Scheduling of heterogeneous data streams in the downlink of a dynamic OFDM-FDMA wireless cell

Magisterarbeit
am Fachgebiet Telekommunikationsnetze
Prof. Dr.-Ing. A. Wolisz
Institut für Telekommunikationssysteme
Fakultät IV Elektrotechnik und Informatik
Technische Universität Berlin

vorgelegt von
Stefan Valentin
am
9. September 2004

Betreuer: Prof. Dr.-Ing. A. Wolisz
Dipl.-Ing. J. Groß

Stefan Valentin
[redacted]

Abstract

In this thesis the performance of a combined link- and physical layer optimization approach [21] is studied for the downlink of a wireless OFDM-FDMA transmission system. This is done by discrete event simulation assuming a *heterogeneous* traffic load scenario, where Web-pages and MPEG-4 coded VBR video streams are transmitted simultaneously using the related transport layer protocols TCP and UDP. For the transmission of a Web-page performance increases of up to 100% were found.

Based on this optimization approach a new scheduling scheme for the heterogeneous traffic load scenario is proposed. It provides the adaption of the bit-rates of the transmitted video streams to the estimated load of the wireless system. In the performance study it is shown, that with a high proportion of video traffic in the cell this approach accelerates the downlink of a Web-page by up to 33% on the average.

Zusammenfassung

In dieser Magisterarbeit wird die Leistung eines Optimierungsverfahrens für den drahtlosen Empfang von Datenströmen unter Voraussetzung des OFDM-FDMA Übertragungsverfahrens untersucht. Dabei handelt es sich um einen kombinierten Ansatz der Informationen aus der Bitübertragungs- und der Sicherungsschicht berücksichtigt. Die Untersuchung erfolgt für ein Szenario mit *heterogener* Verkehrslast in dem die gleichzeitige Übertragung von Web-Seiten, mithilfe des TCP Protokolls, und von MPEG-4 kodierten Video-Strömen, über UDP, angenommen wird. Für dieses Szenario können für die Übertragung einer Web-seite mit dem kombinierten Ansatz Leistungsgewinne von bis zu 100% erreicht werden.

Basierend auf dem o.g. Ansatz wird im Anschluss an dessen Untersuchung ein neues *scheduling*- Verfahren für heterogene Verkehrslasten vorgestellt. Dieses ermöglicht die Adaption der Bit-raten der übertragenen Video-Ströme an die geschätzte Last. Die Untersuchung des Verfahrens zeigt, dass für Szenarien mit hohem Anteil an Video-Strömen die Übertragung einer Web-Seite im Durchschnitt um bis zu 33% beschleunigt werden kann.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The mobility provided by the transmission of speech and data without the need of a cable connection has lead to the wide integration of mobile telephones and computers in every-days life. This has opened a wide field for research and technology, resulting in new developments such as the recently released third generation of mobile telephone networks (UMTS) and Wireless Local Area Networks (WLAN) whose data rate converges more and more to the most widespread standard for cable networks, the 100 MBps Ethernet. With the increase in transmission speed and quality, provided by these techniques new interesting services, e.g. video telephony or the ubiquitous usage of the Internet can be provided to the mobile user.

However, all these benefits add a certain cost to the wireless system. Unlike cable connections, wireless systems suffer from a high variance of the delay and the attenuation of the transmitted signals. As a result the transmission quality permanently varies over time and frequency. This effect increases with higher data rates and mobility. In order to assure the fault-tolerant transmission, while providing high data rates Orthogonal Frequency Division Multiplexing (OFDM) has been developed. OFDM was designed to minimize the distortion of the signal caused by the delay variance. The spectrum is separated into small sub-bands, the so-called *subcarriers*, which are placed in order to assure a high utilization of the given bandwidth.

However, the attenuation to the transmitted signal still varies according to the given speed and position of a wireless terminal (WT). As a consequence a high error rate has to be taken into account, which will increase with the transmission speed. For a static system this results in the under-utilization if both, attenuation and error rate are low. This has lead to the development of *channel-state* dependent scheduling policies. Here the scheduler decides about the distribution of the resource according to the actual capacity of the wireless channel resulting from the actual measured Signal-to-Noise Ratio (SNR). With OFDM the SNR of each subcarrier regarding to a specific WT can be measured. Supposing an Frequency Division Multiple Access (FDMA) scheme each subcarrier can be allocated to those WT for which the best SNR value was detected. Recent research has shown, that this OFDM-FDMA approach highly increases the performance of the system [19].

In addition to the wireless channel another source of variability is the traffic load. To avoid the under-utilization of the channel, here it is beneficial to allocate the resources according to the individual bandwidth requirements of each terminal. Therefore, dynamic multiple access schemes are employed. In combination with OFDM-FDMA the dynamic multiple access scheme provides the adaption to the varying channel as well as to fluctuations in traffic load. Using this so-called *dynamic OFDM-FDMA* [24] approach high performance gains can be achieved compared to a static scheme [21].

For the found improvements of the transmission system it is interesting which fraction of the gain finally reaches the user level. Here the performance investigation on the application layer is much closer to the users perception. For dynamic OFDM-FDMA systems this was done assuming VBR video streams [21] and Web-pages [39]. Both types of traffic have different Quality of Service (QoS) demands to the transmission system and are typically transmitted simultaneously in packet switched networks.

However, a performance study considering only the transmission of a single type of data does not represent the nature of the traffic in the Internet. This has been shown by recent measurements on Internet backbones [17], where a traffic mixture was found. While the main proportion (up to 90%) of this mixture was related to Web-pages, streaming media was measured to be the third most commonly used service.

For the application of a dynamic OFDM-FDMA system for mobile networks with access to the Internet it is interesting how the system performs concerning this *heterogeneous* nature of the traffic. This is investigated in the first performance study given in Chapter 5 of this thesis. Therefore, a traffic load scenario is considered, including both, the streaming of VBR video and the simultaneous transmission of Web-pages. For both types of traffic the performance of several combinations of the dynamic OFDM-FDMA scheduling scheme is studied on the application layer.

The second subject of this thesis is to propose an extension to the dynamic OFDM-FDMA system: the *adaptive Video Queue Management (VQM)*. The design of this new scheme takes heterogeneous traffic into account and provides the adaption of the media streams to the varying system resources. This is known under the topics of *dynamic QoS* and *adaptive filtering* [65]. The basic idea of this adaption strategy is the manipulation of the video data according to the estimated system load. This scheme may be employed especially in dynamic OFDM-FDMA systems, since it frees capacities if congestion occurs due to low channel states. While the adaptive VQM is located on the link layer, application layer related information is extracted from the video streams and considered during the filtering process. After the approach and the algorithm of this scheme are discussed in Chapter 6, the resulting performance is investigated. In Chapter 7 this is done by simulation for several combinations of dynamic and static OFDM-FDMA schemes and for several heterogeneous traffic load scenarios. Beneath the achieved performance increase for the transmission of Web-pages and MPEG-4 coded video streams a further focus is to analyze the adaption capabilities and the influence of the adaptive filtering to the video streams.

In the following chapter a short overview to the components of a wireless communication system is given. Chapter 3 introduces aspects of heterogeneous traffic and the related transport layer protocols. Chapter 4 includes an introduction to scheduling in wireless networks, the problem description of this thesis and a specification of the dynamic OFDM-FDMA schemes. These chapters are followed by the discussion of the proposed video queue management scheme in Chapter 6 and the performance studies in Chapter 5 and 7.

Chapter 2

Basics of wireless communication

Due to the enormous growth rates of mobile telephones and mobile computers during the last decade wireless communication systems have become more and more important. Mobile telephones rule the market as no other communication device with growth rates of over 50% per year [53]. Beneath the transmission of speech the wireless transmission of data is now familiar to everyone who uses a mobile telephone or computer. Since the 802.11b specification for WLAN was published in July 1999, the wireless transmission of digital data with high speed became feasible for private personal communication and was widely adopted. The 802.11b specification was followed by a variety of standards and implementations, e.g. the WLAN specifications 802.11a [28] and 802.11g [29] which both provide up to 5 times higher data rates than 802.11b.

The usage of the wireless channel leads to consequences which are fundamental different from the transmission of data via a cable link:

- While the delay and the attenuation the channel causes to a signal is constant on a cable link a random variation is quite normal with radio channels.
- Movement of the receivers highly increases this variability

The effects causing these problems are mostly of statistical nature and do superpose. This complicates the design of a wireless communication system, since the varying channel leads to high error rates (high attenuation) or under-utilization of bandwidth (low attenuation). In Section 2.2 the influence of these effects on the radio channel and aspects of channel modeling are discussed, while in Section 2.3 and Section 2.4 transmission techniques are presented which address especially the problems of a wireless channel.

Unlike a cable connection the electromagnetic spectrum of the air is a resource which is shared by all radio transmitters. For wireless communication systems this leads to the following specifics:

- The signal of a transmitter is received by all terminals within a distinct propagation radius (depending on the signals power loss). Connecting specific terminals as with switching in cable networks is not possible in a wireless network.

- The used frequency is blocked until the transmission ends. If any other transmitter within the propagation radius sends, both signals are distorted due to interference.

- If more performance is needed in a cable network it can be obtained by using more or better cables. With wireless transmission more bandwidth is hard to obtain. Here it has to be assured by international institutions, such as the *ITU Radio Regulations Board*, that no interference occurs with other systems, e.g. with the radiotelephony of the local airport ground control.

In order to provide a good channel utilization and avoid blocking the limited amount of shared bandwidth should be optimal distributed to the terminals. This problem is discussed in Section 2.5 where multiple access schemes are introduced. Due to the highly limited amount of free capacities in the electromagnetic spectrum [46] it is hard to obtain higher amounts of bandwidth if needed. This lead to the development of cellular systems which will be introduced in the next section.

2.1 Introduction to cellular wireless systems

A traditional radio system uses a high powered transmitter and a tall antenna to cover a large area. For communication systems this means that each connection allocates two carrier frequencies, one for the uplink and one for the downlink, in this large area. With more users this will soon result into a shortage of carrier frequencies.

In cellular systems this problem is solved by separating the covered area into small cells as pictured in Figure 2.1. The transmission within each cell is controlled by one access point. Given that the access point covers only a single cell it sends with much lower power than the transmitter in a traditional system. As denoted by the letters in Figure 2.1 different frequencies are used by neighboring cells in order to prevent interference at the cell boundaries.

Beneath the access point the wireless cell contains a varying number of J Wireless Terminals (WTs). A wireless terminal is a device which is needed by the mobile user to access the wireless network, e.g. a cellular telephone or a mobile computer with a WLAN adaptor. While the position of the access point is fixed the WTs

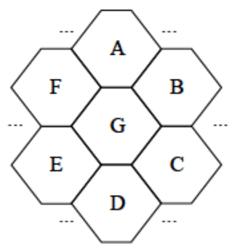


Figure 2.1: Cellular frequency usage concept [53]: Different letters stand for different frequencies which are reused in each set of cells.

may roam in the cell. As long as one WT is located in the cell the access point controls the transmission from and to the terminal via the radio channel. This includes communication between the WTs within the same cell and from or to terminals located in external networks or cells. For this purpose the access point is connected via a cable link to a wired network. An example setup of a single cell is illustrated in Figure 2.2.

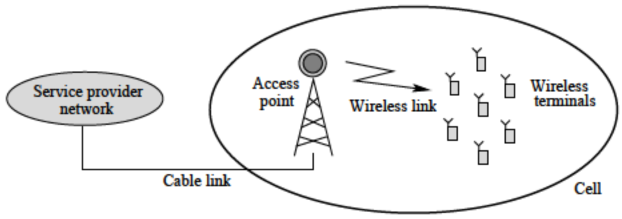


Figure 2.2: Example setup of a single cell with $J = 7$ WTs

The cellular concept is fundamental for all popular terrestrial mobile telephone networks, for example GSM and UMTS [27]. Most WLAN implementations, e.g. the widely used 802.11b and the new 802.11g support the so-called infrastructure mode [47], which is similar to a single cell of a cellular telephone network. However, here the typically application scenario is a *hot spot*, which means that no transparent roaming between the cells is supported.

2.2 Radio channel

Unlike wired channels which show a static and predictable behavior radio channels are extremely random and therefore hard to analyze and model [53]. During its propagation the electromagnetic wave suffers from several effects which are typically summarized using the terms *path-loss*, *shadowing*, and *fading* [1]. These effects and the influence to the transmission system are discussed in this section, together with aspects of channel modeling.

With wireless channels the transmission of data results into the propagation of electromagnetic waves from the sender to the receiver antenna. During the propagation the power of the signal is attenuated. This deterministic effect is known as path loss [1]. Supposing the transmission power P_{tx}, the received power can be characterized by

$$P_{rx} = P_{tx} \times K \frac{1}{d^{\alpha}} \tag{2.1}$$

where K is the reference loss, which occurs over a reference distance and the path loss exponent α characterizes how severe a distance increase attenuates the received signal. Both values α and K depend highly on the chosen environment.

In addition to the loss of power terrestrial transmitted electromagnetic waves are scattered and reflected on fixed or moving obstacles, like buildings, trees or cars. This leads to *multi-path propagation*, which means, that multiple versions of the transmitted signal reach the receiver antenna on different paths. Each version

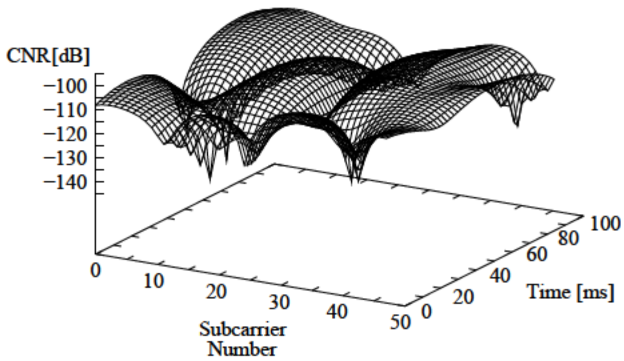


Figure 2.3: Example CNR behavior of a frequency selective and time variant radio channel

has a specific phase, delay (τ) and, resulting from the attenuation (a) due to path loss, a specific amplitude. These waves, called *multi-path waves* superpose at the receiver antenna to a signal which varies widely in amplitude and phase. This effect is described by the term *fading* or *small-scale fading* [53]. Fading leads to drastic stochastic changes of the signal strength and is frequency selective as long as no movement occurs in the propagation environment. With movement of the sender, the receiver or the scatterer the fading becomes time variant [7]. Therefore moving WTs result into the time and frequency variant attenuation $a(f, t)$. With the noise power n this can be expressed in the Channel Gain-to-Noise Ratio (CNR) as in

$$G(f,t) = \frac{a^2(f,t)}{n^2(f,t)} \tag{2.2}$$

which leads together with the transmission power P_{tx} on a specific frequency to the SNR as in Equation 2.3.

$$Q(f,t) = P_{tx}(f,t) \times G(f,t) \tag{2.3}$$

The rapid variation of the CNR in time and frequency due to fading is exemplarily shown in Figure 2.3. The frequency domain is separated into distinct sub-bands called subcarriers. Considering moving WTs the CNR varies independently for each WT.

One possibility to model the stochastic nature of fading is discussed in [1]. Here the time selective behavior of fading is characterized by a *Jakes-like* power spectral density function, while the frequency selective fading is modeled using an exponential power delay profile.

In addition to path loss and fading further propagation phenomena, namely reflection, diffraction (e.g. around buildings), refraction (e.g. through windows and walls), scattering and absorption (e.g. on parks) are summarized by *shadowing*. As fading shadowing influences the signal strength randomly. It can be modeled by the use of a *log normal* distributed probability density function for the attenuation factor with a mean of 0 dB and a standard deviation depending on the considered environment [1]. While shadowing and path loss typically cause a time variance in the scale of a second, fading leads to the most drastic attenuation changes in the time scale of milliseconds or even microseconds [7]. The time scale of fading depends highly on the WT speed. With lower speed the time the CNR roughly stays constant at a given frequency (*coherence time*) is longer than with higher speed [1]. Therefore, the maximum WT speed (v_{max}) is an important design parameter for wireless networks. Another interesting parameter for the severeness of multi-path propagation and the resulting effects is the difference between the delay on the longest and the shortest path

$$\Delta\tau = \tau_{max} - \tau_{min}, \tag{2.4}$$

which is called delay spread. The first version of the signal is typically received on the direct path with a delay of τ_{min} after transmission. It is followed by further reflections until the last reflection is received after τ_{max}. Due to fading different shaped signals combine at the receiver antenna. If each signal is interpreted as a sequence of digital symbols, with a duration of T_s per symbol, the symbols start to overlap if $\Delta\tau$ exceeds T_s. The varying delay spread causes a smearing of the symbols, called Inter-Symbol Interference (ISI) [53], starting with the symbols of the direct and the longest path. With higher $\Delta\tau$ in comparison to the symbol time T_s more and more symbols are affected. Since the mean values for the delay spread can be considered in the design of a system, the standard deviation ($\sigma\Delta\tau$) is the interesting parameter here to express the severeness of multi-path propagation.

2.3 Multi-carrier modulation and OFDM

As mentioned in the preceding section Inter-Symbol Interference occurs at the receiver antenna if the delay-spread exceeds the symbol time. Then the received symbols get more and more distorted if the unbalance between these two parameters rises. This can be caused by strong fading in the propagation environment, resulting in high variations of $\Delta\tau$, but also smaller symbol times will lead to ISI if they drop below the actual $\Delta\tau$. However, with a limited bandwidth sending more symbols within a fixed time interval is the key for the faster transmission of data. With ISI rising symbol-rates result into higher probabilities for the distortion of symbols, which then has to be compensated by equalization. Thus, the complexity and cost of the equalizer rises with the symbol-rate.

To provide higher data rates with practicable equalization complexity Multi Carrier Modulation (MCM) can be applied. Here a broad-band channel is separated into S narrow-band subchannels. The symbols are modulated and transmitted in parallel on the S carrier frequencies of the subcarriers. For the same data rate this leads to a symbol time, which is S times larger than with a Single Carrier Modulation (SCM) scheme, where all symbols are sent sequentially using the full bandwidth. The larger symbol times lower the impact of ISI significantly.

Orthogonal Frequency Division Multiplexing (OFDM) is a specific form of a Multi Carrier Modulation (MCM) scheme. First described in 1966 [8] its implementation complexity became feasible during the last decade due to the progress of

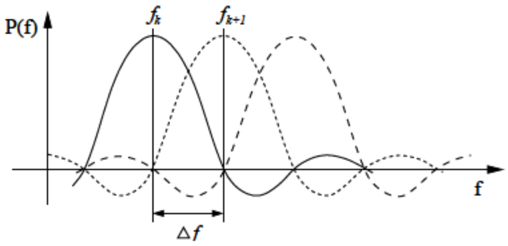


Figure 2.4: Illustration of the power spectrum of three OFDM subcarriers

research and development in signal processing. OFDM is now widely adopted. It is applied for digital audio [13] and video broadcasting [10] as well as for WLAN such as IEEE 802.11a [28], IEEE 802.11g [29] and HIPERLAN/2 [12]. Furthermore it is discussed for new Ultra Wideband (UWB) standards for wireless personal networks in the IEEE work group 802.15.3a [49].

The primarily benefit of OFDM is the high spectral efficiency, which is achieved by the close, overlapping placement of the subcarriers. With OFDM the maximum of the power spectrum for each subcarrier is placed on the zero-crossings of the adjacent spectra. The orthogonal placement of the subcarriers requires a certain distance (Δf) between the carrier frequencies. With the symbol time this leads to a distance of

$$\Delta f = \frac{1}{T_s}. \tag{2.5}$$

Since Δf is fixed with Equation 2.5 the symbol time and therefore the baud rate is constant and equal for all the subcarriers. In Figure 2.4 the orthogonal spacing of the subcarrier power spectra is illustrated. The orthogonal spacing avoids crosstalk, also known as Inter-carrier Interference (ICI), between the subcarriers and saves bandwidth since no guard bands have to be applied. Although the spectra do overlap no crosstalk occurs until orthogonality between the subcarriers is assured, which means, that the carrier frequencies are uncorrelated.

