable for certain scenarios only. In high load scenarios, a small admission area is required; in low load scenarios a large admission area can increase the performance. To account for that I proposed a second quality of service control system, which maintains a variable admission area, depending on the load of the access point. The analytical model is thereby used to estimate the quality of service that each user would achieve if a set of users were allowed to communicate. Based on this estimation, users at a large distance to the access point are blocked until the quality of service of all other users is sufficient. In simulations the superiour performance of this system compared to the more simple control system and the case of not controlling the quality of service at all was shown. Further it was shown that the overall performance of the access point in terms of throughput is increased.

Although, this quality of service control system allows some users to use applications that require quality of service guarantees, the amount of users that can be accepted is limited. Deploying such a system in a Hot Spot, would still mean that many users have to be blocked. The problem here is the limited capacity of the network. One way to increase the capacity is multi-hop communication. Multi-hop communication brings however new problems, as routes from source to destination have to be found and maintained. Routing protocols and algorithms care about that. The routing protocols thereby communicate the information about links between all nodes and allow the nodes in this way to use a routing algorithm to calculate the route. The routing protocols become more complicated and require more resources the more information needs to be transported. The amount of information depends on the routing algorithm and the wireless communication technologies that are used.

Hence, different wireless communication technologies were investigated first. I concentrated on the use of power control and rate adaptation. The WLAN channel access does not work well with power control, but by spending some effort and resources it is possible to use it. Rate adaptation can easily be used, as it is implemented in WLAN cards today, but poses more requirements on the routing protocol: Information about used modulation schemes needs to be propagated. Hence, both technologies can be used, but require an effort to do so. The task was now the see if these technologies provide a performance gain that makes it worth the effort. To judge the performance gain, I formulated an optimization problem to find routes for all flows in the network. Thereby constraints are formulated that guarantee some quality of service to all flows. By solving this problem, it was judged how many flows under the given constraints could be routed. This was used as the performance measure. Using power control with different number of power levels, it is not possible to significantly increase the number of flows. Using rate adaptation it is. Both of these technologies reduce the interference in the network: Power control on a spatial basis, by transmitting with a lower power and hence interfering stations in a smaller area. Rate adaptation on a temporal basis, by transmitting faster and interfering other stations for less time. While the spatial interference reduction does not lead to a significant performance gain, the temporal does. This is because of the access point being a central bottle neck. The access point is utilized less if the links use higher transmission speed. It is however not relieved in cases where the links are used with less power. In these cases the utilization of the links is constant, regardless of the number of different power levels used. By means of this investigation, it was concluded that power control in wireless multi-hop access networks is not worth the effort, while rate adaptation surely is.

After identifying the performance limits under the assumption of global knowledge, the next task was to find routing algorithms that come as close to these limits as possible. To evaluate routing algorithms, a new methodology was proposed. By means of the analytical model from the first part, the load of all links in the network was calculated, considering all flows that are routed. Thereby, the link load is influenced by the transmissions via this link and all transmissions via links from the same interference region. Having calculated the load of all links, it was checked whether any link load exceeds one. If so, an overload situation exists, and it is assumed that the quality of service cannot be kept. Doing that for different routing algorithms enables to compare them regarding their possibility to find routings that guarantee the quality of service of all flows. The methodology was verified by simulations and was shown to be a computational fast means to estimate the performance of routing algorithms.

Some routing algorithms and metrics from related work were evaluated and it was shown that there is space left to the performance of the theoretical upper limit. To exploit this, new routing metrics were proposed, that try to use the existing capacity of the network as efficient as possible. Thereby the interference on a temporal and spatial basis is minimized. In the first metric this is done considering all nodes of a flows interference range. It is usually not possible to have knowledge about all nodes in the interference range, but it is common practice in most routing protocols to maintain a list of all nodes in the transmission range. Hence, the nodes in the interference range were approximated by the nodes in the transmission range in the second metric. In a performance evaluation it was shown that both metrics nearly reach the theoretical performance bounds.

Summarizing, it can be said that wireless Internet access in Hot Spots with standard WLAN equipment is possible but not optimized. The evaluation has shown that the user performance is highly variable and further the capacity of the system is reached when a reasonable number of users connect to the access point directly. The capacity of the network can be increased by multi-hop communication, where the use of rate adaptation is recommended, while power control should not be used. Routing algorithms are proposed that use the capacity of the network most efficiently. The algorithms require only little information and can be embedded in standard ad hoc network routing protocols.

Appendix A

Publications

C. Burmeister, U. Killat, J. Bachmann, *Location-Based Services via IEEE 802.11b Hot Spots*, Proceedings of the 3rd Polish-German Teletraffic Symposium PGTS 2004, Dresden, Germany, September 2004.

C. Burmeister, U. Killat, J. Bachmann, *TCP Throughput Optimized Handover Decisions*, Proceedings of the 11th European Wireless Conference 2005, Nicosia, Cyprus, April 2005.

C. Burmeister, U. Killat, J. Bachmann, *Modeling Rate-Adaptive Wireless LAN with an M/G/1/B Queueing System"*, Proceedings of the International Conference on Communication Systems and Networks (CSN 2005), Benidorm, Spain, September 2005.

C. Burmeister, U. Killat, J. Bachmann, *An Analytical Model of Rate-Adaptive Wireless LAN and its Simulative Verification*, Proceedings of the third ACM International Workshop on Wireless Mobile Applications and Services on WLAN Hotspots (WMASH05), Cologne, Germany, September 2005

C. Burmeister, U. Killat, J. Bachmann, *TCP over Rate-Adaptive WLAN - An Analytical Model and its Simulative Verification*, Proceedings of the 7th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks, Niagara-Falls/Buffalo, USA, June 2006.

