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.

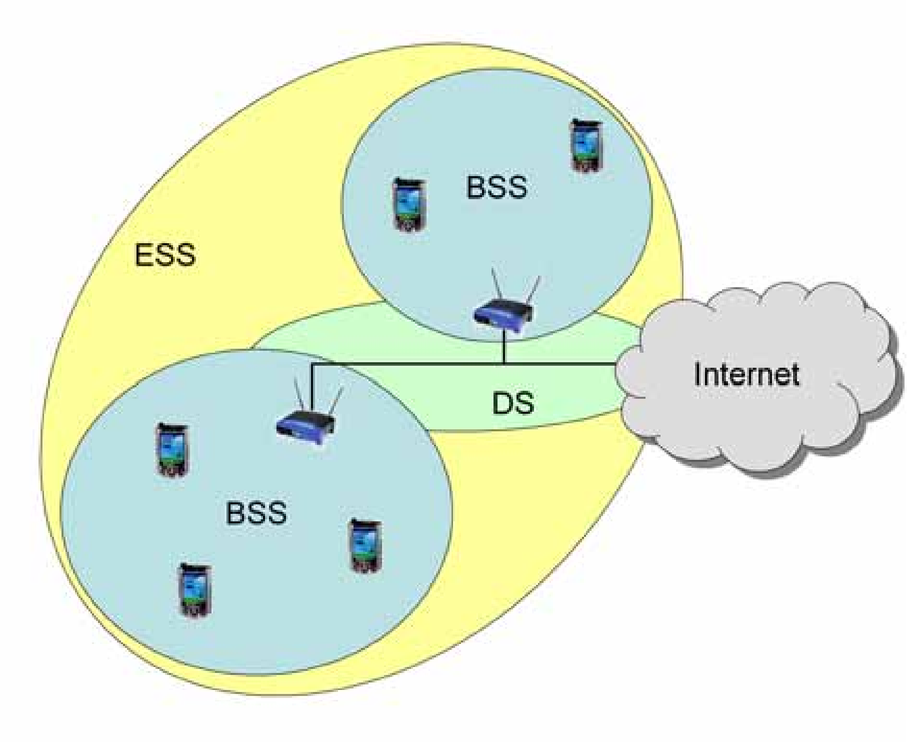

Figure 2.1: IEEE 802.11 architecture overview.

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.