Analytically the power spectrum of a single subcarrier k can be given as the *sinc* function

$$P(f) = T_s \operatorname{sinc}(\pi T_s (f - k\Delta f)) \tag{2.6}$$

where T_s denotes the symbol time. Before the transmission a guard interval is added in front of each symbol to remove ISI completely. This extends the symbol time to $T_b = T_s + T_g$ which denotes the length of the resulting so-called *OFDM-block*. In Figure 2.5 the structure of an OFDM transmitter is illustrated. Per transmission cycle $S \times M$ Bits are modulated on S symbols which are sent simultaneously. Depending on the applied modulation technique a symbol represents M bits and has a specific shape, denoted by $d(t)$ in the time domain. Then each symbol is individually filtered with a rectangular impulse response $g(t)$ of length T_b. In the next step the filtered symbols are modulated to S carrier frequencies, which results in

$$x(t) = \sum_{s=0}^{S-1} (d_s(t) * g(t)) e^{2\pi j f_s t} \tag{2.7}$$

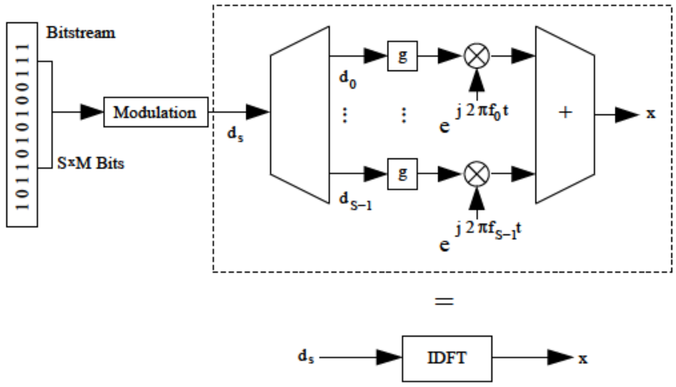


Figure 2.5: Basic structure of the OFDM transmission on the sender side and interpretation as the inverse DFT

for the signal for all subcarriers at the transmitters output. In the frequency domain the rectangular pulse shape resulting from the filtering is represented by the *sinc* function as in Equation 2.6. It can be shown [36], that with discrete signal processing (one OFDM-block is then represented by S samples) the described filtering and modulation process can be realized by applying the inverse Discrete Fourier Transform (DFT) on the S simultaneously processed OFDM-blocks. At the receivers side the inverse procedure is employed using the DFT. To lower the calculation complexity sender and receiver are typically implemented using the Fast Fourier Transform (FFT) algorithm.

Finally, this results into the following benefits which can be achieved by the employment of OFDM:

- Low equalization complexity combined with a insignificant impact of ISI, while providing high data rates
- High spectral efficiency, no guard bands needed, while avoiding crosstalk
- Easy implementation due to the applicable FFT and the symmetric structure of sender and receiver

2.4 Adaptive modulation

The modulation technique defines how an input signal is combined with a carrier frequency to a signal called symbol. In the context of digital data transmission the input signal represents a bit. The number of bits which can be represented by a single symbol depends on the modulation technique. Since higher numbers of bits per symbol result in higher data rates the choice of the best possible modulation technique is important.

One criteria is the expected CNR. If the CNR drops, then the probability (P_s) of a symbol error for a certain modulation technique rises. In this case the Symbol Error Probability (SEP) for higher modulation techniques starts to rise earlier

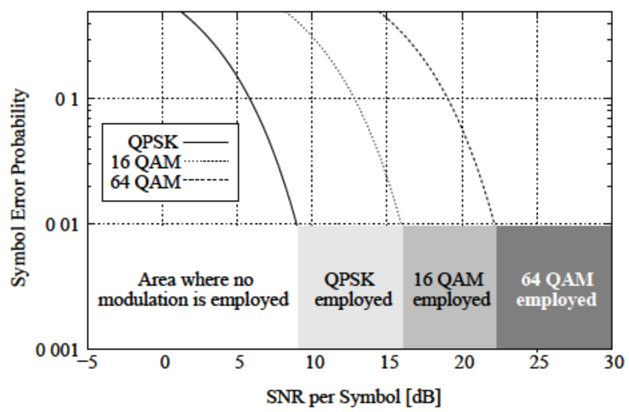


Figure 2.6: Symbol error probability and example SNR modulation ranges as applied for a maximum acceptable P_s of 0.01

than for those techniques representing fewer amounts of bits (Figure 2.6). Thus, with lower CNR one would choose a lower modulation technique to stay within appreciable boundaries for the symbol errors. Otherwise, with a high CNR one can employ a powerful modulation technique, which allows the sender to transmit more bits per subcarrier during T_s.

Due to the random nature of the radio channel choosing one modulation technique permanently will result in high symbol error rates or in under-utilized capacity, depending on the actual CNR. Here it is profitable to adapt the chosen modulation technique to the periodically measured CNR or SNR. The adaption can be done by bit-loading algorithms [37], where also the transmission power is considered. In another approach the transmission power is kept constant on each subcarrier for any modulation technique [23]. On each WT the SNR value $q_{j,s}(t)$ is measured periodically per OFDM subcarrier. Based on this value the highest modulation technique is chosen, which is able to provide a SEP below a certain threshold. Basically, this is equal to assigning each modulation technique a range of SNR values for which this modulation technique will be used. An example of these ranges for a threshold of $P_s < 0.01$ and a transmission power of -7 dBm is shown for several modulation techniques in Figure 2.6.

For each subcarrier this adaption results in different amounts of bits per symbol which may change every adaption cycle. The assigned amounts of bits are called *channel states* and lead to a variable bit-rate per subcarrier while a fixed baud-rate is used for all subcarriers. The added flexibility due to adaptive modulation shows another benefit of separating the whole bandwidth into subcarriers using a MCM scheme. An further improvement on the link layer is discussed under the topic of dynamic FDMA in the following section.

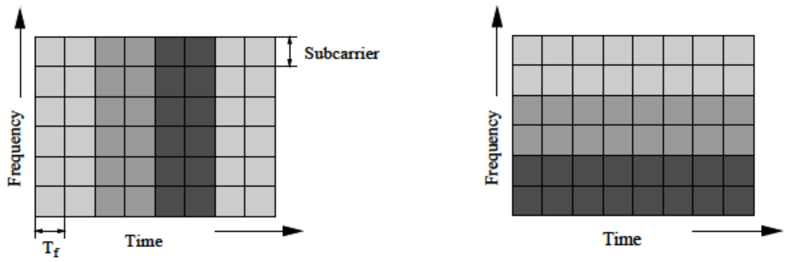
Modulation Technique	Bits per Symbol	SNR Range
No modulation applied	0	$0 \leq q < 4$ dB
BPSK	1	$4 \leq q < 9$ dB
QPSK	2	$9 \leq q < 16$ dB
16-QAM	4	$16 \leq q < 22$ dB
64-QAM	6	$22 \leq q < 28$ dB
256-QAM	8	$28 \leq q < \infty$ dB

Table 2.1: Modulation techniques, the represented number of bits per symbol and the applied SNR ranges to assure a Symbol Error Probability < 0.01 [52] with $P_{tx} = -7$ dBm.

2.5 Multiple access schemes

To provide the capacity of the wireless link to several WTs simultaneously a multiple access scheme is applied on top of the physical layer. The task of a multiple access scheme is to allocate resources of the physical layer to the clients, e.g. the WTs. Here three dimensions can be considered: time, frequency (i.e. subcarriers) and code. If Time Division Multiple Access (TDMA) is employed all subcarriers are assigned to one WT for a limited time interval. If the time is slotted into so-called Medium Access Control (MAC) frames this can be done for each frame of length T_f. This is illustrated in Figure 2.7(a) where the color of a single block denotes one OFDM subcarrier obtained by a specific WT. In contrast to TDMA with FDMA each WT obtains only a subset of N subcarriers (Figure 2.7(b)), however, for the whole transmission time. The third possibility is Code Division Multiple Access (CDMA) where every terminal receives a so-called chipping code which spreads the transmitted signal over the full spectrum. Due to the orthogonality the code causes to the sender signal, every terminal is able to transmit simultaneously. Since it is not discussed in further parts of this thesis CDMA is not illustrated here.

The illustrated static TDMA and FDMA are fundamental concepts in telecommunications. For example FDMA is familiar to anyone who has ever used a radio or television set [57]. Static TDMA is widely used for multiplexed digitized voice or data streams as with DS-1 or ISDN [57]. The problem with this static approach is, that resources are always assigned. The multiple access scheme does not consider whether these resources are needed at the terminal or not. If, for example, one terminal has no data to sent it should be not taken into account during the allocation process. Otherwise this results into an under-utilization of the channel and therefore into a lack of performance. Dealing with this problem requires more complex so-called statistical or dynamic multiple access schemes [59]. Here the multiple access scheme decides about the amount of resources allocated to the terminals. In this thesis the term *scheduler* is used for the entity which makes these decisions and *scheduling policy* denotes the way how this is done. In Figure 2.8 an example is illustrated, where the scheduling policy prioritizes the terminals, e.g. according

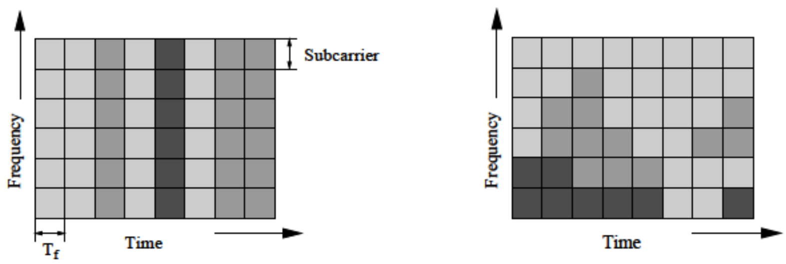


(a) Time Division Multiple Access (TDMA)

(b) Frequency Division Multiple Access (FDMA) with $N = 2$

Figure 2.7: Static multiple access schemes with an OFDM system for 3 WTs

to the amount of data to sent. With TDMA this results into the allocation of more time slots to the terminals with more data to sent while with FDMA higher amounts of subcarriers are received by the higher prioritized terminals. Assuming that the prioritization is correct and the adaption cycle (e.g. every T_f) to the varying amount of sent data is sufficient the channel is optimal utilized. Comparing Figure 2.8(b) with the static multiple access schemes shows, that with FDMA the adaption leads to a combination of the TDMA and the FDMA approach. Dynamic TDMA is used in combination with OFDM in the HIPERLAN/2 broadband radio access network [9] were a centrally scheduled scheme is employed. Centrally scheduled means that the access point controls the assignment of resources to the WTs [9]. In the common WLAN standards 802.11a and 802.11g OFDM is combined with a distributed MAC approach, while also a centralized control, called point coordinator function or infrastructure mode, is supported on top of that [47]. The combination of OFDM and FDMA (*OFDM-FDMA*) has been widely investigated. An overview to the research in this field is given in Section 4.2.



(a) Dynamic Time Division Multiple Access (TDMA)

(b) Dynamic Fre-quen-cy Division Multiple Access (FDMA)

Figure 2.8: Dynamic multiple access schemes with an OFDM system for 3 WTs

Chapter 3

Heterogeneous traffic streams

The availability of the services common in the Internet is also expected by the mobile user. The Internet traffic can be basically characterized as a mixture of Web-page related traffic and bulk-traffic, containing video or audio streams [17]. This *heterogeneous* nature of traffic reflects which services are demanded by the main part of the users connected to the Internet.

Both types of traffic have different characteristics and different demands to the transmission system. Especially the (near) real-time transmission of video streams requires high QoS levels. This has to be taken into account for the design and the evaluation of wireless transmission systems. Furthermore, the bit-rate of the two traffic types, resulting from the user behavior (Web-pages) and the employed encoding scheme (streaming media) shows a fundamental different behavior. Together with the employed techniques on the application layer and aspects of user-behavior modeling this is discussed in Section 3.2. Due to the heterogeneous nature of the application layer traffic the two different protocols Transmission Control Protocol (TCP) and User Datagram Protocol (UDP) are employed on the transport layer. In the following section the mechanisms of these protocols are introduced and discussed in terms of their interaction with wireless networks.

3.1 Internet transport layer protocols

The TCP/IP protocol suite is the foundation of the Internet, an interconnection of many autonomous systems [59], which may be single hosts and Local Area Networks (LAN) or Wide Area Network (WAN) installations. To provide mobile access to the Internet TCP/IP is applied on top of wireless transmission technologies. The protocol suite consists of the Internet Protocol (IP) which provides network oriented functions, e.g. for routing and addressing, and of the two transport protocols TCP and UDP including functions for end-to-end transmission control. The functions and the characteristics of these two transport protocols are discussed in this chapter.

The reliability of TCP makes it a good choice for the transmission of data when data integrity counts more than speed. For example this is the case for

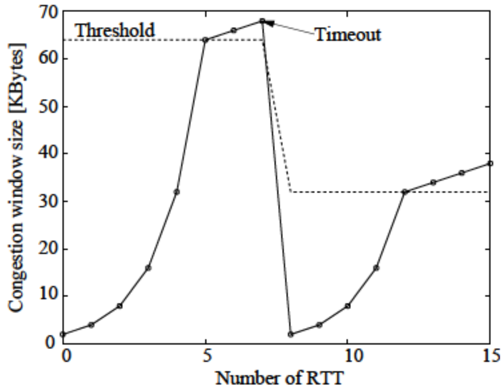


Figure 3.1: Size of the congestion window with TCPs slow-start and congestion avoidance algorithm versus the Round-Trip-Time (RTT)

Web-pages or the most types of files. TCP as basically defined in [31] provides a connection-oriented, reliable, byte-stream service. Data transmission with TCP requires an initial phase for the connection establishment. For this purpose a three-way handshake is implemented, which is detailed described in [58]. To provide reliability TCP includes mechanisms, which assure that the received data is passed error free, complete and in correct order to the application. Therefore, on the sender the application data is broken into chunks, the so called TCP-segments. The size of these segments is defined by TCP until a specific value, the Maximum Segment Size (MSS), is reached. The segments are extended by a 20 Byte TCP-header. For each sent segment a retransmission timer is maintained by TCP until an acknowledgment for this segment, which is identified by a sequence number in the TCP-header, is received. If the timer expires the segment is retransmitted. Each segment contains one checksum including the header and the data which allows TCP to recognize errors during the whole transit (end-to-end). Erroneous packets are discarded and not acknowledged, which causes a retransmission. The way how the retransmission timeout values are derived from the measured round-trip-time is basically described in [34].

To avoid congestion TCP uses a combination of the so-called *slow start* algorithm and a congestion avoidance scheme on the sender side. This combination makes use of a sliding-window (congestion window), whose size is exemplarily illustrated in Figure 3.1. In addition to the congestion window the size of the receiver window is considered. The smaller value determines how many bytes the sender simultaneously loads to the channel.

At the beginning of the transaction a threshold is defined which is typically 64 KByte [58]. If the size of the congestion window exceeds this threshold congestion avoidance is done, otherwise slow start is employed. After it was initialized with MSS the size of the congestion window rises exponential until the threshold or the size of the receiver window is reached or a retransmission timeout occurs. In Figure 3.1 first the threshold is reached, which leads to a linear increase of the window size. If a timeout occurs (at $RTT = 7$ in the example) the algorithm sets the threshold to one half of the actual congestion window size and resets the window size to one MSS. Then again the size of the window is adapted using slow start. This example illustrates how TCPs flow control scheme on the sender side adapts to the given bandwidth (slow start) and how congestion control leads to substantial changes of the sender bit-rate in case of a timeout.

TCPs congestion control was designed under the assumption that a packet loss is mostly caused by a buffer-overflow on the receiver side or the congestion of the medium. Even though this is the case for reliable cable links this is not true for wireless networks. Due to the variability of the radio channel much higher packet loss rates are quite normal on wireless links [59]. Thus, lowering the sender bit-rate using slow-start for each retransmission timeout on a wireless link results in a permanent loss of performance. Instead of doing slow start TCP should repeat the transmission of the erroneous packets immediately. However, typically TCP has no knowledge on the transmission techniques employed on lower layers, which means, that the reason for the packet loss or even the link type is not determinable by TCP. With the breakthrough of wireless networks in personal communication a lot of attention has been paid by the research community to this problem. Many proposals were made, e.g. *indirect TCP* [2] or the so-called *snooping agents* [3]. However, none of them has yet been adopted by the Internet Engineering Task Force (IETF) to the TCP standard. Since their introduction in [34] the flow control mechanisms were revised and extended several times [33], several compatible versions of TCP exist. Common versions implemented in up-to-date operation systems [43] are *Tahoe* and *New-Reno* [16].

In contrast to TCP the User Datagram Protocol (UDP), which was defined in [51], is a simple, connection-less transport layer protocol. Unlike with TCP each output operation, the application process performs to UDP, results into a “message” which is called UDP-datagram. The first 8 Bytes of the datagram contain the UDP-header. The header is followed by the payload, whose size per datagram is defined by the application. The UDP header contains a checksum covering the header and the payload to provide end-to-end error recognition. Unlike TCP the UDP provides no flow control mechanisms nor reliable transmission of data. If an UDP checksum error is recognized, the affected datagram is silently discarded [58]. It is left to the application to recognize lost, duplicated or disordered data or react to congestion. Since the mechanisms which provide reliable transmission and flow control add a certain overhead to the system this lack of features is a clear advantage for the transmission of video streams with strict timing boundaries (e.g. the “real-time” transmission of a video conference). This is illustrated by the following example: On the application server a part of a video is send within an IP datagram, which gets erroneous during the transmission. Assuming TCP to be the transport protocol, this would result in an expired timer and the retransmission of the affected packet. During the time span needed for the timeout and the retransmission the following parts of the video might be already delivered and shown to the user. Then the player software at the receiver has no use for the retransmitted packet anymore. This example shows, that small, predictable transmission delays are more important for the transport of real-time videos than reliability. Thus, on the transport layer, its simplicity makes UDP the ideal protocol for the transmission of streaming media, especially for the considered real-time video streams.

Both transport protocols TCP and UDP assume the Internet Protocol (IP) [30] to be present on the underlying network layer. The TCP segment and the UDP datagram are always encapsulated in an IP datagram which adds a header with the size of 20 Bytes. The datagram is then passed to the lower layers, e.g. the network card driver. Like UDP IP provides an unreliable and connection-less service. Any services related to reliable data transmission are left to the encapsulated protocols as TCP. Since IP is a network layer protocol its scope is to provide functions for the routing, addressing and inter-operability (e.g. fragmentation) of datagrams. Although a newer version exists the actual standard in the Internet is IP Version 4.

3.2 Heterogeneous application layer traffic

Analyzing the traffic structure in the Internet [17] leads to two different classes of traffic: data which is related to Web-pages and streaming media. Both types of traffic have different demands to all layers of the transmission system. As for streaming media data the fast transmission is the critical point, for Web-pages a reliable transmission has to be guaranteed. The characteristics of heterogeneous traffic streams and the related application layer protocols and encoding technologies are discussed in this section. Furthermore aspects of modeling are considered.