C. Burmeister, U. Killat, K. Weniger, *The Performance Gain of Rate Adaptation in Wireless Multi-Hop Access Networks*, Proceedings of the IFIP Fifth Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net 2006), Lipari, Italy, June 2006.

C. Burmeister, U. Killat, J. Bachmann, *Performance of Rate Adaptive Wireless LAN*, Elsevier International Journal of Electronics and Communication (AEUE), Vol.61, Issue 8, 2007.

C. Burmeister, U. Killat, K. Weniger, *Power Control in Wireless Multi-Hop Access Networks - Is it Worth the Effort?*, Proceedings of the 20th International Teletraffic Congress, Ottawa, Canada, June 2007.

Appendix B

List of Acronyms

ACK	Acknowledgement
AP	Access Point
BSS	Basic Service Set
CA	Collision Avoidance
CBR	Constant Bit Rate
CCITT	Comité Consultatif International Téléphonique et Télégraphique
CCK	Complementary Code Keying
CDF	Cumulative Distribution Function
CTS	Clear to Send
COST	Coopération européenne dans le domaine de la recherche scientifique et technique
CRC	Cyclic Redundancy Check
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
DBPSK	Differential Binary Phase Shift Keying
DQPSK	Differential Quadrature Phase Shift Keying
DCF	Distributed Coordination Function
DIFS	Distributed Inter-Frame Space
DS	Distribution System
DSSS	Direct Sequence Spread Spectrum
EDCF	Enhanced Distributed Coordination Function
EIFS	Extended Inter-Frame Space
ESS	Extended Service Set
ETX	Expected Transmission Count
FCS	Frame Check Sequence
FHSS	Frequency Hopping Spread Spectrum
FIFO	First In First Out

GPRS	General Packet Radio Service
GPS	Global Positioning System
GSM	Global System for Mobile Communications
HSDPA	High Speed Downlink Packet Access
HTTP	Hypertext Transfer Protocol
IEEE	Institute of Electrical and Electronics Engineers
ILM	Interfering Load Metric
IP	Internet Protocol
IR	Infra-Red
ISM	Industrial Scientific Medical
LBA	Load Balancing Algorithm
LBT	Listen Before Talk
LIR	Least Interference Routing
LTBQC	Location and Throughput Based QoS Control
MAC	Medium Access Control
MHC	Minimum Hop Count
MIP	Mixed-Integer Linear Program
MMGI	Markov Modulated General Independent
MTM	Medium Time Metric
NAV	Network Allocation Vector
OFDM	Orthogonal Frequency Division Multiplexing
PC	Point Coordinator
PCF	Point Coordination Function
PDF	Probability Density Function
PDU	Protocol Data Unit
PHY	Physical Layer
PIFS	PCF Inter-Frame Space
PLCP	Physical Layer Convergence Protocol
PMD	Physical Medium Dependent
QoS	Quality of Service
RLM	Receiving Load Metric
RTP	Real-Time Transport Protocol
RTS	Ready to Send
SDU	Service Data Unit
SFD	Start Frame Delimiter
SIFS	Short Inter-Frame Space
SLBQC	Simple Location Based QoS Control
SNR	Signal to Noise Ratio
SYNC	Synchronization
TCP	Transmission Control Protocol

TDMA	Time Division Multiple Access
UDP	User Datagram Protocol
UMTS	Universal Mobile Telecommunications System
VoIP	Voice over IP
WIFI	Wireless Fidelity
WLAN	Wireless Local Area Network
WWW	World Wide Web

Appendix C

List of Symbols

$a_{i,j}^p$	node interference indicator between nodes i and j with power level p
b	backoff value
$b_{(i,j),(k,l)}^p$	link interference indicator
B	capacity of the queue at the access point [packets]
$c_{i,j}$	assigned link cost between node i and j
CW	contetion window size
CW_{max}	maximum contention window size
CW_{min}	minimum contention window size
CW_{new}	newly calculated contention window size
CW_{old}	previous contention window size
d	distance
d_{max}	maximum distance of a user to the access point for being accepted
d_f	destination node of flow f
d_n	distance between access point and user n
$D_{i,j}^{m,p}$	transmission delay of a packet from node i to node j with modulation scheme m and power level p
E_c	energy per chip
F	number of flows in the network
f_c	carrier frequency
G	processing gain
$I_{i,j}$	interference indicator used in LIR
k	transmission attempt
K	number of voice users
L	length of the MAC payload
$L_{i,j}$	load of the link between node i and j
L_{ack}	length of a MAC acknowledgement

L_{mac}	length of a MAC header
L_{plcp}	length of a PLCP preamble and header
m	modulation scheme being used
$mmax_{i,j}^{p}$	maximum possible modulation scheme to be used between nodes i and j with power level p
M	number of different modulation schemes
n_c	number f chips
n_b	number of bits
N	number of users
N_0	spectral density of the noise at the receiver
$N_{tx,n}$	number of transmission attempts required from station n
p	TCP packet loss probability
p_a	TCP acknowledgement loss probability
p_b	bit error probability
p_c	collision probability
p_e	transmission attempt error probability
p_i	probability that the channel is idle during a time slot
p_I	packet drop rate of the Internet link
p_k	probability that a packet is transmitted correctly in the k-th attempt
p_q	packet drop probability in a queue
$p_{s_1 t_1; s_2 t_2}$	transition probability from state (s_1, t_1) to state (s_2, t_2)
$p_{s,n}$	probability that station n is silent during a time slot
$p_{tx,n}$	probability that station n starts a transmission in a time slot
$p_{tx=1}$	probability that exactly one station starts sending in a time slot
$p_{tx>1}$	probability that more than one station start sending in a time slot
p_w	TCP packet loss probability on the wireless link
P	number of transmission power levels
P_L	path loss
P_r	received power
P_t	transmitted power
$r_{i,j}^{f,m,p}$	routing indicator
R	bit rate of the currently used modulation scheme
R_c	chip rate
R_f	packet rate of flow f
R_t	transmission rate
$R_{i,j}^{m,p}$	possible packet transmission rate between nodes i and j using modulation scheme m and power level p
RTT	round trip time
RTT_0	basic round trip time