For the transmission of Web-pages the Hypertext Transfer Protocol (HTTP) version 1.0 as proposed in [5] is assumed at the application layer. It is the foundation protocol of the World-Wide-Web (WWW) and can be used to transmit data in the form of hypertext (e.g. HTML), plain text, images or any Internet-accessible file, which is summarized using the term *Web-page* in this thesis. HTTP is a transaction oriented protocol, where each transaction is independently performed between the HTTP client, e.g. a Web-browser, and the Web-server. For reliability TCP is employed, where in HTTP version 1.0 for each transaction a new connection is established, which is terminated as soon as the transaction is completed. The stateless design avoids that the Web-server, which may serve many thousands of clients for a huge Web-site, has to store and manage the state information for each HTTP connection. However, the establishment of a TCP connection per request leads to performance inefficiencies if many small objects are requested [40]. Therefore a persistent version of the protocol was proposed with HTTP v1.1 [14]. In Figure 3.2 a typical example for two HTTP v1.0 transactions is illustrated: After the establishment of the TCP connection, which lasts one Round-Trip-Time (RTT), the Web-browser requests the transaction of a hypertext document. If this document is received it is interpreted. In this example a reference to an image, which is not necessarily located on the Web-server requested during the first connection, is found and requested. Therefore, a second TCP connection is established to the related server. If the image transmission is completed

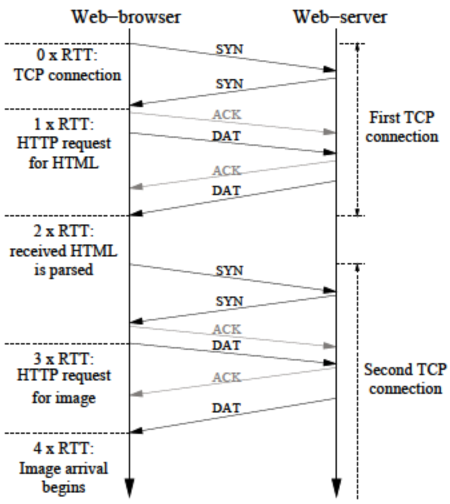


Figure 3.2: Example of HTTP v1.0 transactions during a WWW session [40]

the TCP connection is closed and the hypertext document including the image is shown to the user.

While this structure of a Web-page request has to be considered in the HTTP application layer model, the key aspect is the user behavior: Finally the user decides which Web-page is requested at which time. The user is typically not interested in achieving a constant bit-rate during the download of Web-pages. Therefore, the size of the requested pages, the time between them and as a consequence the bit-rate are typically highly varying over time. Several load models have been introduce to represent the statistic aspects of user behavior, e.g. in [4] and [60]. Especially for the evaluation HTTP v1.0 in wireless systems the *Malaga* traffic model was proposed by Reyes-Lecuona et al. in [54]. *Malaga* classifies the WWW traffic structure in session- and page-levels. A session is defined as a working session of an user downloading several Web-pages from the starting time of the Web-browser until its termination. With *Malaga* sessions are considered to be non-overlapped. On the page level one specific Web-page, containing a set of files (hypertext, images, etc.) is requested and received. For the session level the WWW traffic model defines the parameters session-inter-time, i.e. the time between two consecutive sessions, and the number of requested Web-pages per session. For the page level the page-inter-time, i.e. the time between two page requests, and the page size, i.e. the total amount of downloaded information per Web-page, are given. The distribution functions for the parameters of the Web-model are summarized in Table 3.1. The desired traffic load on the whole system can be justified by changing the mean of the session-inter-time and the number of WTs receiving Web-pages (J_h). The randomness of the user behavior as characterized with *Malaga* leads to a high variance of the data-rate. Figure 3.3 shows an example setup, representing a typical corporate environment exemplarily for a simulated time of 180 seconds. Here the high variance of the shown bit-rate is reflected by the standard deviation which is with $\sigma = 283$ KBps 77% of the mean ($\mu = 366$ KBps).

Session Level Parameter	Distribution Function
Session-inter-time	Exponential
Pages per session	Log normal
Page Level Parameters	Distribution Function
Page-inter-time	Gamma
Page Size	Pareto

Table 3.1: WWW traffic model parameters and the distribution functions as defined in [54]

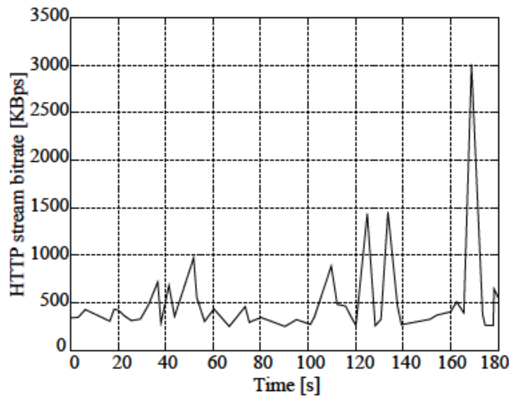


Figure 3.3: Example bit-rate of a WWW transmission characterized using the *Malaga* model [54].

Another class of traffic which is common in the Internet are video data streams. Modern video coding methods, like MPEG-4, are taking advantage of two classes of redundancy within the video streams. At first the spatial redundancy within a single original picture, called a video frame, is exploited using the mechanisms as summarized in Figure 3.4. This encoding process, known as intra-frame coding, leads to encoded video frames which are called I-frames. Secondly the temporal redundancy, i.e. the difference between consecutive frames of the original video, is exploited by methods like motion estimation, motion prediction or frame differencing, summarized by the term inter-frame coding [44].

In addition to the I-frames the inter-frame coding with MPEG-4 leads to two frame types: P- and B-frames. Unlike the intra-frame coded I-frames, which are related to a single picture of the decoded video, these frames are predicted from other frames in the MPEG-4 video stream. Like illustrated by the arrows in Figure 3.5 P-frames are predicted by the use of the previous I- or P-frame, while the *bidirectional* predicted B-frames are depending on the previous and the following I- or P-frame. A consequence of these prediction methods is that I-frames containing the highest amount of information, needed for the reconstruction of the video stream, followed by the P-frames and by the B-frames, which contain the least information. Thus, a single I-frame has to be transmitted using a much greater number of UDP-datagrams than the other two frame types.

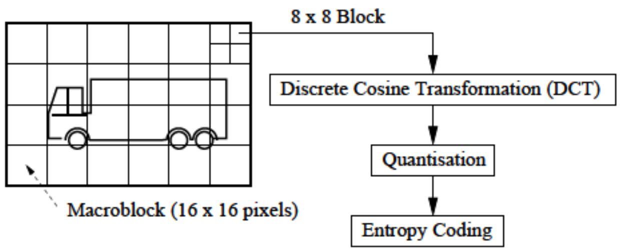
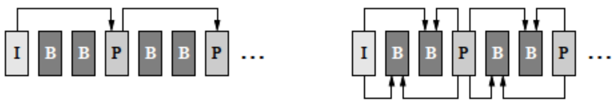


Figure 3.4: Intra-frame coding process



(a) Forward prediction of P-frames

(b) Bidirectional prediction of B-frames

Figure 3.5: Prediction methods as applied in MPEG-4 and the resulting frame relationships

The structure in which the frames are arranged is called Group of Pictures (GOP). For the GOP, which is periodic repeated in the video stream, several types exist [15]. Common with MPEG is the 12-element GOP shown in Figure 3.6. The shape of this structure is interesting if errors occur in the video stream, which is important in Section 4.4.3 when semantic video queue management schemes are discussed.

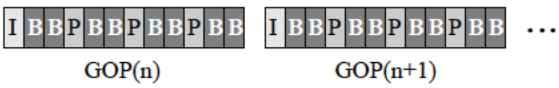


Figure 3.6: Sequence of 12 element GOPs in a MPEG coded video stream

In Figure 3.7 the bit-rate of a single MPEG-4 coded VBR video stream is shown. Here the standard deviation is with $\sigma = 18$ KBps 49% of the mean ($\mu = 37$ KBps). While the mean bit-rate results from the size of the video frames, here the Common Intermediate Format (CIF) format with 352×288 Pixel is assumed, the high fluctuation over time represents the entropy of the original video. Major differences between consecutive frames of the original video, e.g. caused by movement or cuts, lead to higher amounts of information which have to be encoded during the interframe coding process of MPEG. Consequently, the size of the P- and B-frames varies. Reflecting the above mentioned variance of the original video, Variable Bit-Rate (VBR) video streams are separated into *high-motion* and *low-motion* streams. While low-motion is typical for video telephony, the shown bit-rate results from the encoding of a high-motion source, usually representing movies and sport events.

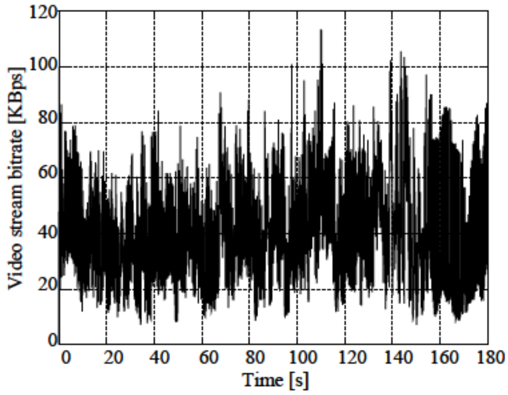


Figure 3.7: Example bit-rate of a single MPEG-4 coded high-motion video stream

The comparison of the video and the bit-rate of a WWW transmission shows significant differences. Although the mean bit-rate is much higher for the transmission of Web-pages than with video transmission, this is not characteristic, because the mean bit-rate of the video stream can be easily raised by changing the frame size. More important is the difference in variance. While the standard deviation for the VBR video is 49% of the mean bit-rate the proportion is much higher (77%) for the WWW transmission. Furthermore, the bit-rate of the VBR video varies much more over time than the Web-page related bit-rate. These differences have to be taken into account if a dynamic scheduling scheme is applied. Since the scheduler has to adapt to the traffic load a high variance over time may cause suboptimal results if the traffic load changes faster than the adaption. Furthermore, the variance of the bit-rate may cause congestion due to limited resources. Both aspects are considered in context of dynamic scheduling in the next chapter were fundamental concepts of scheduling and adaption are discussed.

Chapter 4

Dynamic scheduling in wireless networks

The fundamental differences between wireless and cable transmission lead to the development of scheduling schemes, which are specific for wireless networks. In this chapter the problem of scheduling and the basic concepts are introduced. After an overview of the actual state of the research the basic problem in the focus of this thesis is discussed. Finally, due to their relevance in this thesis, three scheduling schemes, each following a different approach, are investigated and analyzed.

4.1 Introduction

The sharing of resources is a fundamental concept of computer networks. Whether it is the Central Processing Unit (CPU) and the memory on a single node, a hard-disk on a file-server in the LAN or a link somewhere in the Internet backbone, the sharing of a resource provides a great economical benefit and is mostly the basis for the realization of these systems. However, the multiple access to a resource has to be managed. In general this is done by a scheduler, which has to decide which part of the shared resource is allocated to which request. In a wireless network the shared resource is the wireless channel and the requests are made by the traffic streams, which demand bandwidth. Here the task of a scheduling policy is basically to provide the minimal queuing delay as possible, while considering the following conditions:

- Capacity bounds of the wireless channel
- Contention of the Wireless Terminals (WTs) for the channel
- Fairness has to be provided to all WTs
- Given quality levels for several types of application layer traffic, the so-called QoS, have to be assured.

While all these aspects also have to be considered by a scheduling policy for wired systems, the time- and frequency variant nature of the wireless channel adds difficulties here. Due to the significant variation of the CNR (Section 2.2) the capacity boundaries of the wireless channel are typically unknown. Assuming the capacity to be constant leads to a permanently over- or underestimation, causing increased error rates or the under-utilization of bandwidth. Furthermore, the employment of a simple scheduling policy such as First-In-First-Out (FIFO), may cause blocking at the head of the queue if the channel states are bad. This problem was investigated in an early work of Bhagwat et al. [6]. Due to bad states of the wireless channel the packet at the head of the queue may has to be transmitted several times before it is received correctly. Using a FIFO scheduling policy other WTs in a good channel state have to wait until this packet is acknowledged. Bhagwat et al. show that channel utilization and average latency can be improved if recipients with a good channel-state are prioritized. The found improvements demonstrate the advantage of the channel-state dependent approach.

Another aspect which has to be considered by the scheduling policy is the contention of the WTs for the channel. As discussed in Section 2.5 a static allocation of the channel resources, i.e. the subcarriers, would result in under-utilization of the wireless channel and would lead to performance problems. This is a well known problem of telecommunication systems and is treated by various so-called dynamic- or statistic multiple access schemes. With wireless systems here again the variance of the channel-states have consequences for the scheduling: If the state of a specific WT decreases, less data can be correctly received. Compared to the WT with better channel-states this WT is lower prioritized by the channel. With channel-state dependent scheduling the lower priority propagates from the physical layer to the scheduler on the link layer. Thus, the channel-state dependent scheduling policy has also take aspects of *fairness* into account.

The varying channel capacity leads to problems if a given quality level has to be assured on the application layer. Decreasing channel-states will cause higher latency, lower throughput and consequently a lower QoS perceived by the user. This is especially the case with *soft real-time* [1] traffic, e.g. real-time video or audio streams, since here transmission delays directly lead to gaps in the continuously transmitted streams. To avoid the permanent reservation of bandwidth in the background if application layer QoS has to be assured for streaming media, extensions to schedulers were proposed. Due to the high demands of real-time video streaming, particularly video is considered for this approach. Here the varying bit-rate of the video stream is the source of variability. An early approach was formulated by Yeadon et al. in 1996 considering the transmission of MPEG video streams in a wired network. In [65] several classes of so-called *continuous media filters* are applied on special filtering nodes if congestion occurs in the network. Analyzing the cost Yeadon et al. came to the conclusion that the complexity of these filters

[1] The term *soft* denotes that the miss of the real-time boundaries results in a quality decrease in opposite to *hard* real-time applications, e.g. aircraft safety control systems, where the miss of a deadline is unacceptable [45].

allows the application to provide QoS for MPEG video. This approach is known under the topic *dynamic* QoS. With dynamic QoS the bit-rate of the streaming media is adapted to the actual resources available at the system. However, with this approach the scheduler has to consider application layer information on the link layer.

4.2 Related work

In this section an overview to the research that was done recently in the field of dynamic scheduling is given. Basically, the following three approaches are considered:

- Channel-state dependent scheduling: adapting the scheduling policy to the measured CNR
- Dynamic multiple access schemes: allocate resources considering traffic load per terminal
- Adaptive filters: analyzing and filtering media streams in order to the measured traffic load

For wire-less OFDM-FDMA systems the concept of channel-state dependent scheduling has been investigated on the link layer by Gross et al. in [19]. The approach makes use of the OFDM subcarrier states, provided by the physical layer, and enables the adaption of the spectral position of the subcarriers chosen per WT to the varying CNR. For the calculation of the subcarrier assignments the Advanced Dynamic Algorithm (ADA) was presented. The performance has been evaluated on the link layer using discrete event simulation. It has been found, that the throughput gain due to the adaption to the varying CNR in comparison to the static OFDM-FDMA scheme is from 30 to 40%. The subcarrier assignment approach and ADA are further discussed in Section 4.4.1. Another channel-state dependent approach for OFDM-FDMA systems was presented in [39]. Here Le-Ngoc et al. proposed and investigated an scheduling scheme combining the channel adaptive assignment of OFDM subcarriers and dynamic multiple access. In the performance study two types of application layer traffic were considered: non-real time traffic, related to WWW browsing, and soft-real time VBR video streams. Both types of traffic were investigated separately in two homogeneous scenarios. For WWW traffic a performance increase of 56% and for video streams of 57% in terms of satisfied users was found for the dynamic system.

In [50] an algorithm is shown which calculates subcarrier assignments and power control in an OFDM-FDMA system. However, in contrast to the channel-state dependent schedulers the adaption algorithm is oriented on the individual resource requirements. It is a modification of Wongs algorithm, presented in [64]. Both algorithms follow the dynamic FDMA approach, aiming at the improvement of the power consumption for wireless transmission. The simulation results of Pietrzyk et al. show a superior performance compared to Wongs proposal and static OFDM-FDMA. Another dynamic multiple access scheme, considering the individual resource requirements for the allocation of subcarriers in an OFDM system, was proposed in [20]. The scheme has been studied separately and in combination with the above mentioned channel-state depended approach by Gross et al. The performance was analyzed on the application layer for the transmission of MPEG-4 coded video streams. In terms of the number of supportable WTs at a given video quality level the the combined dynamic approach (channel-state depended scheduling in combination with dynamic multiple access) outperforms the static case for up to 300% while 2/3 of this gain was achieved by the dynamic multiple access scheme. In this thesis the above mentioned combined dynamic approach is called *dynamic OFDM-FDMA* and is further discussed in Section 4.4.1 and 4.4.2.

For the transmission of soft video streams in wireless systems [41], [18] and [35] proposed extensions to wireless scheduling policies. In [41] the demands of the transmission of multimedia to a scheduler in a wireless system is analyzed. Considering the found performance gap in application layer QoS Meng et al. propose extensions to the Idealized Weighted Fair Queueing (IWFQ) [56] scheduling policy. Here the scheduler takes application layer information of the investigated MPEG-4 and H.263 video streams into account. Following this approach Klaue et al. [35] extended the scheduling policy of the dynamic OFDM-FDMA scheme by Gross et al. Here the video packets in the access point queue of a cellular OFDM-FDMA system are weighted according to the semantics of a MPEG video stream. The transmission order and the deadline of a packet are then defined by these weights. After the deadline is hit for a specific packet, it is removed by the scheduler. This approach is called *semantic VQM* and is further analyzed in Section 4.4.3.

A similar extension to a scheduling policy is given in [18]. Francini et al. extend the Frame-Based Fair Queueing (FBFQ) scheduler with a packet dropper and a call admission control algorithm in order to provide bandwidth and delay guarantees for voice traffic. In contrast to the above mentioned extensions here the capacity of the system is measured continuously. If low capacity is found the media streams are modified in order to lower the traffic load. This leads to a different approach also considering application layer information which is known under the topic of *dynamic* QoS and *adaptive filtering* [26]. Here the term adaption stands for the modification of the streams according to the traffic load. After proposed by Yeadon et al. in [65] several classes of adaptive filters were investigated. In [25] a modification of the simple drop filter for MPEG coded video streams is proposed. Here the dropping is performed by considering semantic aspects and the structural relevance of the video frames in the stream. This filter was investigated for an end-to-end approach, where server and client control the filtering nodes within the wired network. The key of dynamic QoS is the bounded cost of the filtering and the quality of the adaption algorithm. This was analyzed in [26] by Hemy et al. for wired networks. Here an adaption algorithm is proposed, which is performed at the client, while the filtering of MPEG video streams is done by nodes in the network. Hemy et al. show, that this dynamic QoS scheme leads to significant lowering of the network load and the video frame drop rate if the wired network is overloaded. Furthermore, for an experimental setup the processor usage of the adaptive filter was investigated. Therefore, it was shown that the processor load highly depends on the bit-rate of the transmitted video.

The research in the field of dynamic scheduling is the basis for the following chapter where the problem that is considered in this thesis is described.

4.3 Problem description

Unlike a cable link a wireless system suffers from variation of the CNR caused by effects on the radio channel. This was discussed in Chapter 2 where also the transmission techniques OFDM and adaptive modulation were introduced which address this problem on the physical layer. However, the adaption to the CNR using adaptive modulation forwards the variation in terms of channel states to the link layer. In Section 4.1 the channel-state dependent scheduling approach was introduced which avoids the problem of head-of-line blocking in a queue if the states to a specific WT are low. As discussed in Section 4.2 recent research was done on the algorithms and it was shown, that channel-state dependent schedulers show good performance results on the link layer and in homogeneous application layer scenarios.

Another problem which has to be considered by the scheduler is the avoidance of under-utilization of the channel resources. This has been discussed in Section 2.5 and 4.1 of this thesis under the topic of dynamic multiple access. Dynamic multiple access is based on the adaption of the scheduling policy according to the variance of the traffic load. Although this concept is not specific for wireless networks its application is useful due to the highly limited resources if the channel is in a low state. In Section 4.2 several proposals for dynamic schedulers in wireless systems were presented. If this dynamic multiple access scheme is combined with a channel-state depended approach the scheduler considers two sources of variability: the channel variation on the physical layer and the variation of the traffic load, resulting from the upper (transport and application) layers. To both sources the scheduling policy can be adapted. This aspect is illustrated in Figure 4.1.

In this thesis the performance of a combined dynamic scheduling approach, called dynamic OFDM-FDMA is studied on the application layer of a wireless system. This is done assuming heterogeneous traffic load to be present which consists

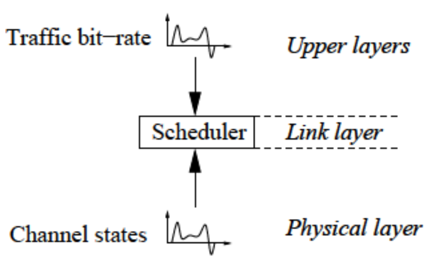


Figure 4.1: Sources of variability and the related layers

of data that is related to Web-pages and VBR video streams. As mentioned in Section 4.2 these two types of traffic have been studied separately in combination with dynamic schedulers. However, this has not been done so far for the *simultaneous* transmission of Web-pages and video streams. In Chapter 5 of this thesis it is shown how dynamic and static scheduled OFDM-FDMA systems perform if heterogeneous traffic is assumed.

This traffic scenario is much closer to the traffic flows found on Internet backbones [17], than the separate consideration of Web- or video streams. Furthermore, this traffic mixture is interesting due to the totally different behavior shown by Web-page and VBR video streams. As discussed in Section 3.2 the transmission of Web-pages leads to much higher variances of the bit-rate than the transmission of video streams. However, the encoding process of video streams results into heavy fluctuations of the bit-rate over time. Furthermore, both parts of the heterogeneous traffic have totally different claims to the transmission system. In contrast to the transmission of video streams, where the video quality directly depends on the transmission delay, the transmission of Web-pages is not time critical from the application point of view. However, here the reliability in terms of an error-free reception counts. This is reflected by the transport layer protocols, which are applied for both stream types. The transmission of Web-pages is based on TCP which provides reliable transmission regardless to the transmission delay. Furthermore, TCPs flow control adapts the sending rate in order to avoid congestion. In contrast to TCP the UDP, as applied for the video transmission, includes no flow control. If it is demanded by the sender application, as much bandwidth as provided by the lower layers is used. As a result in heterogeneous traffic scenarios UDP streams tend to rule out TCP under high load situations [42].

The employment of adaptive filtering in such high load situations was studied for wired networks (Section 4.2). Here advantages in terms of lowering network load and increasing transmission quality were found. As mentioned in Section 4.2, with streaming media adaptive filtering can be applied. For wireless systems the adaptive filtering has not been employed for high load scenarios. In these approaches the support of more users or QoS guarantees was aimed. Thus, in Chapter 6 of this thesis an adaptive filtering scheme, called *adaptive Video Queue Management (VQM)*, is proposed. Adaptive VQM provides the adaption of the video stream bit-rate to the estimated traffic load. The objective of adaptive VQM is to avoid congestion in the system, while still assuring high quality for the filtered video (in Section 5.4.2 it is discussed what “high quality” means). Therefore, a load-estimator is combined with a filter, which considers semantic elements in the video streams. In wireless networks the adaptive VQM can be applied on the access point in order to prevent and handle high load situations, e.g. if many WTs are requesting video streams. For the dynamic OFDM-FDMA system the performance of this approach is evaluated in Chapter 7. Assuming heterogeneous traffic, per scenario two aspects are of particular interest: At first, the amount of capacity which can be provided by the adaptive VQM to all WTs in the cell. Secondly, it is important how this “freed” capacity is then distributed according to the stream types. Precisely, what benefit can be found for the transmission of a Web-page compared to the gain for the send video streams. Furthermore, the applied filter should do not decrease the video quality. Thus, the effect of the filtering to the video quality is analyzed more detailed.

4.4 Analysis of the investigated dynamic scheduling schemes

Three dynamic scheduling schemes which are the basis of the investigated dynamic OFDM-FDMA system considered in this thesis are discussed in this section. This includes the *dynamic subcarrier assignment* and the *dynamic subcarrier allocation* as both proposed by Gross et al. in [19] and [20]. The dynamic subcarrier assignment scheme makes use of the ADA to calculate subcarrier assignments in order to the channel-states. The dynamic subcarrier allocation is a dynamic FDMA scheme, which allocates the number of subcarriers per WT based on the queue lengths and is combined with the dynamic assignment scheme to dynamic OFDM-FDMA. An example of the two dynamic schemes is shown in Figure 4.2 where the color of a single block denotes one subcarrier obtained by a specific WT for the duration of T_f.

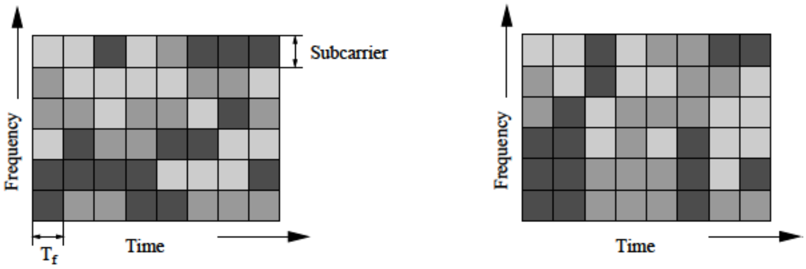
With the dynamic subcarrier assignment each of the three WTs receives the same number of subcarriers ($N = 2$) per T_f but their position in the frequency domain varies. In comparison to static OFDM-FDMA this flexible choice of the subcarrier assignments enables adaption to the varying CNR. The dynamic subcarrier assignment algorithm can choose for each WT those subcarriers with the highest CNR and therefore, due to adaptive modulation, with the highest amount of Bits per symbol.

Combining the dynamic assignment with the dynamic allocation of subcarriers leads to the modification of N, as shown in Figure 4.2(b). This enables another adaption technique: With the dynamic subcarrier allocation N is adapted to the traffic load. If the size of all outstanding packets in the access point queue for a specific WT is high then more subcarriers will be allocated to this WT.

In addition to the scheduling schemes which adapt the spectral position and the amount of the OFDM subcarrier per WT two VQM schemes are analyzed in this section. At first a simple FIFO approach is discussed. Secondly the semantic VQM as proposed by Klaue et al. in [35] is analyzed. The characteristics of these two VQM have been considered during the development of the adaptive VQM.

4.4.1 Dynamic subcarrier assignment

As discussed in Section 2.2 the effects on the radio channel lead to a varying CNR over time and frequency ($G(f, t)$). In an OFDM system the given bandwidth is separated into subcarriers. Therefore, the resulting CNR per subcarrier (s)



(a) Dynamic subcarrier assignment with $N = 2$

(b) Dynamic assignment and dynamic allocation of subcarriers

Figure 4.2: Dynamic OFDM-FDMA schemes for 3 WTs

regarding to a specific WT (j) can be given by $g_{j,s}(t)$ as in Equation 4.1.

$$g_{j,s}(t) = \frac{a_{j,s}^2}{n_{j,s}^2} \tag{4.1}$$

For all S subcarriers regarding to all J WTs the CNR-matrix $\hat{G}(t)$ of dimension $J \times S$ can be derived from the $G(f,t)$ which was given in Equation 2.2:

$$\hat{G}(t) = \begin{pmatrix} g_{1,1}(t) & g_{1,2}(t) & \cdots & g_{1,S}(t) \\ g_{2,1}(t) & g_{2,2}(t) & \cdots & g_{2,S}(t) \\ \vdots & \vdots & & \vdots \\ g_{J,1}(t) & g_{J,2}(t) & \cdots & g_{J,S}(t) \end{pmatrix} \tag{4.2}$$

If the time is slotted into MAC frames of length T_f this leads to one $\hat{G}(t)$ per T_f. By multiplying each value of $\hat{G}(t)$ with the applied transmission power P_{tx} that is constant for all the subcarriers during the whole transmission, the matrix $\hat{Q}(t)$ is obtained containing the SNR values for the transmission towards all J WTs on all S subcarriers. From $\hat{Q}(t)$ the adaptive modulation scheme derives the applied modulation technique, which results to a number of bits (m) represented by a symbol of the chosen modulation alphabet. (The procedure was described in Section 2.4.) Since the resulting m called the *subcarrier state* is specific for the considered WT and subcarrier all $m_{j,s}$ can be expressed by the $J \times S$ matrix $M(t)$. The CNR-matrix is then used by the ADA-algorithm to calculate the subcarrier assignment for each WT for the next downlink phase $(t+1)$. This means that the measured CNR values in $\hat{G}(t)$ may be outdated if the subcarrier assignments are applied. This problem is discussed under the topic of *realistic channel knowledge* in [22].

In this thesis the ADA as proposed in [19] is considered for the dynamic assignment of subcarriers. ADA makes use of priorities for each WT. For each WT a weight value (w) is calculated where a single weight value for the specific WT j is obtained by the sum of the channel gain values of the i WTs with a lower priority than j:

$$w_{j,s}(t) = \sum_{\forall i} g_{i,s}(t). \tag{4.3}$$

After the calculation of the weights for all J WTs the assignment algorithm selects the subcarriers with the highest possible weight ratio, defined by

$$\frac{g_{i,s}(t)}{w_{i,s}(t)}. \tag{4.4}$$

This is done N times for each WT, starting with the WT which received the highest priority. Therefore each WT obtains N subcarriers, where to the WTs with higher priorities better subcarriers are assigned. Finally the priorities are rotated to provide fairness.

ADA provides a heuristic and not the optimal solution for the subcarrier assignment as the well known Hungarian algorithm [38]. However, the optimal solution can not be obtained the acceptable timing bound that is smaller than T_f. In [19] it was shown that the throughput gain achieved with solution provided by ADA is most only 5% off from the optimal solution.

4.4.2 Dynamic subcarrier allocation

For the description of the dynamic subcarrier assignment scheme it was supposed that a constant number of subcarriers (N) is given to each WT during the assignment procedure. As discussed in Section 2.5 static multiple access schemes are far from being optimal.

Considering static FDMA each WT receives N subcarriers per downlink phase regardless of data has to be send to a WT or not. If, for example, no data is send to i WTs for a downlink phase $i \times N$ subcarriers are not utilized. In contrast, a dynamic FDMA scheme then allocates these $i \times N$ subcarriers to those WTs which have to receive data. This will clearly result in an improvement of the throughput rates to these WTs.

In the OFDM-FDMA system investigated in thesis the dynamic subcarrier allocation scheme as proposed in [20] is considered. It provides the adaption to the varying traffic load by measuring the queue length (d_j), i.e. the sum of all the packet sizes in the access point queue, for a single WT. The dynamic allocation scheme then compares the obtained d_j values to the sum of all d_j and calculates the amount of subcarriers allocated per WT (s_j) as given in Equation 4.5.

$$s_j(t) = 1 + \left\lfloor (S - J_a) \frac{d_j(t)}{\sum_{\forall J} d_j(t)} + 0.5 \right\rfloor \tag{4.5}$$

Since prior to this allocation that is derived from the queue lengths each WT which has data to receive during the next downlink phase whose sum is denoted by J_a obtains one subcarrier. Thus, only $S - J_a$ subcarriers can be weighted in Equation 4.5. The allocation is done once per MAC frame before the downlink phase starts, for which the achieved subcarrier allocations are applied.

4.4.3 Video queue management

In general the term Video Queue Management (VQM) is used to describe a number of functions employed on queues, which contain packets of video stream data. It is supposed that these queues are located at the access point and even in scenarios with heterogeneous traffic the access point knows which queues contain video packets. Furthermore the VQM on the access point is able to extract parameters from these packets which have semantic relevance for the related video stream. For example, with an MPEG coded video stream this may be the type of the video frame which is (partially) included in an UDP datagram. Extracting this information the VQM on the access point knows whether this UDP datagram and the related IP datagram belongs to an I-,P- or B-frame and may adapt its scheduling policy in order to this information. In context of a VQM the scheduling policy decides which priority is given to the transmission of a particular video packet.

A simple form is a FIFO scheduling policy where the access point transmits the packets in order of the reception. This is the basic approach of the first VQM introduced here: the *simple VQM*. The only function which distinguishes the simple VQM from the plain FIFO queue is the removal of packets related to outdated video frames. This function is the consequence of the strict boundaries for the transmission delay on the downlink if videos in real-time application scenarios, e.g. video conferences, are considered. If the transmission delay exceeds 180ms the distorting effect of acoustic echoes becomes annoying to the user [62]. Furthermore, time lags in the video stream may become noticeable since the video decoder has to wait until all packets per video frame have arrived. For the prevention of problems due to boundless transmission delays it is assumed that video frames (or parts of them) which are received later than a given global deadline ($\hat{D}$) are not decoded at the receiver anymore. Therefore, the simple VQM may remove video packets whose transmission delay exceeds $\hat{D}$. Because the VQM is located at the access point, i.e. between sender and receiver, it has no knowledge about the absolute end-to-end transmission delay of a packet. Thus, the value for $\hat{D}$ has to be much smaller than the overall acceptable transmission delay of 180ms. In this thesis $\hat{D} = 100$ms is chosen which is quite pessimistic. The simple approach of this VQM can be formalized as in Figure 4.3. Considering a single packet with the sending

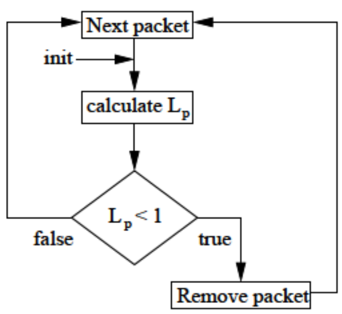


Figure 4.3: Algorithm of the simple VQM as performed independently for each queue

time T_{tx} the algorithm calculates the "lifetime-index" L_p as in Equation 4.6.

$$L_p(t) = \frac{\hat{D}}{t - T_{tx}} \tag{4.6}$$

If then $L_p < 1$ which means that the transmission delay is larger than the global deadline the packet is removed from the access point queue. The algorithm is performed for each queue independently. It begins with the first packet at the tail of the queue and ends if all packets are inspected.

In addition to the errors which may occur during the encoding and the transmission of the video frame the removal of IP datagrams due to exceeded deadlines causes errors in the MPEG-4 video stream. Therefore, it is important to know how the decoder on the receiver reacts in case of errors. If an error occurs, the decoder has to resynchronize to the bit-stream. Then the missing information has to be replaced. Therefore, an *error concealment* scheme is performed. In the simplest case the replacement is done with monochrome blocks while more complex error concealment schemes considering the data from the preceding and following video frames [44]. Usually this is done on the Macro-block level which means that in case of an error the whole affected Macro-block is rejected. Beneath the quality of the decoder the effect of errors to a MPEG-4 video stream highly depends on the type of the affected frame. If an error occurs in an I-frame, it propagates through the video streams until the next I-frame appears. The reason for the so called *error propagation* is that all P- and B-frames within the considered 12-GOP are predicted from the preceding I-frame (Figure 3.5). This lasts into 11 also affected frames due to one erroneous I-frame. The effect of errors within a P-frame is similar: If the first P-frame in the 12-GOP is affected the error propagates through all B-frames and the following P-frame. In case of an error in the second P-frame of the 12-GOP only the two last B-frames are also affected (Figure 3.5(b)).

The approach for a more sophisticated VQM scheme is to avoid the removal of I-frames and then P-frames due to the exceeded deadline. This means that semantic information of the video stream, i.e. the video frame type, is considered within the scheduling policy. This leads to the second VQM discussed in this thesis called the *semantic VQM*. As proposed in [35] the basic idea of the semantic VQM is that semantically more important frames are prioritized. Therefore, the VQM manages the order of the transmission of the IP datagrams in the access point queue: If the queue contains datagrams related to I-frames they are transmitted first followed by datagrams containing parts of P-frames. Finally the datagrams which are related to B-frames are transmitted. Furthermore, with the semantic VQM the deadline which defines when a packet is deleted from the queue becomes frame type dependent. Therefore, a weighting factor (w) is added to Equation 4.6 which can be formalized as in

$$\hat{L_p}(t) = w(\Phi)\frac{\hat{D}}{t - T_{tx}} \tag{4.7}$$

for the calculation of the "lifetime-index". For the weights the values in Equation 4.8 are considered which are defined in [21]. The Φ denotes the frame type, which is extracted from the considered packet.

$$w(\Phi) = \begin{cases} 1, & \Phi = \text{I frame} \\ 0.75, & \Phi = \text{P frame} \\ 0.5, & \Phi = \text{B frame} \end{cases} \quad (4.8)$$

This means, that P-frame related packets are dropped 25% and B-frame packets are dropped 50% earlier than I-frame packets. Therefore capacity for I- and P-frames is provided by the deletion of packets related to B-frames. Since B-frames are no prediction-source for other frames the errors caused by the removal of the packets will not propagate.

The algorithm of the semantic VQM, as it is independently performed for each access point queue containing video data, is illustrated in Figure 4.4: At first all video packets in the queue are sorted according to the type of the related video frame. Therefore, the packets which are related to I-frames are shifted to the tail of the queue, followed by the packets related to P- and B-frames. Then basically the same iteration over all packets in the queue as in the simple VQM starts. However, w is considered in the calculation of $\hat{L_p}$ in step 2. If the frame specific deadline for a packet is hit ($\hat{L_p} < 1$), it is dropped during step 3. The algorithm ends if the head of the queue is reached. Although this algorithm is performed once per downlink phase in the considered system this is not essential. The more infrequent execution would decrease the calculational overhead caused by the semantic VQM to the access point. However, this would also lower the adaption capabilities of the scheme to the actual queue content.

The two discussed VQM schemes are the building blocks and the yardstick for the VQM scheme proposed in Chapter 6. A newly developed VQM should clearly outperform both of them. This is investigated in the performance study in Chapter 7.

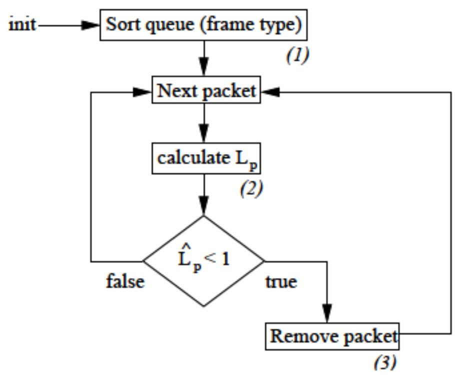


Figure 4.4: Algorithm of the semantic VQM as performed independently for each queue

Chapter 5

Performance study of dynamic OFDM-FDMA

In recent research the performance of the dynamic OFDM-FDMA approach was investigated on the link layer and on the application layer where a homogeneous traffic scenario with video streams was considered. On the link layer throughput gains vs. the static OFDM-FDMA scheme were found from 30 to 40% [20]. For the video stream scenario a gain of up to 300% in terms of supportable WTs at a given quality level was achieved [20].

The performance study in this chapter now investigates which portion of the benefit achieved by the improvement on the link layer reaches the application layer if *heterogeneous* traffic is considered. Therefore, a number of WTs requesting Web-pages is assumed while the other part of the WTs is receiving video streams at the same time. Three combinations of the dynamic scheduling schemes are investigated:

- Static OFDM-FDMA: static subcarrier assignment and static subcarrier allocation, the spectral position and the number of the subcarriers a single WT obtains (N) are fixed

- Semi-dynamic OFDM-FDMA: dyn-amic sub-carrier assign-ment and static sub-carrier alloc-ation, N is fixed

- Dynamic OFDM-FDMA: dynamic subcarrier assignment and dynamic subcarrier allocation

In order to reflect the users perception the performance is measured on the application layer. Thus, separate metrics for the transmission of Web-pages and video streams are applied which are introduced in Section 5.4. Furthermore, the investigated scenario and the chosen parameters have to be discussed.

Additionally, this chapter gives a short overview to the simulation framework and to the resulting timing and encapsulation structure which is relevant for the whole thesis.

5.1 Considered system and scenario

For the performance studies in this thesis a single wireless cell as described in Section 2.1 is considered. The cell has a radius of R and contains one access point and J wireless terminals. While the position of the access point is fixed the WTs are moving with a maximum speed of v_{max}. The access point which is connected to the Internet via a cable link, forwards data between application servers located in the Internet and the WTs using a wireless OFDM link. Furthermore, various combinations of the dynamic scheduling schemes as described in Section 4.4 are applied.

For all simulations heterogeneous traffic is assumed. Thus, the WTs are requesting two different types of application layer data from the servers: J_h terminals are requesting Web-pages using the HTTP and the TCP protocol and to J_v WTs MPEG-4 coded video streams are transmitted by the use of UDP. In Figure 5.1 an example setup of the considered system is illustrated, different colors for the servers and the WTs denote different types of application layer data. From setup to setup the amount of HTTP and video receiving WTs is varied. However, during the simulation of a single setup the application type of a specific WT is fixed.

In this thesis only the downlink transmission of data from the access point to the WTs is considered. The time is slotted into so-called MAC frames of length T_f. It is supposed that the subcarrier assignments and allocations are calculated once per T_f.