s_f	source node of flow f
T_b	channel busy time
T_{bo}	backoff time
T_c	contention time
T_{ca}	channel access time
T_{difs}	time of a DIFS
T_{eifs}	time of a EIFS
T_i	channel idle time
T_{plcp}	transmission time of PLCP preamble and header
T_q	queueing time
T_{sifs}	time of a SIFS
T_{slot}	time slot time
T_{ta}	transmission time of the TCP acknowledgement packet
T_{td}	transmission time of the TCP data packet
T_{tx}	transmission time
TP	throughput
TP_{min}	minimum throughput to fulfill QoS requirements
U	utilization of the channel
X_σ	random variable describing the effect of obstacles on the received power
λ_i	overall packet arrival rate at the access point
λ_i	packet arrival rate for user i
$\lambda_{tx,n}$	MAC layer transmission rate of station n
$\lambda_{tx,0}$	MAC layer transmission rate of the access point
μ	service rate
$\pi_{s,t}$	stationary probability of state (s,t)
ρ	user density

Bibliography

[1] I. Aad, C. Castelluccia, *Differentiation Mechanisms for IEEE 802.11*, Proceedings of the Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2001), Anchorage, USA, April 2001.

[2] M. Abolhasan, J. Lipman, J. Chicharo, *A routing strategy for heterogeneous mobile ad hoc networks*, Proceedings of the 6th IEEE Circuits and Systems Symposium on Emerging Technologies: Frontiers of Mobile and Wireless Communications. Shanghai, China, 2004.

[3] A. Adya, P. Bahl, J. Padhye, A. Wolman, L. Zhou, *A Multi-Radio Unification Protocol for IEEE 802.11 Wireless Networks*, Proceedings of the First International Conference on Broadband Networks (BROADNETS'04) , Washington, DC, USA, 2004.

[4] S. Agarwal, A. Ahuja, J.P. Singh, R. Shorey, *Route-lifetime assessment based routing (RABR) protocol for mobile ad-hoc networks*, Proceedings of the IEEE International Conference on Communications (ICC 2000),New Orleans, USA, 2000.

[5] S. Agarwal, R.H. Katz, S.V. Krishnamurthy, S.K. Dao, *Distributed Power Control in Ad-hoc Wireless Networks*. Proceedings of the 12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2001), San Diego, CA, USA, 2001.

[6] E. Altman, K. Avrachenkov, C. Barakat, R. Nunez-Queija, *State-dependent M/G/1 type queueing analysis for congestion control in data networks,* Computer Networks: The International Journal of Computer and Telecommunications Networking, Vol.39, No.6, August 2002.

[7] B. Awerbuch, D. Holmer, H. Rubens, *High throughput route selection in multi-rate ad hoc wireless networks*, Proceedings of the First Working Conference on Wireless On-demand Network Systems and Services (WONS 2004), Trento, Italy, 2004.

[8] A. Banchs, X. Perez-Costa, D. Qiao, *Providing Throughput Guarantees in IEEE 802.11 e Wireless LANs*, Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC2002), Orlando, USA, March 2002.

[9] A. Banchs, L. Vollero, *A delay model for IEEE 802.11e EDCA*, IEEE Communications Letters, Vol.9, No.6, pp.508-510, June 2005.

[10] A. Banchs, P. Serrrano, A. Azcorra, *End-to-end delay analysis and admission control in 802.11 DCF WLANs,* Computer Communications, Vol.29, No.7, pp.842-854, April 2006.

[11] M. Barry, A. T Campbell, A. Veres, *Distributed Control Algorithms for Service Differentiation in Wireless Packet Networks,* Proceedings of the Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2001), Anchorage, USA, April 2001.

[12] B. Bellalta, M. Oliver, M. Meo, M. Guerrero, *A simple model of the IEEE 802.11 MAC protocol with heterogeneous traffic flows,* The International Conference on *Computer as a tool* (EuroCon 2005), Belgrade, Serbia & Montenegro, November 2005.

[13] G. Bianchi, L. Fratta, and M. Oliveri, *Performance analysys of IEEE 802.11 CSMA/CA medium access control protocol,* Proceedings of the Seventh IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'96) Taipei, Taiwan, October 1996.

[14] G. Bianchi, *IEEE 802.11-saturation throughput analysis,* IEEE Communications Letters, Vol.2, No.12, pp.318-320, December 1998.

[15] G. Bianchi, *Performance analysis of the IEEE 802.11 distributed coordination function,* IEEE Journal on Selected Areas in Communications, Vol.18, No.3, pp.535-547, March 2000.

[16] G. Bianchi, I. Tinnirello, *Remarks on IEEE 802.11 DCF performance analysis* IEEE Communications Letters, Vol.9, No.8, pp.765-767, August 2005.