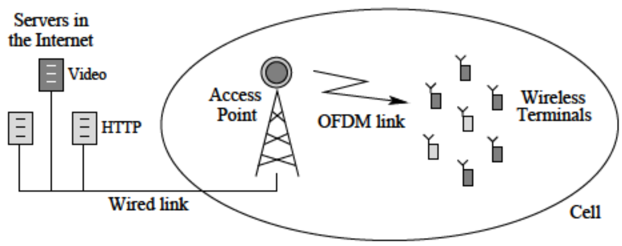


Figure 5.1: Example setup of the considered cellular system

5.2 Simulation framework

All results presented in the performance studies were obtained via discrete event simulation. Figure 5.2 shows the structure of the simulated system and illustrates how the modules are interconnected.

The system consists of application servers, the access point and the wireless terminals. Each application server module consists basically of a traffic generator and a transport protocol submodule. [1] Depending on the type of the sent traffic

[1]The employed IP modules are not shown here. Since point-to-point connections are assumed in the simulator, IP does nothing more than adding 20 Bytes of overhead per transport layer packet.

this is either a WWW traffic generator, according to the *Malaga* WWW traffic model (Section 3.2) combined with a TCP submodule or an video traffic model and a UDP submodule. The video traffic generator is based on trace-files which were produced using MPEG-4 coded high-motion VBR video streams. From the server modules messages are sent to the queues which are located at the access point. While different types of traffic use different queues, one queue for each WT is assumed. The modules for the queues, the application servers and for the WTs are replicated according to the chosen numbers of video and WWW terminals.

Beneath the queues the basic functions of the access point can be divided into two modules. In the channel module the CNR is generated considering the effects which were discussed in Section 2.2. From the CNR the channel states are derived using the adaptive modulation scheme as described in Section 2.4 while a constant transmission power is assumed. The channel states are then sent to the scheduler module. It contains three submodules where the subcarrier assignments (AS), the allocation of subcarriers (AL) and the VQM are performed. Several submodules can be used, containing e.g. static or dynamic algorithms. According to the results of the submodules for the used subcarriers (AS), the number of subcarriers (AL) or the index of the removed video packets (VQM) the scheduler interacts with the queues. This interaction and calculation process is done cyclic every T_f and is followed by the transmission of the message at the head of each queue. This is done simultaneous to the J WT modules where the messages are received, processed (according to the used transport protocol) and statistics for the measured values are calculated.

In addition to the above mentioned modules, for all simulations a link layer model is supposed which basically provides MAC-frame encapsulation, error-coding and signaling. It is detailed described and analyzed in [22] and [23]. Since it does not affect the studies presented in this thesis, it is not shown and not discussed here. The simulator was implemented using the OMNeT++ discrete event simulation framework [61] while the transport layer protocols the IP-suite for OMNeT++ [63], developed at the university of Karlsruhe was integrated.

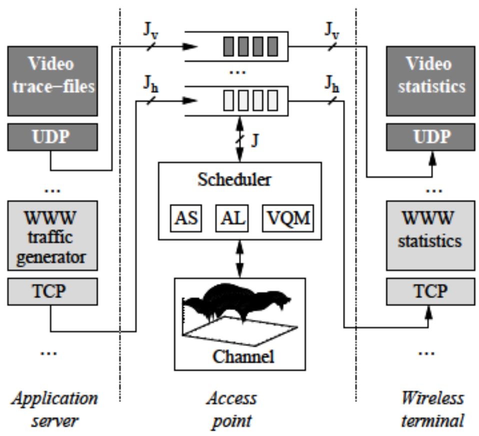


Figure 5.2: Basic structure of the applied discrete event simulator

5.3 Timing structure and encapsulation

The delays considered in the simulated system are summarized in this section.

If a WT demands a video stream, a request is sent to the video server. Since no uplink is considered in the system model the request is sent directly from the access point to the video server for which a duration of T_i is needed. After a service delay T_h at the video server, e.g. for encoding or hard-disc seeking purposes, the first part of the MPEG-4 coded video is send within an UDP datagram. Depending on the size of the transmitted video frame the UDP datagram may contain a full video frame or a part of it. After the encapsulation the resulting IP datagram is sent via the wired link to the access point were it is queued for the duration of T_q.

Depending on the results of the adaptive modulation a specific part of the IP datagram can be sent to the WT during T_f. This part is modulated and the resulting symbols are transmitted simultaneously on the OFDM subcarriers via the wireless link. The transmission of the whole encapsulated IP datagram lasts T_w. As a simplification it is not assumed that the decapsulation and data processing, e.g. the decoding of the received video frame causes any delays on the WT. Thus, the delays which occur during the transmission can be summarized as follows:

$$RTT = 2 \times T_i + T_h + T_q + T_w. \tag{5.1}$$

For the request and the transmission of a Web-page the timing structure is the same. However, since TCP is employed on the whole transmission path the whole Web-page or parts of it are encapsulated within a TCP segment.

5.4 Transmission metrics

In this section the metrics which are investigated in all performance studies are defined. Although in the considered system the queuing delay at the access point (T_q) is the critical factor the transmission quality the user perceives at the application layer is what finally counts. This leads to two types of application layer metrics, considered separately for the transmission of Web-pages and MPEG-4 video.

5.4.1 HTTP transmission metrics

In order to measure the performance provided by the transmission system to those WTs receiving Web-pages the metrics as listed in Table 5.1 are considered. In [40] Menascé et al. came to the conclusion that latency and throughput are the two most important performance metrics for Web-systems. These two classes are considered for the metrics for the HTTP transmission in the simulated system. The throughput is measured in Bits per second (Bps) on the application layer of each WT separately. To get closer to the users perception then the latency of a requested Web-page is monitored by measuring the transmission time of a single page. The faster a Web-page is received and displayed, the more the user

HTTP Metric	Unit of Measurement
TCP bit-rate	Bits per second [Bps]
Web-page transmission time	Seconds [s]
Standard deviation of the Web-page transmission time	Seconds [s]

Table 5.1: Considered HTTP transmission metrics

is satisfied with the transmission system. [2] Furthermore, the transmission delay of a Web-page should be predictable for the user. This is the fewer the case, the higher the delay varies. This variation is represented by the standard deviation (σ) of the Web-page transmission time.

For the presentation in the performance studies the results obtained for all J_h WTs during the whole simulation time are averaged over J_h. In addition to the mean values the variance of the results was investigated for 95% confidence. Due to their small size (typically less than 0.5% of the plotted value) no confidence intervals are shown in the diagrams given in the performance studies.

5.4.2 MPEG-4 video transmission metrics

During the transmission the video is continuously decoded and shown to the user. Since then the users subjective impression of the video quality is what finally counts in this thesis the measurement of video quality is based on a user-preceptive approach.

Typically a subjective evaluation has to be done by tests, where humans have to watch and grade the video, which are time intensive and costly. Here a different approach is used: With the Distortion In Interval (DIV) metric, as proposed in [35], the subjective impression of an user can be mapped to an objective metric. The DIV metric is based on the Mean Opinion Score (MOS), which is standard for the subjective rating of still pictures [32]. By the use of Table 5.2 [48] the subjective metric MOS can be derived from the objective Peak Signal-to-Noise Ratio (PSNR), which is widely used for the evaluation of picture quality [55]. Therewith, the MOS for a single picture of the decoded video can be obtained by calculating its PSNR value.

$$PSNR(a, b) = 20 \log \frac{2^k - 1}{MSE(a, b)} \tag{5.2}$$

This is done by the use of Equation 5.2 where the decoded video frames a and b are compared. Here MSE denotes the mean square error between the two frames and k the color depth of the video in bits. The result of Equation 5.2 is given in dB. The frame a denotes the video frame in prior to the transmission. Frame b is the encoded, transmitted (and maybe damaged) and decoded video frame. If a is

[2] Here it is assumed, that the program on the application layer e.g. the Web-browser does not causes any delays.

assumed to have always a high $PSNR$ the comparison shows the influence of the transmission errors.

PSNR [dB]	MOS grade	Video quality
> 37	5	Excellent
$31 - 37$	4	Good
$25 - 31$	3	Fair
$20 - 25$	2	Poor
< 20	1	Bad

Table 5.2: PSNR to MOS conversion as in [48]

To obtain one value for the quality of a large number of still pictures, as in a video stream, the simplest approach would be to obtain one MOS value per video frame and then calculate the mean. However, this straight-forward approach usually gives no adequate reflection of the users impression. For example: If the first seconds of the video are lost during the transmission and the quality of frames received and decoded during the following larger time interval is excellent, a good MOS value is obtained on the average. This will clearly not match the subjective user impression. To avoid the false estimation as performed by time-averaging metrics the DIV counts the number of transmitted video frames with a lower MOS grade than before the transmission process. This is done for a given time window, which is shifted over the whole MOS vector.

In this thesis the threshold of 20% lost frames for the time interval of 20 seconds, as both defined in [21], is considered. Results for the DIV higher than the threshold are considered to lead to an unacceptable decrease of the video quality. In the DIV plots presented in the performance studies this is indicated by a horizontal line at the 20% mark. Although this is a very stringent metric, its representation of the grade of distortion caused by the transmission system is closer to the perception of the user than time-averaging metrics [35].

In addition to the DIV the packet loss rates for the transmitted video streams are considered in the performance studies. This is done separately for each video frame type. Compared to the DIV this illustrates the effect of error-propagation through the video stream. Furthermore it shows, whether the correct packets are removed by the filter of the VQMs.

For both video metrics the average over the number of video receiving WTs is shown in the performance studies.

5.5 Parameterization

In this section the chosen values for the parameters considered during the simulation are presented. At first the scenario independent parameters are shown. These parameters were left constant for all simulations which were calculated within the scope of this thesis. Secondly the values for those parameters are listed, which are chosen for the performance study of the dynamic OFDM-FDMA schemes in this chapter.

5.5.1 Scenario independent parameters

In Table 5.3 a selection of the applied parameters is shown, which were left constant during the whole simulation. The full set of the scenario independent parameters is shown in Appendix A.

The table starts at the lowest simulated layer the radio channel. As discussed in Section 2.2, the maximum speed and the delay spread have significant effects to the system due to the resulting fading and ISI. While the given speed is comparable to walking, the chosen value for the cell radius and the standard deviation of the delay spread represent an indoor environment [1]. On the physical layer $S = 48$ subcarriers are simulated in the 5.2 GHz band. Furthermore the bandwidth, the symbol time T_s and the transmission power are given. These values were chosen according to the IEEE 802.11a WLAN standard [28]. For the delay spread the standard deviation is given. Since mean values can be considered in the design of the system, the variation is the interesting parameter here. Since T_s is nearly 27 times higher than the standard deviation of the delay spread no significant distortion due to ISI has to be assumed. The SEP threshold P_s defines the SNR boundaries for the applied modulation techniques during the adaptive modulation. This value has been chosen from [53]. The length of the MAC frame T_f corresponds to HIPERLAN/2 [11], while the length of the downlink phase field results from the link layer model which was proposed in [23]. The maximum size of the TCP segments represents the maximum amount of data which is sent within a single segment. Due to the length of the data field in a frame of the very common Ethernet (IEEE 802.3) this is typical sent to 1460 Bytes. On the application layer, the session-inter-time defines the time between two consecutive WWW sessions. Although, this parameter is defined by the *Malaga* WWW traffic model [54] its value is not given. For this thesis it was chosen in order to provide considerable traffic load. If the session-inter-time is set to 0 s the resulting traffic could be interpreted a single (much longer) session. The MPEG-4 stream parameter result from the decoding of 180 s of a video in the CIF format with the size of 352×288 Pixel per video frame.

5.5.2 Parameters for the OFDM-FDMA performance study

For the traffic load model heterogeneous traffic is considered: Here a fixed number of WTs (J_h) receives Web-pages, while a number of WTs (J_v) simultaneously receives MPEG-4 coded VBR video streams. For each scenario and OFDM-FDMA scheme a series is recorded by varying J_v from 0 to $J_{max} - J_h$. Since with static subcarrier allocation the number of subcarriers one WT obtains is $N = S/J$, the maximum amount of WTs is chosen in order to provide at least one subcarrier per WT. Table 5.4 lists those parameters which are specific for this performance study.

Channel/Scenario parameter	Chosen value
Cell radius (R)	100 m
Maximum speed of the WTs (v_{max})	1 m/s
Delay spread standard deviation ($\sigma\Delta\tau$)	0.15 μs
Physical layer parameter	Chosen value
Total bandwidth (B)	16.25 MHz in the 5.2 GHz band
Number of subcarriers (S)	48
Max. transmit power/subcarrier (P_{tx})	0.2 mW (-7 dBm)
Symbol time (T_s)	4 μs
SEP threshold (P_s)	0.01
MAC timing parameter	Chosen value
MAC frame time (T_f)	2ms (500 Symbols)
Downlink phase, payload field time	0.968ms (242 Symbols)
Transport layer parameter	Chosen value
TCP Maximum segment size	1460 Bytes
WWW Traffic model parameter	Chosen value
Session-inter-time	$\mu = 3$ s
MPEG-4 video stream parameter	Chosen value
Mean bit-rate	951 KBps
Video duration	4500 frames, i.e. 180 s

Table 5.3: Selection of the applied scenario independent parameters

Since the focus of this performance study is not to investigate the VQM schemes, as proposed in Section 4.4.3 and Chapter 6, the simple VQM is applied for all the simulations of the heterogeneous scenario.

Scenario Parameter	Value
Maximum amount of WTs per cell (J_{max})	48
No. WWW receiving WTs per cell (J_h)	12
No. video receiving WTs per cell (J_v)	varied: [0 : 36]

Table 5.4: Scenario parameters as applied for the OFDM-FDMA performance study

5.6 Results for heterogeneous traffic streams

The following results are obtained for the heterogeneous scenario where in addition to the 12 WTs which receive HTTP, a variable amount of video receiving WTs (J_v) is considered. Since with higher amounts of J_v the limited amount of subcarriers has to be distributed to more and more WTs generally the average bit-rate decreases (Figure 5.3) for all considered OFDM-FDMA schemes. The highest gain for the bit-rate versus the static scheme is 80%, achieved for $J_v = 36$. The maximum gain can be found for the dynamic assignment which is with 55% much higher than the 25% obtained with the dynamic allocation. For $J_v \geq 24$ the gain achieved with the full dynamic approach decreases while it rises slightly for the semi-dynamic approach. The absolute values show the low capacity, which remains on the application layer. For $J_v = 0$ the Web-pages are sent to each of the 12 WTs with 166 KBps on the average, while this value decreases until 98 KBps are reached ($J_v = 48$).

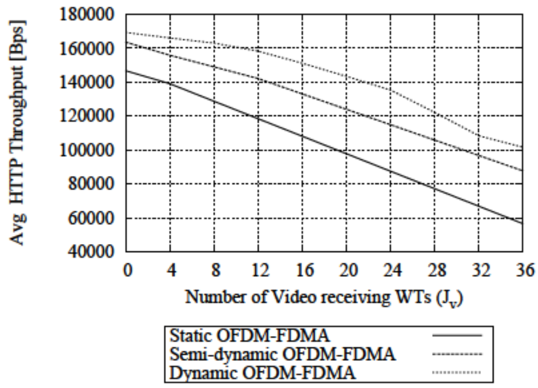


Figure 5.3: Average bit-rate of HTTP receiving WTs for different OFDM-FDMA schemes in a heterogeneous traffic load scenario

In Figure 5.4 the average transmission time of a Web-page per WT is shown. Due to the higher contention for the channel resource it rises with J_v. On the average the dynamic OFDM-FDMA scheme leads to the fastest transmission of a Web-page, followed by the semi-dynamic and by the static scheme. The maximum value is achieved at $J_v = 36$ for the static case. It is 100% higher than the value obtained for the dynamic OFDM-FDMA, where the component achieved with the dynamic assignment is 84%. As for the bit-rate a decrease of the performance gain achieved with dynamic allocation starting at $J_v = 24$ shows up.

While the standard deviation of the page transmission time (Figure 5.5) basically shows the same qualitative behavior as the page transmission time, the highest gains are achieved for this metric. Here the maximum benefit achieved at $J_v = 36$ for dynamic OFDM-FDMA is 215%, while 161% are achieved by the dynamic assignment. The gain achieved with the dynamic allocation of subcarriers clearly slopes for $J_v \geq 24$.

The investigation of the DIV metric, which expresses the amount of distortion caused by the transmission system to the video stream, is interesting since it shows the effect of the choice of the OFDM-FDMA scheme has to the video quality. In Figure 5.6 the positive impact is significant for the DIV results obtained with dynamic OFDM-FDMA for low numbers of video receiving WTs. For example, for $J_v = 4$ the DIV is 37% lower with the dynamic approach and 25% lower with semi-dynamic OFDM-FDMA. If J_v rises, the DIV also rises for all transmission schemes. The horizontal line represents the upper boundary for the DIV. It marks

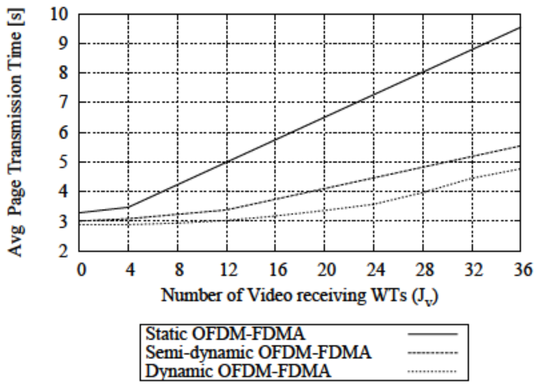


Figure 5.4: Average transmission time per Web-page for different OFDM-FDMA schemes in a heterogeneous traffic load scenario

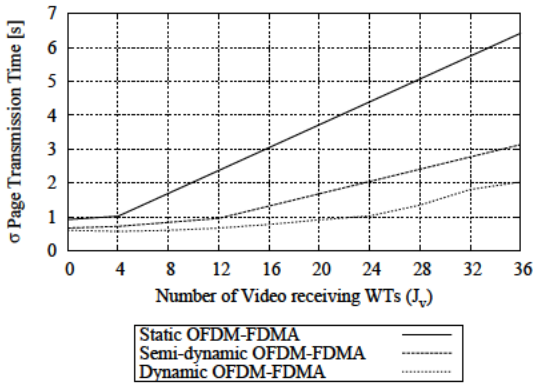


Figure 5.5: Standard deviation of the transmission time per Web-page for different OFDM-FDMA schemes in a heterogeneous traffic load scenario

the level of 20%, since larger DIV values result in an unacceptable video quality [21]. While for static OFDM-FDMA all results stay above this line, the upper boundary is reached at $J_v = 8$ for the semi-dynamic and at $J_v = 17$ for the dynamic approach.

Since most of the users will give up watching the video stream, video scenarios with permanent high DIV rates are not typical. However, with best-effort networks as the Internet, where requests are not rejected on the transmission path, such situations may occur especially on links with low capacity. Thus, the chosen scenarios with high J_v is not unrealistic for the considered wireless system.

For the performance study in this chapter the high load can also be interpreted as any other type of VBR bulk traffic, e.g. the transmission of files or real-time audio streams, which is in contention to the WWW traffic.

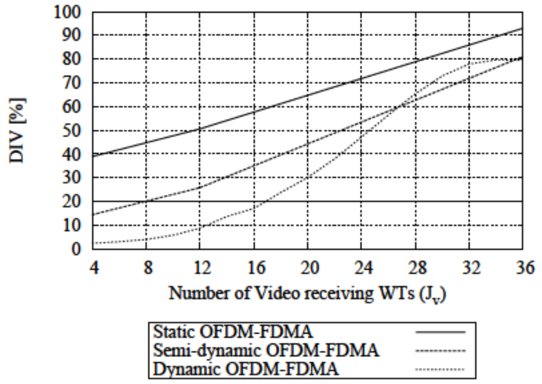


Figure 5.6: DIV for different OFDM-FDMA schemes in a heterogeneous traffic load scenario

5.7 Interpretation

The investigation of the HTTP transmission metrics shows that static OFDM-FDMA is clearly outperformed by the dynamic assignment of subcarriers with semi-dynamic OFDM-FDMA. The exploitation of the statistical multiplex with the dynamic subcarrier allocation leads to an additional performance increase, achieved with dynamic OFDM-FDMA. The maximum achieved gains are compared in section 5.7, where the highest gain of 215% is achieved for the standard deviation of the Web-page transmission time. Considering all HTTP metrics leads to the conclusion that for all HTTP metrics the maximum benefit is obtained if high traffic load is simulated using high numbers of WTs in the cell. The major part of the gain is typically achieved by the dynamic subcarrier assignment. Adding the dynamic allocation the maximum is reached with the full dynamic OFDM-FDMA approach for the highest simulated number of WTs.