[17] R. Bruno, M. Conti, E. Gregori, *Throughput Evaluation and Enhancement of TCP Clients in Wi-Fi Hot Spots,* Proceedings of the First IFIP Working Conference on Wireless On-demand Network Systems (WONS 2004), Madonna di Campiglio, Italy, January 2004.

[18] R. Bruno, M. Conti, E. Gregori *Analytical Modeling of TCP Clients in Wi-Fi Hot Spot Networks,* The Third IFIP-TC6 Networking Conference, Athens, Greece, May 2004.

[19] F. Calì, M. Conti, E. Gregori, *Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit,* IEEE/ACM Transactions on Networking Vol.8, No.6, pp.785-799, December 2000.

[20] G. Cantieni, Q. Ni, C. Barakat, T. Turletti, *Performance Analysis of Finite Load Sources in 802.11b Multirate Environments,* Elsevier Computer Communications Journal, Special Issue on Performance Issues of Wireless LANs, PANs, and Ad Hoc Networks, 2004.

[21] G.R. Cantieni, Q. Ni, C. Barakat, T. Turletti, *Performance analysis under finite load and improvements for multirate 802.11*, Computer Communications, Vol.28, No.10, pp.1095-1109, June 2005.

[22] R. Chang, W. Chen, Y. Wen, *Hybrid wireless network protocols*. IEEE Transactions on Vehicular Technology, Vol.52, No.4, pp.1099-1109, 2003.

[23] P. Chatzimisios, A.C. Boucouvalas, V. Vitsas, *Packet delay analysis of IEEE 802.11 MAC protocol*, Electronics Letters, Vol.39, No.18, pp.1358-1359, September 2003.

[24] P. Chatzimisios, A.C. Boucouvalas, V. Vitsas, *IEEE 802.11 packet delay-a finite retry limit analysis*, IEEE Global Telecommunications Conference (GLOBECOM'03), San Francisco, USA, December 2003.

[25] P. Chatzimisios, A.C. Boucouvalas, V. Vitsas, *Performance analysis of IEEE 802.11 DCF in presence of transmission errors*, IEEE International Conference on Communications (ICC'04), Paris, France, June 2004.

[26] P. Chatzimisios, A.C. Boucouvalas, V. Vitsas, *Performance analysis of the IEEE 802.11 MAC protocol for wireless LANs*, International Journal of Communication Systems, Vol.18, No.6, pp. 545-569, August 2005.

[27] P. Chatzimisios, A.C. Boucouvalas, V. Vitsas, *IEEE 802.11 Wireless LANs: Performance analysis and protocol refinement*, Eurasip Journal on Wireless Communications and Networking, Vol.2005, No.1, pp.67-78, March 2005.

[28] L. Chen, W.B. Heinzelman, *QoS-aware routing based on bandwidth estimation for mobile ad hoc networks*. IEEE Journal on Selected Areas in Communications, Vol.23, No.3, pp.561-572, March 2005.

[29] L. Cheng, A.G. Bourgeois, B.H. Yu, *Power Management in Wireless Ad Hoc Networks Using AODV*. Proceedings of the Sixth International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing and First ACIS International Workshop on Self-Assembling Wireless Networks. SNPD/SAWN 2005, Towson, USA, 2005.

[30] S.K. Cheung, J.L. van den Berg, R.J. Boucherie, R. Litjens, and F. Roijers, *An analytical packet/flow-level modelling approach for wireless LANs with quality-of-service support*, Proceedings of the 19th International Teletraffic Congress (ITC19), Beijing, China, September 2005.

[31] C.F. Chiasserini, M. Meo, *An innovative routing scheme for 802.11-based multi-hop networks*. Proceedings of the 60th IEEE Vehicular Technology Conference (VTC2004-Fall), Los Angeles, USA, 2004.

[32] K. Chin, J. Judge, A. Williams, R. Kermode, *Implementation experience with MANET routing protocols*, SIGCOMM Computer Communication Review, Vol.32, No.5, ACM Press, New York, NY, USA, 2002.

[33] E. Damasso, L.M. Correia (Editors), *Digital mobile radio: COST 231 view on the evolution towards 3rd generation systems*. European Commission - COST Telecommunications, Brussels, Belgium, 1998.

[34] D.S.J.D. De Couto, D. Aguayo, J. Bicket, R. Morris, *A high-throughput path metric for multi-hop wireless routing*. Proceedings of the 9th Annual International Conference on Mobile Computing and Networking (MobiCom 2003), San Diego, USA, 2003.

[35] J. Deng, R.S. Chang, *A Priority Scheme for IEEE 802.11 DCF Access Method*, IEICE Transactions on Communications, Vol.E82-B, No.1, pp. 96-102, 1999.

[36] R. Draves, J. Padhye, B. Zill, *Routing in multi-radio, multi-hop wireless mesh networks*. Proceedings of the 10th Annual International Conference on Mobile Computing and Networking (MobiCom 2004), Philadelphia, USA, 2004.

[37] R. Draves, J. Padhye, B. Zill, *Comparison of routing metrics for static multi-hop wireless networks*, Proceedings of the 2004 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM2004), New York, USA, 2004.

[38] R. Dube, C.D. Rais, K. Wang, S.K. Tripathi, *Signal stability-based adaptive routing (SSA) for ad hoc mobile networks*. IEEE Personal Communications Vol.4, No.1, pp.36-45, 1997.

[39] C.H. Foh, M. Zukerman, *Performance Analysis of the IEEE 802.11 MAC Protocol*, Proceedings of the European Wireless Conference (EW 2002), Florence, Italy, February 2002.