However, this second aspect has to be further examined. While this is clearly present with the semi-dynamic OFDM-FDMA, where the gain rises with J_v, a different behavior is shown for dynamic OFDM-FDMA: With this scheduling scheme the gain starts to sink faster from a specific point ($J_v > 24$) in the heterogeneous scenario. The reason for this is the decreasing amount of subcarriers which can be assigned. As given in Equation 4.5 with dynamic assignment each WT which has data in its access point queue achieves at least one subcarrier. Assuming that this is the case for all WTs, with $J_h + J_v = 24$ this lasts to 12 freely allocatable subcarriers, while with $J_v = 36$ no subcarrier can be allocated. The gain which is still achieved for the highest J_v shows that not for all WTs is data in the queues during the whole simulation. However, according to the queue length the probability that a WT receives data rises. Despite from the sender bit-rate the queue length depends on the data rate on the wireless link, which is lower with less amounts of subcarriers allocated per WT. Assuming a constant bit-rate, this leads to a vicious circle for high number of terminals: Less amounts of allocatable subcarriers are followed by rising queue lengths which are followed again by less amounts of allocatable subcarriers.

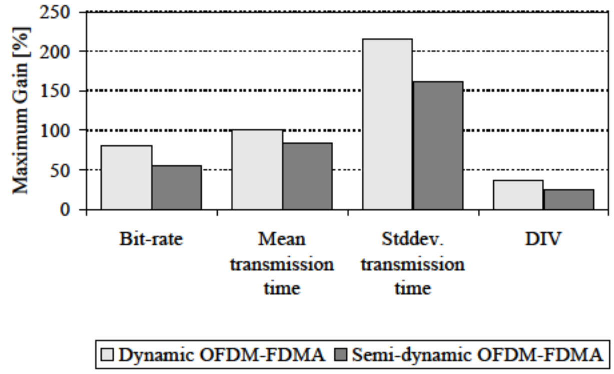


Figure 5.7: Maximum gains achieved with dynamic and semi-dynamic OFDM-FDMA in comparison to static OFDM-FDMA in a heterogeneous traffic load scenario

The influence of scheduling schemes to the video quality is expressed by the DIV rate. The comparison of the DIV rates for all the applied OFDM-FDMA schemes leads to the conclusion that for setups with small numbers of video receiving WTs, which means small load to the transmission system, dynamic OFDM-FDMA outperforms the static for up to 37% and the semi-dynamic scheme for up to 25%. Although in setups with high traffic load the DIV is slightly lowered this is not beneficial. Due to the high traffic load the distortion of the video is generally too high for being acceptable for the user ($DIV > 20\%$). As for the HTTP metrics, the DIV rate gain achieved with the dynamic subcarrier allocation decreases more clearly for higher traffic load than the gain achieved with semi-dynamic OFDM-FDMA.

Chapter 6

Adaptive video queue management

In the preceding chapter it was shown, that for the heterogeneous scenario the highest performance can be achieved with the combined dynamic OFDM-FDMA approach. Although this adaption to the channel states and to the traffic load leads to performance gains of up to 215% for the transmission of HTTP, there are still problems. Especially for setups with high traffic load the high queuing delays at the access point lead to significant drawbacks for the transmission of data. With heterogeneous traffic here the performance gain achieved with the improved subcarrier allocation decreases. Although this is the case for video and HTTP especially the HTTP related data streams suffer.

In this chapter a Video Queue Management (VQM) scheme is introduced which directly addresses the problems of heterogeneous traffic scenarios under high load in combination with dynamic OFDM-FDMA. This scheme, whose basic idea is given in the next section, is called *adaptive VQM*. Its algorithms and parameters are explained in greater detail in the second section of this chapter.

6.1 Basic idea

In Section 4.4.3 the *simple VQM* using a FIFO scheduling policy and the *semantic VQM*, which adapts its scheduling policy according to video semantics, were discussed. One important aspect of the semantic VQM is, that the removal of semantically less important parts of the video stream, e.g. B-frames, provides additionally capacity to more important parts such as I-frames. This is essential in scenarios with high traffic load, which occurs in one of the following cases:

- Many receiving WTs: As discussed in Chapter 5 the performance gain achieved with the dynamic allocation scheme drops due to the lack of subcarriers.

- Data streams with a high bit-rate: The queue length for the affected WTs rises. Due to dynamic allocation these WTs receive higher amounts of subcarriers.

- Bad channel states: Due to low subcarrier CNR only small amounts of bits can be sent during T_s.

Since the resources of the radio channel are limited, congestion may occur which leads to higher transmission delays, regardless which type of application layer data is transported. Due to its strict timing boundaries (Section 4.4.3) especially real-time video transmission will suffer: The rising transmission delay of the IP datagrams, containing parts of video frames, leads to more and more datagrams which will hit the deadline. Therefore they are discarded, either by the simple or the semantic VQM. In this chapter a further stage of the semantic VQM, called *adaptive VQM*, is introduced. The basic idea of this VQM is to free capacity in the case of high traffic load by the preemptive removal of packets from the video queues. In addition to the consideration of semantic information which is extracted from the video packets the scheduling policy is adapted to the amount of traffic load, the system estimates. The objective of this scheme is to avoid congestion in the access point queues, from which all types of application layer traffic will suffer.

In the following section all parts of the adaptive VQM are introduced and it is explained how they work together. Furthermore, aspects as traffic load estimation, adaptive video filters and the related parameters and algorithms are discussed. In Chapter 7 the performance of the adaptive VQM in comparison to the simple and the semantic VQM is investigated.

6.2 Optimization approach

The adaptive VQM provides the adaption of the scheduling policy to the traffic load on the access point. To fulfill this task basically two new components are added to the plain semantic VQM: The first has to estimate the traffic load to the monitored system and is called *system load estimator*. The second component, called *adaptive video filter* works on the video queue under the objective to lower the bit-rate of the contained video stream while assuring small decreases of the video quality. Therefore, semantic parameters, regarding to the perceived quality at the application layer, are considered by the filter algorithm. The steps of the adaptive VQM are listed in Table 6.1. In general the estimation of the traffic load can be treated as a simple accounting calculation at the access point, i.e. the amount of free resources is compared to the needed resource capacity:

$$\frac{\text{Bits to transmit}}{\text{Bits transmittable per } T_f} \tag{6.1}$$

While the upper part of the fraction always changes with VBR traffic, changes in the denominator are specific for dynamic scheduling systems. If the result of Equation 6.1 is greater than 1 a lack of resources occurs for the considered frame of length T_f. The problem of this approach is, to decide which amount of Bits is

Step	VQM Component
1	Video packet lifetime calculator
2	Sorting of the packets according to lifetime
3	Non-load-adaptive packet removal
4	Traffic load estimator
5	Adaptive video filter

Table 6.1: Basic steps of the adaptive VQM

considered in the upper part of the fraction. Neither the full length of the queue nor the mean number of Bits send-able per T_f will produce useful results. Therefore, in the adaptive VQM a different approach is used, which is presented in the following discussion of the basic algorithm. As shown in Figure 6.1, in the first step the calculation of the *lifetime-index* (L_p) is done for all video packets in the queue. This is done exactly as shown in Equation 4.7 for the plain semantic VQM. Then all the video packets in the queue are sorted according to L_p (step 2). This assures that those packets with the smallest lifetime-index are sent first, considering that for the calculation of L_p, also the frame type is used. After the queues are sorted the basic algorithm of the semantic VQM (Figure 4.4) is performed. This third step of the adaptive VQM is called *non-load-adaptive packet removal* and is marked by a dashed line in Figure 6.1. During this iteration over all the packets in the queue all packets with $L_p < 1$ are removed. The size of the removed packets is totalized independently for each WT to $S_j(t)$. This value is an indirect measure of the needed capacity according to Equation 6.1. It is used during step 4 for the estimation of the traffic load. Here it is compared to the free downlink capacity of the MAC frame from the preceding downlink phase ($F_j(t-1)$), as given in Equation 6.2.

$$R_j(t) = \frac{S_j(t)}{F_j(t-1)} \tag{6.2}$$

Since the VQM runs prior to the dynamic OFDM-FDMA schemes only the value measured for F_j in the preceding downlink phase can be used for the comparison. The $S_j(t)$ value already expresses the result of the traffic load to the system. If $S_j(t)$ is high, many video packets, or those related to more important frame types of greater size, exceed their deadline. If this is combined with a small estimated downlink capacity ($F_j(t-1)$), then additional filtering of the video queues will lower the probability of congestion and therefore, the S_j in the next cycle. If the resulting load-rate $R_j(t)$ exceeds a given threshold α, then the access point is considered to be overloaded and the adaptive video filter is executed.

The objective of the adaptive video filter is to free system capacity by lowering the bit-rate of the video streams within the access point queue. For this purpose several types of video filters, e.g. *re-quantization-*, *low-pass-* or *color-to-monochrome-filters* were proposed for the usage with streaming data in wired networks [65]. However, the complexity of these filters might be problematic for the access point. Typically the access point has to serve many WTs and is still

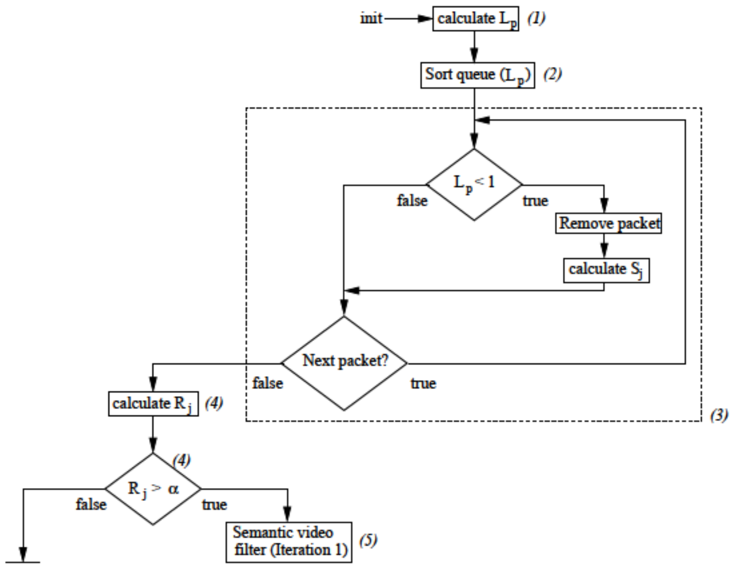


Figure 6.1: Basic algorithm of the adaptive VQM as performed independently for each queue

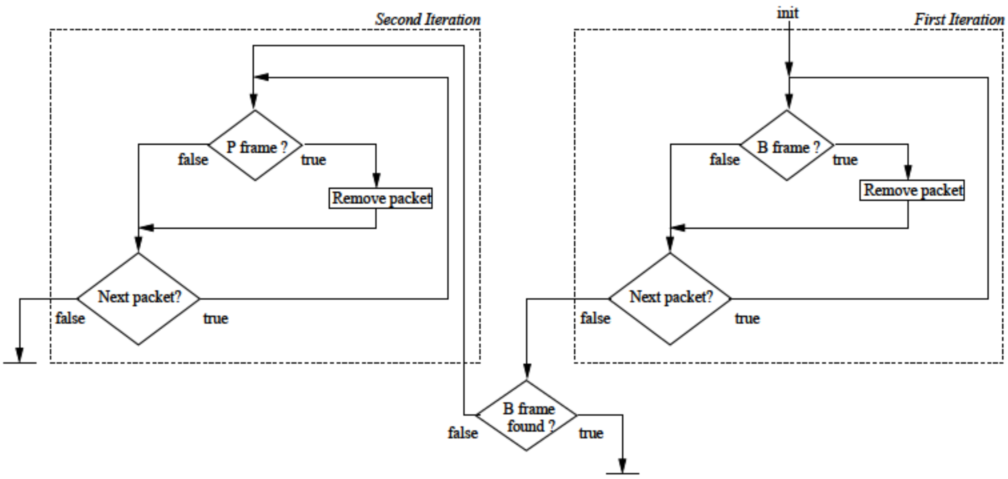


Figure 6.2: Algorithm of the adaptive video filter as invoked by the system load estimator independently for each queue

in a high-load situation if the filter is performed. Therefore, the simpler class of *packet-drop* filters is chosen for the adaptive VQM. Packet-drop filters delete packets which are related to video frames due to enable adaption.

The filtering algorithm of the adaptive VQM considers again the type of the video frame to which the packet is related. As shown in Figure 6.2 it consists of two iterations: In the first iteration the filter inspects the whole queue regarding to the frame types and deletes all packets which are related to B-frames. The algorithm returns if it reaches the head of the queue. If no B-frame is found during the first iteration, a second iteration starts, where the filter deletes all P-frames in the queue. Although the removal of all packets in the queue may cause drastic distortion effects to the video, this has one benefit: If the queue length rises, typically the risk of congestion rises. Therefore, it is profitable to delete more packets if the queue is longer. Due to the strong impact to the quality of the video (error propagation) no I-frames related data is removed by the algorithm. Thus, the filter provides the explicit, preemptive prioritization of I-frame related packets if high load occurs. Compared to the semantic but reactive strategy this approach should minimize the removal of I-frame related packets due to exceeded deadlines. In order to lower the complexity, the filter does make no use of further semantic parameters in the video stream. For example, considering the GOP structure the second iteration could be modified in order to only remove every second P-frame per GOP. On the one hand, this would provide a finer granularity of the changes to the video quality. On the other hand, this results in more calculation complexity due to the maintenance of a GOP index for each packet. Keeping the algorithm complexity as small as possible allows the execution of the adaptive VQM every T_f. This is not essential but will provide the most actual information to the load estimator.

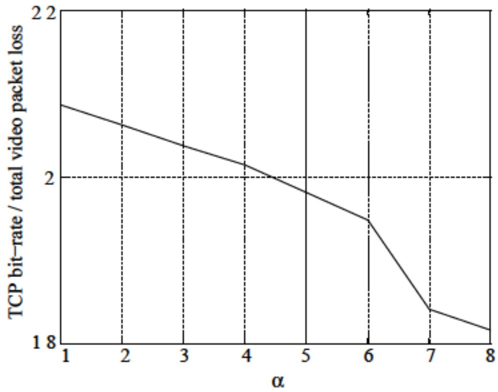


Figure 6.3: The selection of α considering HTTP and video metrics

The choice of the threshold α is a critical point of the adaptive VQM. Its value is determined by the comparison of the TCP bit-rate to the overall packet loss for a varying α, as shown in Figure 6.3. Since with rising α less packets are removed by the adaptive video filter the plotted fraction of bit-rate over packet loss drops. From $\alpha = 1$ to 5 this is nearly linear followed by higher negative slopes for $\alpha \geq 6$.

Due to these results, obtained for a scenario with $J_h = J_v = 12$ and dynamic OFDM-FDMA, the value $\alpha = 5$ is chosen. This value is used for all the simulated scenarios as shown in Chapter 7 and represents not more than a heuristic one has optimize. However, the adaption of α is not considered in this thesis.

Chapter 7

Performance study of the proposed queue management method

The main goal of this study is to investigate the performance of the adaptive VQM, which is designed to optimize the transmission of heterogeneous traffic. Section 5.6 shows, that for this type of traffic load the highest gain can be achieved with dynamic OFDM-FDMA. Therefore, it is of particular interest how the adaptive VQM performs in combination with dynamic OFDM-FDMA. After the methodology of this investigation is discussed in the next section, the parameters for the simulated scenarios are listed in Section 7.2. The presentation of the simulation results is separated into two parts: At first, in Section 7.3 the performance of the adaptive VQM is analyzed in combination with three different OFDM-FDMA schemes for scenarios with a fixed number of HTTP receiving WTs per cell. Secondly, in Section 7.4 the combination of dynamic OFDM-FDMA and adaptive VQM is investigated for scenarios with varying amounts of video and HTTP receiving WTs in the cell.

7.1 Methodology

Although in this performance study basically the same procedures and methods are applied as introduced in Chapter 5, several differences have to be discussed. For the combination of the adaptive VQM with dynamic OFDM-FDMA the performance is analyzed by the use of two methods for defining the structure of the heterogeneous traffic: At first, in Section 7.3.1, the number of Web-page receiving WTs (J_h) is left constant, while the amount of WTs receiving video (J_v) is increased. This method is known from the former performance study given in Chapter 5. Therewith, the traffic load from setup to setup increases due to the higher amount of simultaneous received video streams per cell. Secondly, to provide a better evaluation of the influence different traffic structures have to the performance of the VQM, J_v and

J_h are changed from setup to setup, while the ratio

$$J_r = \frac{J_v}{J_h} \tag{7.1}$$

is fixed. For varying J_v this leads to a fixed proportion of HTTP receiving WTs per cell. For example, 25% of the WTs per cell are receiving video streams if $J_r = 1/4$. An investigation for several J_r is done for dynamic OFDM-FDMA in Section 7.4, where the results are shown versus J_v. Since the maximum amount of WTs per cell (J_{max}) is limited the highest possible J_v usable per simulation is

$$J_v = J_{max} - \frac{J_h}{J_r}. \tag{7.2}$$

In addition to the investigations for dynamic OFDM-FDMA, the dependency of the VQM from the FDMA scheme is evaluated. Therefore, the two combinations adaptive VQM with static OFDM-FDMA (Section 7.3.2) and adaptive VQM with semi-dynamic OFDM-FDMA (Section 7.3.3) are analyzed.

In order to evaluate the performance of the adaptive VQM the simulation results, obtained for the adaptive VQM, are compared to those results, obtained for the simple and the semantic VQM. This is done for all metrics which are defined in Section 5.4 for HTTP and video separately. For the HTTP receiving WTs, the metrics average HTTP bit-rate, average Web-page transmission time and standard deviation of the Web-page transmission time are considered. The quality of the MPEG-4 coded video streams is evaluated by the use of the DIV metric. Additionally the loss rates of video related UDP packets are shown for each MPEG frame type to illustrate the influence of the different VQM to the video streams.

In the investigations with constant J_h, for each of the mentioned metrics the mean values, obtained for rising numbers of J_v are shown. However, no absolute results are shown if the VQM are analyzed for several J_r. Since here various setups are investigated, only the performance gain, which is achieved with the adaptive VQM in comparison to the simple VQM (Section 7.4.1) and to the semantic VQM (Section 7.4.2), is shown. This is done for rising amounts of J_v for all HTTP and video metrics except for the video packet loss rates.

7.2 Parameters for the VQM performance study

The chosen scenario is basically the same as for the performance study of the OFDM-FDMA schemes in Section 5.5. However, the composition of the traffic structure is modified. In the first case J_h is set to a fixed number of 12 WTs, while J_v is varied from 1 to 36. Secondly, scenarios with varying J_h and J_v are examined, while the ratio J_r (Equation 7.1) is constant. According to Equation 7.2 J_v is raised until J_{max} is reached. For example, with $J_r = 1/4$ every increment of J_v results in 4 additional HTTP receiving WTs, which leads to the maximum of $J_v = 9$ and $J_h = 36$. For J_r the values as listed in Table 7.1 are considered. For all further modeling parameters the values shown in Section 5.5 are applied.

Scenario Parameter	Value
J_{max}	48
J_h if constant	12
J_v if J_h is constant	varied $[1:36]$
J_r	$1/4, 1, 4$

Table 7.1: Scenario parameters as applied for the VQM performance studies

7.3 Analysis for a fixed number of HTTP traffic streams

In this section the performance of all VQM is compared in combination with the dynamic, the static and the semi-dynamic OFDM-FDMA scheme. Assuming a fixed J_h, the simulation results for the simple, the semantic and the adaptive VQM are shown versus rising J_v. The main topic of the discussion is to expose the influence of the different VQM methods to the Web- and video metrics, whereby the aspects of load-adaption and filtering are considered in particular. While in the following sections several details of the results are discussed in addition to the presentation, a discussion considering all results is finally given in Section 7.3.4.

7.3.1 Results for dynamic OFDM-FDMA

The bit-rate of the average TCP throughput, achieved for all the Web-page receiving WTs, in Figure 7.1 shows a clear benefit for the adaptive VQM. While with higher traffic load due to higher amounts of video receiving WTs in the cell (J_v) the throughput rate strongly decreases for the simple and the semantic VQM, the decrease is much smaller with the adaptive VQM. For the case $J_v = 36$, where the highest traffic load is produced, the adaptive VQM outperforms the semantic VQM for 24% and the simple VQM for 32% for the bit-rate.

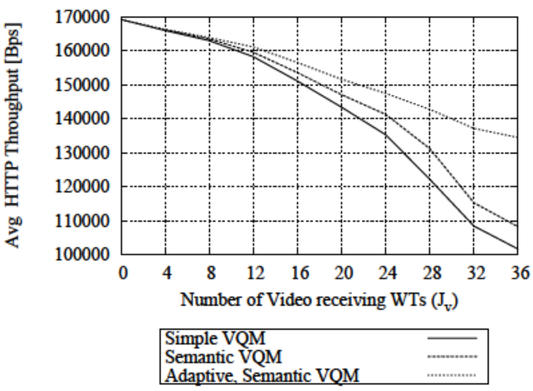


Figure 7.1: Average bit-rate of HTTP receiving WTs for different video queue management methods with dynamic OFDM-FDMA

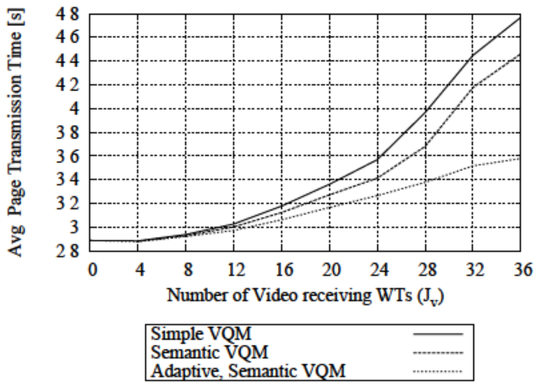


Figure 7.2: Average transmission time per Web-page for different video queue management methods with dynamic OFDM-FDMA

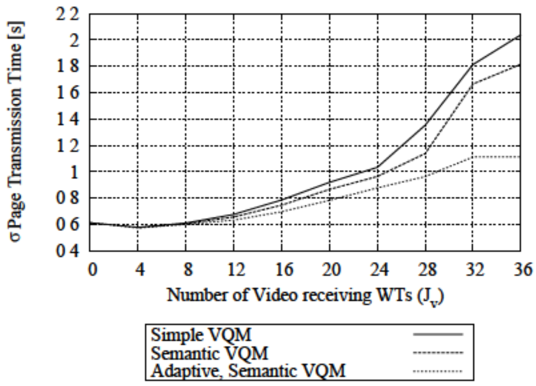


Figure 7.3: Standard deviation of the transmission time per Web-page for different video queue management methods with dynamic OFDM-FDMA

The average transmission time per Web-page for rising J_v is shown in Figure 7.2. As for the TCP throughput the adaptive VQM outperforms the other VQM schemes. For J_v values greater than 8 the simple VQM leads to the longest page transmission times, followed by the semantic VQM. The highest gain by the adaptive VQM is achieved when the maximum traffic load is simulated as for $J_v = 36$. Compared to the semantic VQM it is 25% lower and 33% lower in comparison to the simple VQM.

The plot of the standard deviation of the page transmission time (Figure 7.3) shows a similar quality as the page transmission time. Therefore, in case of higher J_v, the standard deviation is the lowest with the adaptive VQM. For the maximum of J_v this leads to the gain for the predictability of 63% in comparison to the semantic and of 83% in comparison to the simple VQM.

Considering the additional video filtering as performed in the adaptive VQM, it is of great interest, how this impacts the video quality. Figure 7.4 shows the average DIV, which is a metric for the amount of distortion the transmission system causes to the video, as perceived by the user. Therefore, higher DIV values express higher

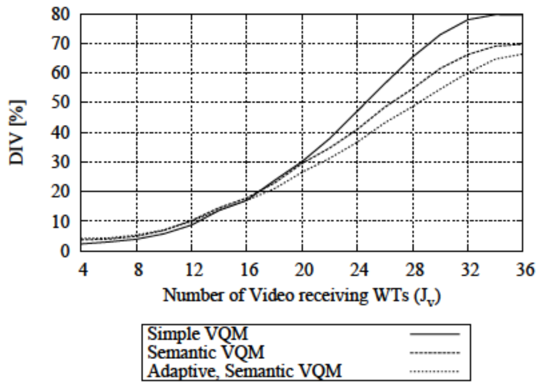


Figure 7.4: DIV for different video queue management methods with dynamic OFDM-FDMA

amounts of distorted video frames. Generally for rising amounts of video receiving WTs (J_v) the DIV value rises. Compared to the simple VQM, it is interesting that for J_v smaller than 16 the DIV is slightly higher with the semantic and the adaptive VQM (up to 2% for $J_v = 4$). For both schemes the results for J_v greater than 16 are lower and therefore better than the DIV for the simple VQM. However, here the upper boundary for the DIV of 20% [21] as marked by the horizontal line is crossed by all the VQM. Results for the DIV higher than this threshold are considered to lead to an unacceptable decrease of the video quality. This leads to the conclusion that with all VQM not more than 16 video receiving WTs can be supported. However, in cases of high load, the semantic and the adaptive VQM lead to benefits for the HTTP transmission metrics.

The video packet loss rates for the different VQM as shown in Figure 7.5 illustrate the influence of the video filter as applied in the adaptive VQM. In Figure 7.5(a) the total percentage of the lost video packets, regardless to the video frame type, is shown. Although for $J_v < 26$ the simple VQM leads to the smallest loss rates, compared to the other VQM schemes, more errors are caused by lost packets to the semantically more important I- and P-frames (Figure 7.5(b) and 7.5(c)). This is especially the case with higher numbers of video receiving WTs, where even the total loss rate of the simple VQM exceeds those of the semantic and the adaptive VQM. For example, for $J_v = 36$, 4% more I- and 6% more P-frame related packets are lost with the simple VQM than with the adaptive VQM. The behavior of the DIV for high J_v, can be explained by comparing the B-frame loss rates of the VQM (Figure 7.5(d)). Since here the packet loss rate is up to 10% higher for the semantic approach, it is shown wherefrom the capacity to prioritize the transmission of I- and P-frames is taken. Considering the absolute values shows that the most lost packets are related to B-frames. The lower DIV, which is achieved by the adaption to the traffic load in comparison to the plain semantic VQM, can be explained by the packet loss rates for the scenarios for J_v greater than 24. Compared to the values obtained with the semantic VQM the percentage of the removed I- and P-frames with the adaptive VQM is lower (Figure 7.5(b) and

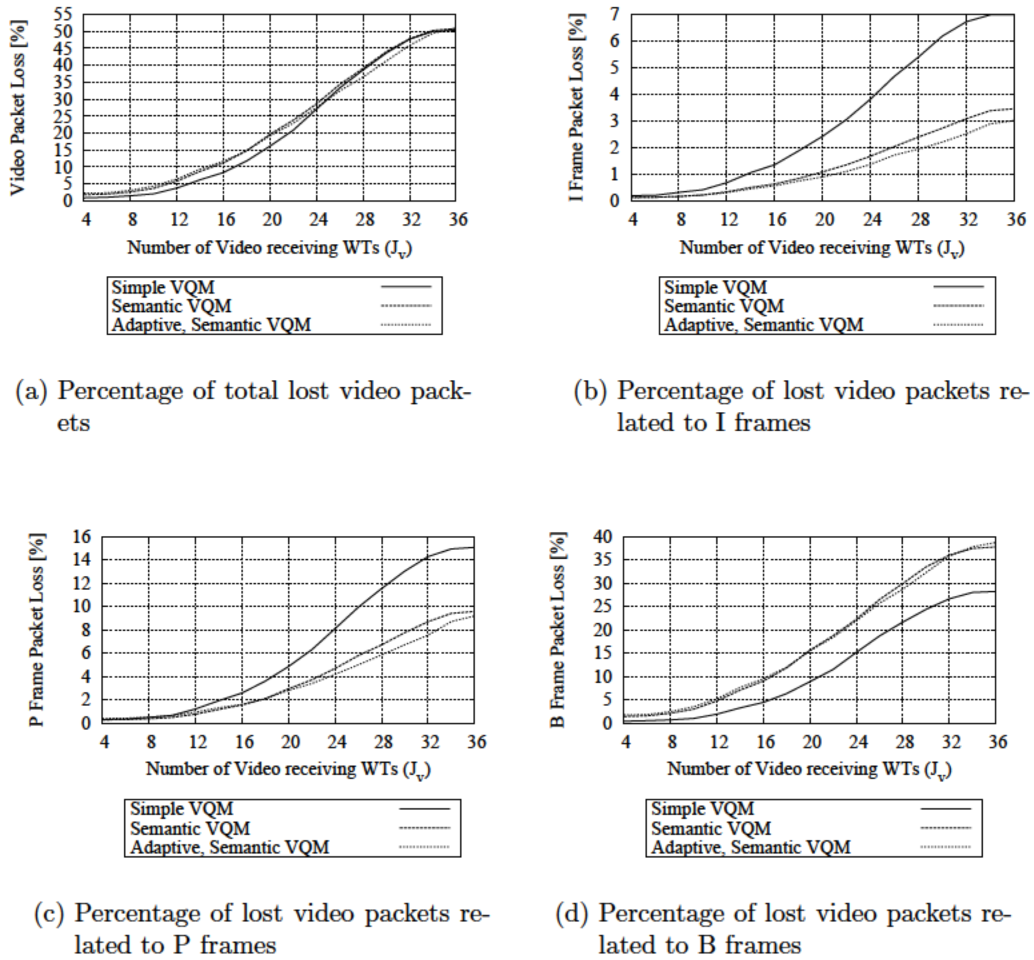


Figure 7.5: Video packet loss rates for different video queue management methods with dynamic OFDM-FDMA

7.5(c)). Therefore, with high traffic load the preemptive removal of video packets leads to less errors in the semantically more important frames. Due to error propagation a single error in an I- or P-frame affects a high number of frames within the video stream.

7.3.2 Results for static OFDM-FDMA

In contrast to the HTTP throughput gain, which can be achieved in combination with dynamic OFDM-FDMA, the employment of the adaptive or the semantic VQM in combination with static OFDM-FDMA leads to no benefit for the HTTP receiving WTs. The results for the TCP bit-rate presented for the three examined VQM in Figure 7.6 are basically the same.

This behavior is also presented by the results for the standard deviation and the mean of the transmission time per Web-page. As shown in Figure 7.7 and

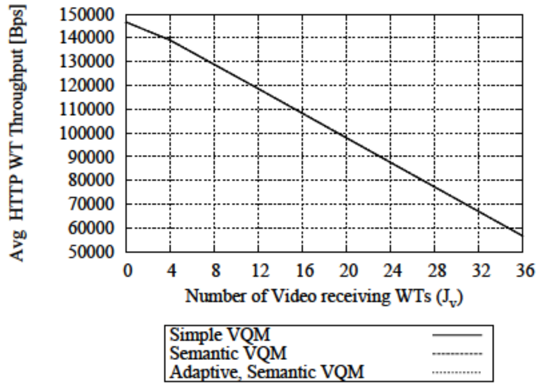


Figure 7.6: Average bit-rate of HTTP receiving WTs for different video queue management methods with static OFDM-FDMA

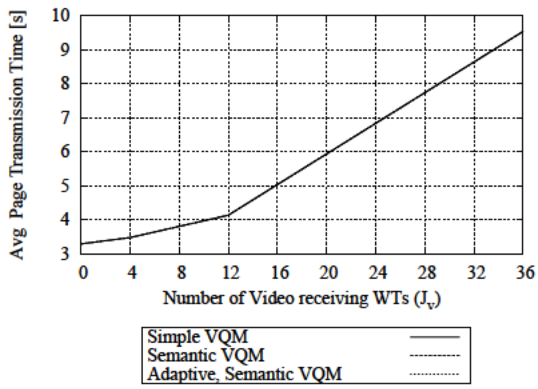


Figure 7.7: Average transmission time per Web-page for different video queue management methods with static OFDM-FDMA

Figure 7.8 no gain can be achieved for the transmission of Web-pages if the semantic or the adaptive VQM is applied.

This is not the case regarding the results for the video metrics. The simulation results obtained for the DIV (Figure 7.9), basically show the same tendency than the results for dynamic OFDM-FDMA: For all VQM the DIV values rise for higher amounts of video streams in the cell. Since the minimum DIV of 37% is larger than the given threshold of 20% static OFDM-FDMA is clearly the worst choice for this scenario, regardless which VQM is applied. In addition to the 12 HTTP receiving WTs no video stream can be transmitted in acceptable quality. However, the behavior of the VQM in combination with static FDMA is interesting, since it shows the influence of the filtering process. The difference between the semantic and the simple VQM rises with J_v until the maximum of 6% is reached for $J_v = 36$. It rises slightly faster for the semantic VQM. For $J_v < 16$ the adaptive and the semantic VQM results differ for up to 1% from the results for the simple VQM. In contrast to the dynamic OFDM-FDMA case, where also a slight decrease is achieved with both semantic VQM schemes, here the plain semantic VQM slightly

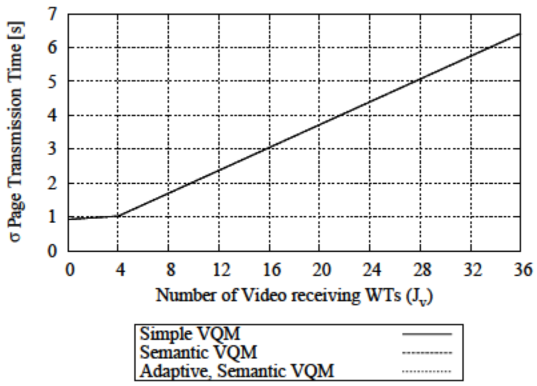


Figure 7.8: Standard deviation of the transmission time per Web-page for different video queue management methods with static OFDM-FDMA

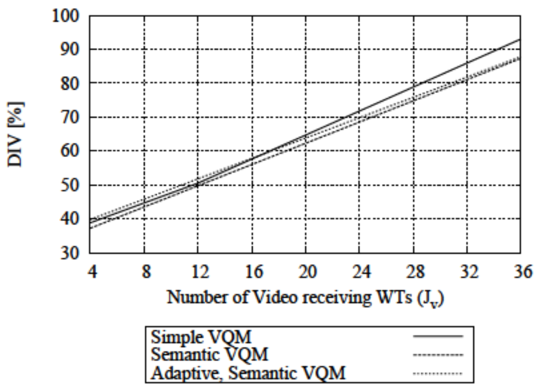


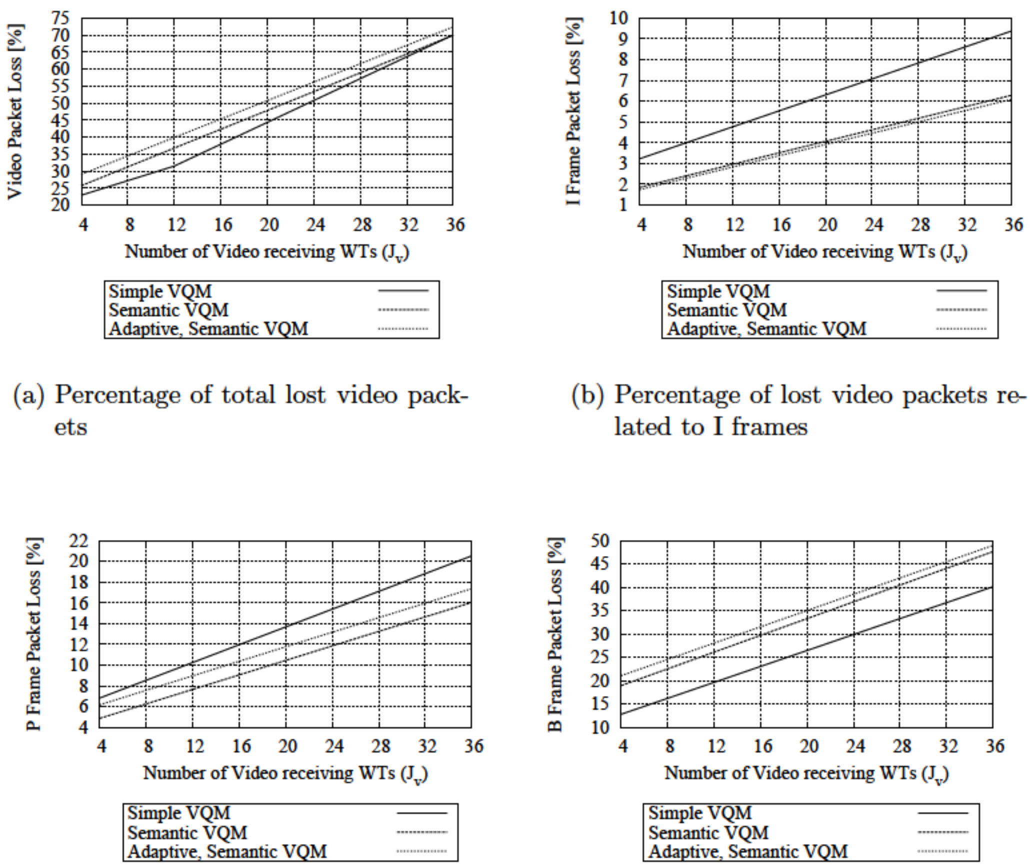
Figure 7.9: DIV for different video queue management methods with static OFDM-FDMA

outperforms the adaptive strategy for low J_v.

A further investigation is possible by taking a look at the packet loss rates in Figure 7.10. The percentage of total lost video packets is the highest for the adaptive VQM followed by the semantic approach. As discussed for dynamic OFDM-FDMA in section 7.3.1 higher total loss rates do not necessarily lead to higher DIV values. The packet loss rates for the semantically more important I and P frames have to be considered in particular. The highest amount of packets related to I frames is lost with the simple VQM (9% for $J_v = 36$). The I frame loss rates achieved with the semantic approach are much lower: Up to 6.3% of the transmitted video packets are lost if the semantic VQM is applied, while the adaptive approach results into a slightly lower packet loss of 6%.

The loss rates for the P frame related packets rise for all VQM with increasing J_v. The highest loss rate of 21% and the highest slope are obtained with the simple VQM. It is followed by the results for the adaptive VQM with a maximum P frame packet loss of 17% and by the semantic VQM, whose maximum loss is slightly lower. While the difference between the simple VQM and the other two VQM slopes for low numbers of WTs it rises for the adaptive VQM in comparison to the plain semantic approach. Here up to 2% more P frame related packets are lost with the adaptive VQM than with the semantic VQM. This explains the higher DIV rate, in comparison to the semantic VQM, with the adaptive VQM for small J_v. Comparing the results in Figure 7.9 and Figure 7.10(c) it shows again, how the packet loss rate for the semantically important P frames affects the DIV. For higher J_v the loss rate for the simple VQM shows a higher increase than those for adaptive VQM. This directly results into a higher DIV increase for the simple approach.

The loss rate for the B frame related packets in Figure 7.10(d) show, that the adaptive VQM leads to the highest packet loss rate of up to 49%, which is 9% higher than the loss rate, achieved with the semantic VQM. In contrast to the combination with dynamic OFDM-FDMA, with static FDMA this offset is constant for all J_v. Since, as for the P frames, the amount of removed packets does not depend on J_v, no load-adaptive removal of P and B frame related packets is performed by the adaptive filter.