[40] M. Gerharz, C. de Waal, P. Martini, P. James, *Strategies for finding stable paths in mobile wireless ad hoc networks*. Proceedings of the 28th Annual IEEE International Conference on Local Computer Networks (LCN 2003), Bonn, Germany.

[41] T. Goff, N.B. Abu-Ghazaleh, D.S. Phatak, R. Kahvecioglu, *Preemptive routing in Ad Hoc networks*. Proceedings of the 7th Annual International Conference on Mobile Computing and Networking (MobiCom 2001), Rome, Italy, 2001.

[42] J. Gomez, A.T. Campbell, M. Naghshineh, C. Bisdikian, *Conserving transmission power in wireless ad hoc networks*. Proceedings of the 9th International Conference on Network Protocols (ICNP 2001), Riverside, USA, 2001.

[43] D. Gu, J. Zhang, *A new measurement-based admission control method for IEEE 802.11 wireless local area networks*, Proceedings of the 14th IEEE Conference on Personal, Indoor and Mobile Radio Communications (PIMRC 2003), Beijing, China, September 2003.

[44] M. Gudmundson, *Correlation model for shadowing fading in mobile radio systems*. Electronics Letters, Vol.27, No.23, pp.2145-2146, November 1991.

[45] P. Gupta, P.R. Kumar, *The Capacity of Wireless Networks*. IEEE Transactions on Information Theory, Vol.46, No.2, 2000.

[46] A. Gurtov, S. Floyd, *Modeling wireless links for transport protocols*, ACM SIGCOMM Computer Communication Review, Vol.34 , No.2, pp.85-96, April 2004.

[47] Z. Hadzi-Velkov, B. Spasenovski, *Saturation throughput: delay analysis of IEEE 802.11 DCF in fading channel*, Proceedings of IEEE International Conference on Communications (ICC'03), Seattle, USA, May, 2003.

[48] H. Hassanein, A. Zhou, *Routing with load balancing in wireless Ad hoc networks*. Proceedings of the 4th ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWIM 2001), Rome, Italy, 2001.

[49] J. He, Z. Tang, Z. Yang, W. Cheng, C.T. Chou, *Performance evaluation of distributed access scheme in error-prone channel*, Proceedings of the 2002 IEEE Region 10 Conference on Computers, Communications, Control and Power Engineering (TENCON'02), Beijing, China, October 2002.

[50] J. He, H.K. Pung, *Performance modelling and evaluation of IEEE 802.11 distributed coordination function in multihop wireless networks*, Computer Communications, Vol.29, No.9, May 2006.

[51] M. Heusse, F. Rousseau, G. Berger-Sabbatel, A. Duda, *Performance anomaly of 802.11b*, Proceedings of the Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2003), San Francisco, USA, April 2003.

[52] G. Holland, N. Vaidya, P. Bahl, *A Rate-Adaptive MAC Protocol for Multi-Hop Wireless Networks*. Proceedings of the 7th Annual International Conference on Mobile Computing and Networking (MobiCom 2001), Rome, Italy, 2001.

[53] H.Y. Hsieh, R. Sivakumar, *A Hybrid Network Model for Cellular Wireless Packet Data Networks*. Proceedings of the IEEE Global Telecommunications Conference (Globecom 2002), Taipei, Taiwan, 2002.

[54] Y.C. Hu, D.B. Johnson, *Design and demonstration of live audio and video over multihop wireless ad hoc networks*. Proceedings of the 2002 Military Communications Conference (MILCOM 2002), Anaheim, California, USA, 2002.

[55] Y.C. Hu, D.B. Johnson, *Exploiting congestion information in network and higher layer protocols in multihop wireless ad hoc networks*. Proceedings of the 24th International Conference on Distributed Computing Systems, Tokyo, Japan, 2004.

[56] IEEE Computer Society LAN/MAN Standards Committee, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, The Institute of Electrical and Electronic Engineers, IEEE Std. 802.11-1997, New York, USA, 1997.

[57] IEEE Computer Society LAN/MAN Standards Committee, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications - Amendment 1: High-Speed Physical Layer in the 5 GHz Band*, The Institute of Electrical and Electronic Engineers, IEEE Std. 802.11a-1999, New York, USA, 1999.

[58] IEEE Computer Society LAN/MAN Standards Committee, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications - Amendment 2: Higher-Speed Physical Layer (PHY) Extension in the 2.4 GHz Band*, The Institute of Electrical and Electronic Engineers, IEEE Std. 802.11b-1999, New York, USA, 1999.

[59] IEEE Computer Society LAN/MAN Standards Committee, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications - Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements*, The Institute of Electrical and Electronic Engineers, IEEE Std. 802.11e-2005, New York, USA, 2005.

[60] IEEE Computer Society LAN/MAN Standards Committee, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications - Amendment 4: Further Higher-Speed Physical Layer Extension in the 2.4 GHz Band*, The Institute of Electrical and Electronic Engineers, IEEE Std. 802.11g-2003, New York, USA, 2003.

[61] ILOG, *ILOG CPLEX 8.1 Reference Manual*. Available from http://www.ilog.com, 2004.

[62] Intersil Corporation, *Wireless LAN Integrated Medium Access Controller with Baseband Processor*, Data Sheet of the ISL3873, File Number 8015.2, USA, 2001.

[63] H. Jiang, J.J. Garcia-Luna-Aceves, *Performance comparison of three routing protocols for ad hoc networks*. Proceedings of the Tenth International Conference on Computer Communications and Networks (ICCCN 2001), Scottsdale, USA, 2001.

[64] S. Jiang, D. He, J. Rao, *A prediction-based link availability estimation for mobile ad hoc networks*. Proceedings of the Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2001), Anchorage, USA, 2001.

[65] J. Jun, P. Peddabachagari, M. Sichitiu, *Theoretical maximum throughput of IEEE 802.11 and its applications*, Proceedings of the Second IEEE International Symposium on Network Computing and Applications (NCA'03), Cambridge, USA, April 2003.

[66] E.-S. Jung, N.H. Vaidya, *A Power Control MAC Protocol for Ad Hoc Networks*. Wireless Networks, Vol.11, No.1-2, Springer Netherlands, 2005.

[67] A. Kamerman, L. Monteban, *WaveLAN-II: A high-performance wireless LAN for the unlicensed band.* Bell Labs Technical Journal, 1997.

[68] V. Kanodia, C. Li, A. Sabharwal, B. Sadeghi, E. Knightly, *Distributed Multi-Hop Scheduling and Medium Access with Delay and Throughput Constraints*, Proceedings of the 7th Annual International Conference on Mobile Computing and Networking (MOBICOM 2001), Rome, Italy, June 2001.

[69] S. Keshav, *The Packet Pair Flow Control Protocol*. Technical Report 91-028, International Computer Science Institute, Berkeley, USA, 1991.

[70] H. Kim, J.C. Hou, *Improving protocol capacity with model-based frame scheduling in IEEE 802.11-operated WLANs*, Proceedings of the 9th Annual International Conference on Mobile Computing and Networking (MOBICOM 2003) San Diego, USA, September 2003.

[71] P.C. Kokkinos, C.A. Papageorgiou, E.A. Varvarigos, *Energy-Aware Routing in Wireless Ad-Hoc Networks*. Proceedings of the Sixth IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks (WoWMoM 2005), Taormina, Italy, 2005.

[72] M. Kubisch, S. Mengesha, D. Hollos, H. Karl, A. Wolisz, *Applying ad-hoc relaying to improve capacity, energy efficiency, and immission in infrastructure-based WLANs*. Proceedings of Kommunikation in Verteilten Systemen (KiVS'03), Leipzig, Germany, 2003.

[73] A. Kumar, E. Altman, D. Miorandi, M. Goyal, *New insights from a fixed point analysis of single cell IEEE 802.11 WLANs*, Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2005), Miami, USA, March 2005.

[74] Y.-L. Kuo, C.-H. Lu, E.H.K. Wu, G.-H. Chen, *An admission control strategy for differentiated services in IEEE 802.11*, IEEE Global Telecommunications Conference (GLOBECOM'03), San Francisco, USA, December 2003.

[75] Y. Kwon, Y.Fang, H.Latchman, *Performance Analysis for a New Medium Access Control Protocol in Wireless LANs,* Wireless Networks, No.10, pp.519-529, 2004.

[76] E.L. Lawler, D.E. Wood, *Branch-and-bound method: A survey.* Operations Research, Vol.14, No.4, 1966.

[77] H. Lei, A.A. Nilsson, *An M/G/1 queue with bulk service model for power management in wireless LANs*, Proceedings of the 2nd ACM International Workshop on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks (PE-WASUN'05), Montreal, Canada, October 2005.

[78] B. Li, R. Battiti, *Performance Analysis of an Enhanced IEEE 802.11 Distributed Coordination Function Supporting Service Differentiation*, 4th COST 263 International Workshop on Quality of Future Internet Services (QoFIS 2003), Stockholm, Sweden, October 2003.

[79] Y. Li, H. Man, *Three load metrics for routing in ad hoc networks*. Proceedings of the 60th IEEE Vehicular Technology Conference (VTC2004-Fall), Los Angeles, USA, 2004.

[80] Y. Lin, H. Wu, R. Fan, S. Cheng *Modeling multiple TCP connections established between a busy server and many receivers*, IEEE Global Telecommunications Conference (GLOBECOM'03), San Francisco, USA, December 2003.

[81] L. Lin, H. Fu, W. Jia, *An efficient admission control for IEEE 802.11 networks based on throughput analyses of (un)saturated channel*, IEEE Global Telecommunications Conference (GLOBECOM'03), St. Louis, USA, December 2005.

[82] R. Litjens, F. Roijers, H. den Berg, R.J. Boucherie, M. Fleuren, *Performance analysis of Wireless LANs: An integrated packet/flow level approach*, Technical Report, Memorandum No. 1676, Department of Applied Mathematics, University of Twente, Netherlands, May 2003.

[83] Y. Lu, W. Wang, Y. Zhong, B. Bhargava, *Study of distance vector routing protocols for mobile ad hoc networks*. Proceedings of the First IEEE International Conference on Pervasive Computing and Communications (PerCom 2003), Dallas, USA, 2003.

[84] H. Luo, R. Ramjee, P. Sinha, L. Li, S. Lu, *UCAN: a unified cellular and ad-hoc network architecture*. Proceedings of the 9th Annual International Conference on Mobile Computing and Networking (Mobicom'03), San Diego, USA, 2003.

[85] H. Lundgren, E. Nordström, C. Tschudin, *Coping with communication gray zones in IEEE 802.11b based ad hoc networks*. Proceedings of the 5th ACM International Workshop on Wireless Mobile Multimedia (WOWMOM'02), New York, USA, 2002.

[86] S. Mengesha, H. Karl, *Relay Routing and Scheduling for Capacity Improvement in Cellular WLANs*. Proceedings of Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt'03), Sophia-Antipolis, France, 2003.

[87] S. Mengesha, H. Karl, A. Wolisz, *Capacity Increase of Multi-hop Cellular WLANs Exploiting Data Rate Adaptation and Frequency Recycling*. Proceedings of the Third Mediterranean Ad-Hoc Networking Workshop (MedHocNet 2004), Bodrum, Turkey, 2004.

[88] D. Miorandi, E. Altman, *On the effect of feedback traffic in IEEE 802.11 b WLANs*, Research Report, INRIA, Sophia Antipolis, France, August 2003.

[89] D. Miorandi, A.A. Kherani, E. Altman, *A queueing model for HTTP traffic over IEEE 802.11 WLANs,* Computer Networks, Vol.50, No.1, pp.63-79, January 2006.

[90] J.P. Monks, V. Bharghavan, W.M.W. Hwu, *A power controlled multiple access protocol for wireless packet networks,* Proceedings of the 20th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2001), Anchorage, USA, 2001.

[91] S. Narayanaswamy, V. Kawadia, R.S. Sreenivas, P.R. Kumar *Power control in ad-hoc networks: Theory, architecture, algorithm and implementation of the COMPOW protocol.* Proceedings of the European Wireless 2002, Florence, Italy, 2002.

[92] *ns (Network Simulator),* http://www.isi.edu/nsnam/ns/.

[93] M. Özdemir, A.B. McDonald, *An M/MMGI/1/K queuing model for IEEE 802.11 ad hoc networks,* Proceedings of the 1st ACM International Workshop on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks (PE-WASUN'04), Venice, Italy, October 2004.

[94] J. Padhye, V. Firoiu, D. Towsley, J. Kurose, *Modeling TCP Throughput: A Simple Model and its Empirical Validation,* Proceedings of the Annual Conference of the Association for Computing Machinery's Special Interest Group on Data Communication (ACM SIGCOMM'98), Vancouver, Canada, 1996.

[95] W. Pattara-atikom, S. Banerjee, P. Krishnamurthy, *Starvation Prevention and Quality of Service in Wireless LANs,* Proceedings of the 5th International Symposium on Wireless Personal Multimedia Communications (WPMC'02), Honolulu, USA, October 2002.

[96] S. Pilosof, R. Ramjee, D. Raz, Y. Shavitt, P.Sinha, *Understanding TCP fairness over Wireless LAN,* The 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2003), San Francisco, USA, 2003.

[97] D. Pong, T. Moors, *Call admission control for IEEE 802.11 contention access mechanism,* IEEE Global Telecommunications Conference (GLOBECOM'03), San Francisco, USA, December 2003.

[98] R.J. Punnoose, P.V. Nikitin, J. Broch, D.D. Stancil, *Optimizing wireless network protocols using real-time predictive propagation modeling.* Proceedings of the Radio and Wireless Conference (RAWCON 99), Denver, USA, 1999.

[99] M.B. Pursley, H.B. Russell, J.S. Wysocarski, *Energy-efficient transmission and routing protocols for wireless multiple-hop networks and spread-spectrum radios,* Proceedings of the IEEE/AFCEA EUROCOMM 2000 - Information Systems for Enhanced Public Safety and Security, Munich, Germany, May 2000.

[100] D. Qiao, K.G. Shin, *Achieving efficient channel utilization and weighted fairness for data communications in IEEE 802.11 WLAN under the DCF*, Proceedings of the Tenth IEEE International Workshop on Quality of Service (IWQoS 2002), Miami Beach, USA, May 2002.

[101] T.S. Rappaport, *Wireless Communications: Principles and Practice,* Prentice Hall, USA, 1996.

[102] L. Romdhani, C. Bonnet, *Energy consumption speed-based routing for mobile ad hoc networks*. Proceedings of the 24th International Conference on Distributed Computing Systems, Tokyo, Japan, 2004.

[103] S.M. Senouci, G. Pujolle, *Energy efficient routing in wireless ad hoc networks*. Proceedings of the IEEE International Conference on Communications (ICC2004), Paris, France, 2004.

[104] S.H. Shah, K. Chen, K. Nahrstedt, *Dynamic Bandwidth Management for Single-hop Ad Hoc Wireless Networks,* Proceedings of the IEEE Internatioanl Conference on Pervasive Computing and Communication, Dallas, USA, March 2003.

[105] J. Song, V. Wong, V.C.M. Leung, *Load-aware on-demand routing (LAOR) protocol for mobile ad hoc networks*. Proceedings of the 57th IEEE Semiannual Vehicular Technology Conference (VTC 2003-Spring), Jeju , Korea, 2003.

[106] J.A. Stevens, *Spatial reuse through dynamic power and routing control in common-channel random-access packet radio networks*. Ph.D. Thesis, University of Texas at Dallas, USA, 1988.

[107] Y.C. Tay, K.C. Chua, *A Capacity Analysis for the IEEE 802.11 MAC Protocol*, Wireless Networks, Vol.7, No.2, pp.159-171, 2001.

[108] O. Tickoo, B. Sikdar, *A queueing model for finite load IEEE 802.11 random access MAC,* IEEE International Conference on Communications (ICC'04), Paris, France, June 2004.

[109] O. Tickoo, B. Sikdar, *Queueing analysis and delay mitigation in IEEE 802.11 random access MAC based wireless networks,* Proceedings of the 23rd Conference of the IEEE Communications Society (INFOCOM 2004), Hong Kong, China, March 2004.

[110] C. Toh, *A novel distributed routing protocol to support ad-hoc mobile computing*. Proceedings of the Fifteenth IEEE Annual International Phoenix Conference on Computers and Communications, Scottsdale, USA, 1996.

[111] N.H. Vaidya, P. Bahl, S. Gupta, *Distributed Fair Scheduling in a Wireless LAN*, Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (MOBICOM 2000), Boston, USA, August 2000.

[112] A. Veres, A.T. Campbell, M. Barry, L.H. Sun, *Supporting service differentiation in wireless packet networks using distributed control*, IEEE Journal on Selected Areas in Communications, Vol.19, No.10, pp.2081-2093, October 2001.

[113] V.M. Vishnevsky, A.I. Lyakhov, *IEEE 802.11 Wireless LAN: Saturation Throughput Analysis with Seizing Effect Consideration*, Cluster Computing, Vol.5, No.2, pp.133-144, April 2002.

[114] V. Vitsas, P. Chatzimisios, A.C. Boucouvalas, P. Raptis, K. Paparrizos, D. Kleftouris, *Enhancing performance of the IEEE 802.11 distributed coordination function via packet bursting*, IEEE Global Telecommunications Conference Workshops (GlobeCom'04 Workshops), Dallas, USA, December 2004.

[115] B. Walke, *Mobile Radio Networks*, Wiley & Sons Ltd., 2nd ed., New York, USA, 2001.

[116] S.Y. Wang, J.Y. Liu, C.C. Huang, M.Y. Kao, Y.H. Li, *Signal Strength-Based Routing Protocol for Mobile Ad Hoc Networks*. Proceedings of the IEEE 19th International Conference on Advanced Information Networking and Applications (AINA 2005), Taiwan, March 2005.

[117] H. Wei, R.D. Gitlin, *Two-hop-relay architecture for next-generation WWAN/WLAN integration*. IEEE Wireless Communications, Vol.11, No.2, pp.24-30, 2004.

[118] J. Weinmiller, A. Festag, M. Schlager, A. Wolisz, *Performance Study of Access Control in Wireless LANs - IEEE802.11 DFWMAC and ETSI RES 10 HIPERLAN*, Mobile Networks and Applications Journal, MONET, Vol.2, 1997.

[119] K. Wu, J. Harms, *Load-sensitive routing for mobile ad hoc networks*. Proceedings of the Tenth International Conference on Computer Communications and Networks (ICCCN 2001), Scottsdale, USA, 2001.

[120] H. Wu, Y. Peng, K. Long, S. Cheng, J. Ma, *Performance of reliable transport protocol over IEEE 802.11 wireless LAN: analysis and enhancement*, Proceedings of the Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2002), New York, USA, 2002.

[121] Y. Xiao, J. Rosdahl, *Throughput and delay limits of IEEE 802.11*, IEEE Communications Letters, Vol.6, No.8, pp.355-357, August 2002.

[122] Y. Xiao, *Performance analysis of IEEE 802.11 e EDCF under saturation condition*, IEEE International Conference on Communications (ICC'04), Paris, France, June 2004.

[123] Y. Xiao, H. Li, *Evaluation of distributed admission control for the IEEE 802.11e EDCA*, IEEE Communications Magazine, Vol.42, No.9, pp.20-24, 2004.

[124] Y. Xiao, H. Li, S. Choi, *Protection and guarantee for voice and video traffic in IEEE 802.11e wireless LANs*, Proceedings of the 23rd Conference of the IEEE Communications Society (INFOCOM 2004), Hong Kong, China, March 2004.

[125] Y. Yang, R. Kravets, *Throughput guarantees for multipriority traffic in ad hoc networks*, Proceedings of the IEEE International Conference on Mobile Ad-hoc and Sensor Systems (MASS 2004), Fort Lauderdale, October 2004.

[126] J. Yin, X. Wang, D.P. Agrawal, *Modeling and optimization of wireless local area network* , Computer Communications, Vol.28, No.10, pp.1204-1213, June 2005.

[127] H. Zhai, X. Chen, Y. Fang, *A call admission and rate control scheme for multimedia support over IEEE 802.11 wireless LANs*, Wireless Networks, Vol.12, No.4, pp.451-463, August, 2006.

[128] L. Zhang, S. Zeadally, *HARMONICA: enhanced QoS support with admission control for IEEE 802.11 contention-based access*, Proceedings of the 10th IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS'04), Toronto, Canada, May 2004.

[129] C. Zhu, M.S. Corson, *QoS routing for mobile ad hoc networks*. Proceedings of the Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2002), New York, USA 2002.

[130] E. Ziouva, T. Antonakopoulos, *CSMA/CA performance under high traffic conditions: throughput and delay analysis*, Computer Communications, Vol.25, No.3, pp.313-321, February 2002.

Lebenslauf

Name:	Carsten Burmeister
Geburtsdatum :	21.05.1973
Geburtsort:	Hamburg
1980–1983	Grundschule Fleestedt, Seevetal
1983–1992	Alexander von Humboldt Gymnasium, Hamburg
	Abschluss: Abitur
1992–1993	Zivildienst
1993–1999	Studium an der Technischen Universität Hamburg-Harburg
	Abschluss: Diplom-Ingenieur Elektrotechnik
1999–2002	Research Engineer
	Panasonic Research and Development Center Germany, Langen
2002	Stanardisierungsingenieur
	Mobilcom Multimedia GmbH, Büdelsdorf
2002–2006	Wissenschaftlicher Mitarbeiter
	Technische Universität Hamburg-Harburg, Institut für Kommunikationsnetze
2007–heute	Entwicklungsingenieur
	Lufthansa Technik AG, Hamburg