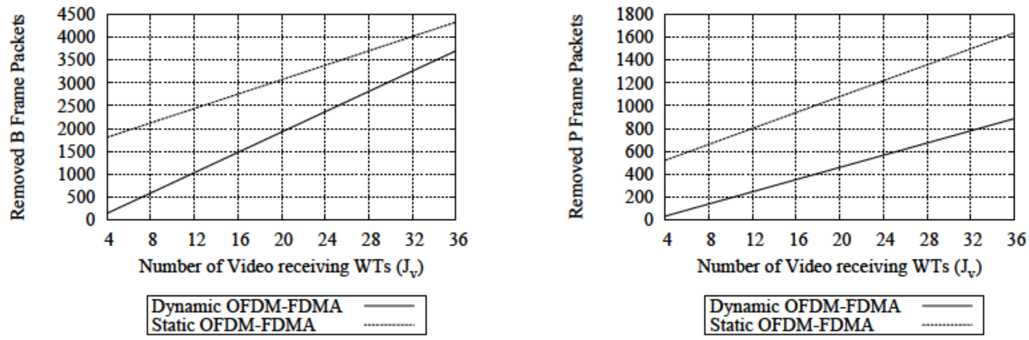
(a) Percentage of total lost video packets

(b) Percentage of lost video packets related to I frames

(c) Percentage of lost video packets related to P frames

(d) Percentage of lost video packets related to B frames

Figure 7.10: Video packet loss rates for different video queue management methods with static OFDM-FDMA



(a) Amount of B frame related packets removed due to exceeded semantic deadlines

(b) Amount of P frame related packets removed during the second iteration of the adaptive filter

Figure 7.11: Average number of removed packets per WT with adaptive VQM for two OFDM-FDMA schemes

The reason for the higher P frame loss rates with the adaptive approach in comparison to the semantic VQM can be found by analyzing the loss rates separately for the iterations of the adaptive filter (Figure 6.2). As illustrated in figure 7.11(a), with static FDMA up to 1136% more B frame related packets exceed the weighted deadline for $J_v = 4$ and are, therefore, removed in step 3 of the adaptive VQM (Figure 6.1). These packets can not be considered by the adaptive filter in step 5 of the VQM. This leads to a more frequently invocation of the second iteration of the filtering algorithm, where P frame packets are removed from the queue. Thus, as shown in figure 7.11(b), 1428% more P frame related packets are removed during the second iteration with static FDMA for $J_v = 4$, while due to the more frequent execution of the adaptive filter this percentage rises for higher J_v. This explains the high P frame packet loss rate with the adaptive approach, which results into a more drastic decrease of the video quality than the removal of B frame related packets.

From the results for the HTTP metrics it follows that in combination with static OFDM-FDMA no benefit can be achieved with both semantic VQM for the transmission of Web-pages. There, are two reasons for this: Firstly, due to the general higher capacity limitation compared to dynamic OFDM-FDMA the earlier removal of packets considering MPEG semantics leads only to a small amount of freed capacities. This leads to the discussed drawbacks for both types of traffic streams. The second aspect is, that the obtained capacity can not be used for the transmission of Web-pages by TCP. This can be explained by taking the average number of TCP retransmission timeouts per WT into account which is shown in Figure 7.12. Here, the influence of the OFDM-FDMA scheme to TCP becomes clear. With lower capacities, which means rising J_v as well as less effective scheduling schemes, the amount of timeouts rises significantly. For example for $J_v = 36$

the amount of timeouts occurring with the static scheduling is 14 times higher than with the semi-dynamic scheme. Furthermore, it is shown that for dynamic OFDM-FDMA no timeouts occur – regardless which amount of traffic is loaded to the system.

In Section 3.1 it has been discussed how TCP reacts if a retransmission timeout occurs. After the significant decrease of the bit-rate, the slow-start algorithm is performed. Due to the halved congestion window threshold the following linear congestion avoidance phase leads to a slower increase of the bit-rate than before the timeout. Considering the extended congestion avoidance phases, which become longer the more timeouts occur, it becomes clear why TCP is not capable to profit from the resources freed by the adaptive VQM if retransmission timeouts occur.

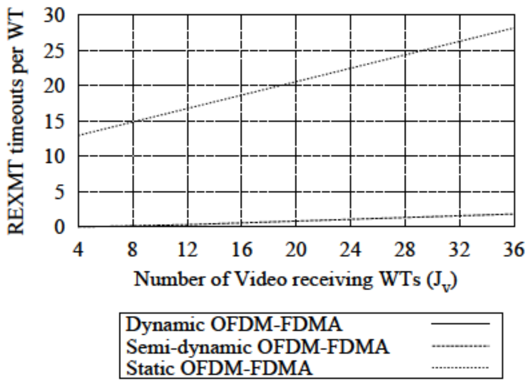


Figure 7.12: Average number of TCP retransmission timeouts for 12 HTTP receiving WTs for different OFDM-FDMA schemes

7.3.3 Results for semi-dynamic OFDM-FDMA

The investigation of semi-dynamic OFDM-FDMA in combination with the three VQM leads to the same qualitative behavior of the results for the HTTP metrics as the investigation of the static FDMA scheme. Since no explicit difference between the VQM can be found for the TCP throughput (Figure 7.13) and for the mean and the standard deviation of the Web-page transmission time (Figure 7.14, Figure 7.15), no gain is achieved with the semantic or the adaptive VQM. The reason for this is again the occurrence of TCP retransmission timeouts. Considering Figure 7.12 for the semi-dynamic OFDM-FDMA it shows up that even low amounts of timeouts make the capacity, freed by the VQM, unavailable to TCP related traffic streams. However, as known from Chapter 5, even for the WWW streams the channel adaption leads to a general benefit compared to the static case.

The results for the DIV in Figure 7.16 clearly benefit from the improved subcarrier assignment. The values are up to 37% smaller than with the static FDMA scheme (Section 5.6). Thus, up to 8 video receiving WTs can be supported at a acceptable video quality. The maximum difference of the results of both semantic

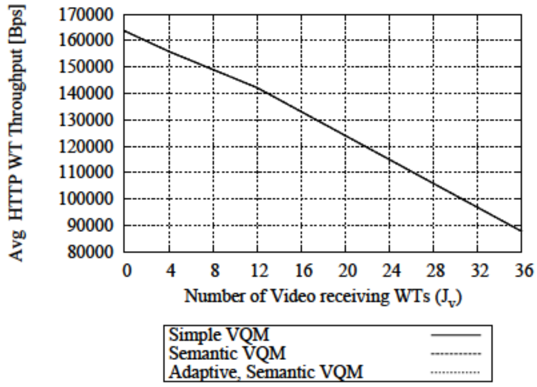


Figure 7.13: Average bit-rate of HTTP receiving WTs for different video queue management methods with semi-dynamic OFDM-FDMA

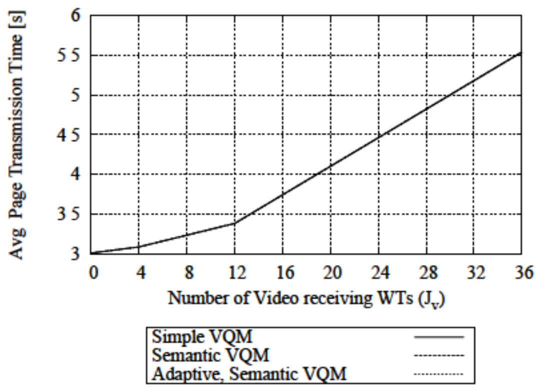


Figure 7.14: Average transmission time per Web-page for different video queue management methods with semi-dynamic OFDM-FDMA

VQM to the results of the simple VQM is 10% for $J_v = 36$. This is higher than with static OFDM-FDMA. Both semantic VQM approaches slightly outperform the simple VQM for 1% on the average if J_v is small.

These small improvements for the video quality are reflected by the results for the video packet loss rates in Figure 7.17. As with static FDMA the adaptive VQM leads to the smallest amount of lost packets related to I frames (3% for $J_v = 36$), while its P frame packet loss rate is up to 1% higher than the loss rate for the plain semantic approach. Both packet loss rates for the semantic strategies increase less for rising J_v than those for the simple approach. In addition to the smaller absolute values this explains the improvement of the video quality with the adaptive and the semantic VQM. The results for all shown B frame packet loss rates are smaller than those for the static VQM and closer to the results for dynamic OFDM-FDMA. The highest amount of packets is removed with the semantic methods, whereby the adaptive VQM results in up to 4% more removed packets than the plain semantic approach. It is interesting, that the slope of the adaptive VQM increases stronger for rising J_v than the slope for the semantic VQM. This shows a slight adaption

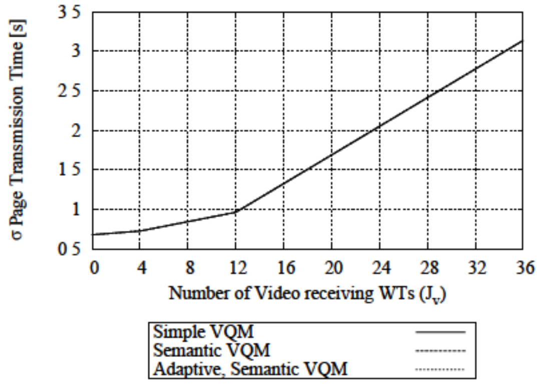


Figure 7.15: Standard deviation of the transmission time per Web-page for different video queue management methods with semi-dynamic OFDM-FDMA

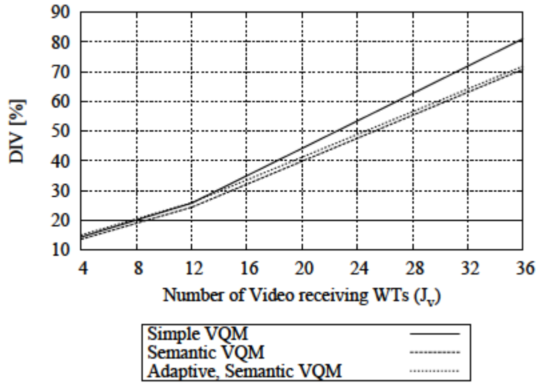
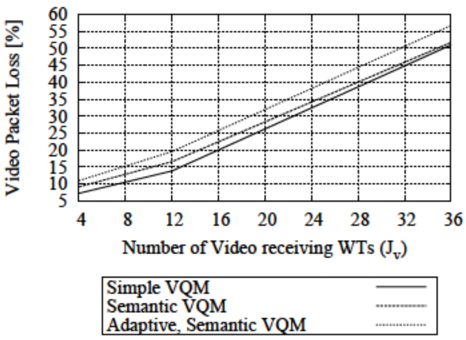


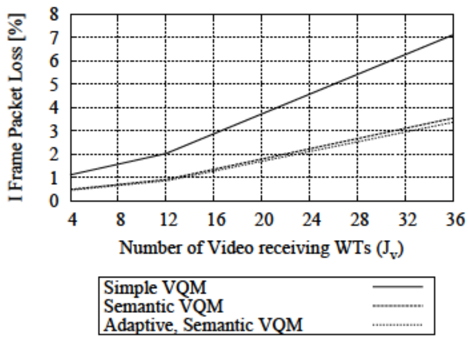
Figure 7.16: DIV for different video queue management methods with semi-dynamic OFDM-FDMA

to the traffic load, which rises with J_v. However, this does neither provides a significant improvement of the HTTP metrics nor an adequate increase of the video quality.

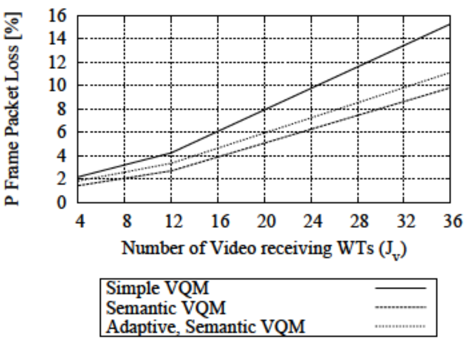
Although the benefit in this scenario basically results from the improved OFDM-FDMA scheme a very slight improvement is achieved for the DIV with both semantic VQM. Comparing the packet loss rates with those for the static FDMA scenario shows, that in this scenario adaption is performed by the load-adaptive removal of B frames. However, the best results are achieved with the combination of adaptive VQM and dynamic OFDM-FDMA, whose performance is further analyzed in Section 7.4.



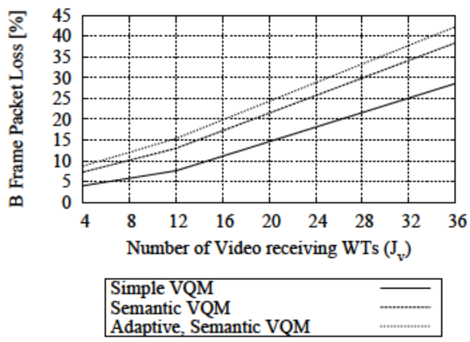
(a) Percentage of total lost video packets



(b) Percentage of lost video packets related to I frames



(c) Percentage of lost video packets related to P frames



(d) Percentage of lost video packets related to B frames

Figure 7.17: Video packet loss rates for different video queue management methods with semi-dynamic OFDM-FDMA

7.3.4 Interpretation

In this section the performance of the simple, the semantic and the adaptive VQM was examined for the three OFDM-FDMA schemes. With the results for the HTTP metrics presented for the combination of adaptive VQM and dynamic OFDM-FDMA in Section 7.3.1 and Figure 7.18 it is shown that the adaptive VQM provides a clear benefit in cases where the Web-page receiving WTs and a high number of video receiving WTs are competing for the limited resources of the transmission system. The high queuing delay in the high load scenarios leads to falling throughput rates and rising transmission times for all the considered VQM. However, with the adaptive VQM the downgrade is much flatter than with the two non-load-adaptive VQM which leads to gains for up to 83% for the HTTP transmission metrics. This is not the case if static or semi-dynamic OFDM-FDMA is employed. In Section 7.3.2 and Section 7.3.3 no gains are achieved for the transmission of HTTP regardless which VQM is employed. As discussed in Section

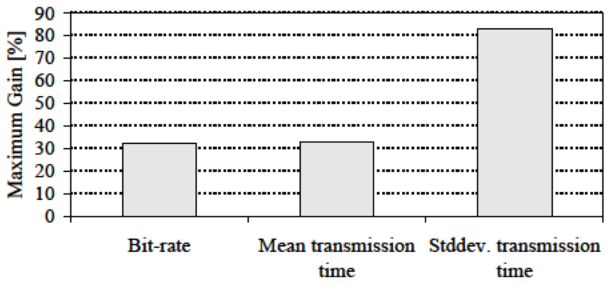


Figure 7.18: Maximum gains achieved for the WWW traffic metrics with the adaptive VQM versus the simple VQM for the dynamic OFDM-FDMA scheme and a constant J_h

7.3.2 this is the result of the low transmission capabilities at the link layer. Due to higher transmission delays the amount of video packets which hit their deadline rises significantly (7.11(a)). These packets are removed from the access point queue and can not be considered by both semantic VQM. Therefore, in comparison to the simple VQM the capacity gain, achieved by the adaptive filtering or the semantic removal of video packets, is lower. As shown for the DIV metric in Figure 7.16 and Figure 7.9 the slight capacity gain is utilized for the transmission of video streams, which leads to a small DIV gain of 1% for the semi-dynamic approach. Here UDP, employed for the video transmission, outperforms TCP. Due to its flow control schemes: congestion avoidance and slow-start, the TCP on the Web-server is not capable to make use of the freed capacities within the small time scale. Thus UDP, where no flow control is implemented, utilizes these resources prior to TCP.

For the DIV rate basically no benefit was found. However, it is important that the adaptive and the semantic VQM do not lower the video quality significantly. Considering the DIV obtained for $J_v < 16$ a quality decrease of up to 2% shows up for both semantic approaches in combination with dynamic OFDM-FDMA. Although for low J_v the loss rates for I and P frame related packets are lower, the B frame packet loss rate is higher than with the simple VQM (Figure 7.5). This implies that for $J_v < 16$ too many B frame packets are removed which results in lower quality than with the simple VQM. Therefore, the semantic weight for the B frames, defined in Equation 4.8, is too low for this scenario. However, due to the low DIV in this range the slight quality decrease is not perceptible for the viewer.

The video packet loss rates shown for the semi-dynamic OFDM-FDMA scheme in Figure 7.17 illustrate how the adaptive VQM benefits from dynamic subcarrier assignment. While due to the high B frame packet loss rates with static OFDM-FDMA the adaption basically fails, it starts with the semi-dynamic OFDM-FDMA.

7.4 Analysis for varying numbers of HTTP traffic streams

In this section the combination of adaptive VQM with dynamic OFDM-FDMA is further examined. In order to analyze the dependency of the VQM from the traffic-load scenario, three different J_r setups with varying J_h and J_v are investigated. The results are shown versus J_v, while the maximum number of J_v depends on J_r (Equation 7.2). Since various setups are investigated, the simulation results are presented in a shorter form. In place of absolute values the gain of the adaptive VQM in comparison to the other VQM schemes is shown. Although the shown gains may increase for rising J_v it has to be considered that the results for all metrics typically become worse if more video streams are “loaded” to the system. In order to investigate only the gains for acceptable video quality no results are presented for the DIV if the threshold of 20% is exceeded.

In Section 7.4.1 the adaptive approach is compared to the simple VQM, while in Section 7.4.2 adaptive and semantic VQM are compared. Finally, all results are considered in the discussion in Section 7.4.3.

7.4.1 Comparison of the simple VQM to the adaptive VQM

In Figure 7.19 the gain achieved by the adaptive VQM versus the simple approach is plotted for the TCP bit-rate of the HTTP receiving WTs. Considering the increase of the results for rising J_v shows clearly, how the gain of the adaptive VQM depends on the scenario, which is predetermined by J_r. The higher the amount of video streams, the higher is the achieved gain. For example, the comparison of the results at $J_{max} = 48$ for $J_r = 1/4$ and $J_r = 1$ leads to the conclusion that 50% more video streams in the cell result in the additional gain of 21% for the adaptive VQM on the average. For higher amounts of WTs the gain decreases for those scenarios with higher proportions of J_h, starting at $J_v = 8$ for $J_r = 1/4$ and at $J_v = 20$ for $J_r = 1$. This slope is not shown for the highest J_r considered, where 75% of the WTs are receiving video streams. Here the gain at J_{max} is with 26% only slightly higher than the gain of 25% for $J_r = 1$.

The high gain for the adaptive VQM versus the simple approach for the TCP throughput is clearly reflected by the results for the Web-page transmission time metrics, shown in Figure 7.20 and Figure 7.21. For the mean and the standard deviation of the Web-page transmission time the gain rises for higher J_r and slopes for the two lower J_r if J_{max} is nearly reached. Although the highest gain for the mean of 27% is achieved for $J_r = 4$ the maximum benefit of 73% for the standard deviation of the Web-page transmission time is reached for $J_r = 1$.

The comparison of the DIV, obtained for the adaptive and the simple VQM, shows in Figure 7.22 how the combination of semantic weighting and the load-adaptive filtering influences the video quality. As with the investigation for a fixed J_h in Section 7.3.1, here a quality decrease for low J_v shows up. This decrease is higher for lower J_r and nearly achieves a maximum of 3% for $J_r = 1/4$. However,

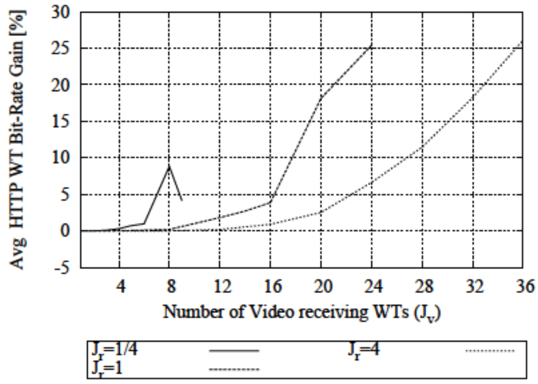


Figure 7.19: HTTP bit-rate gain achieved with adaptive VQM compared to the simple VQM for different J_r

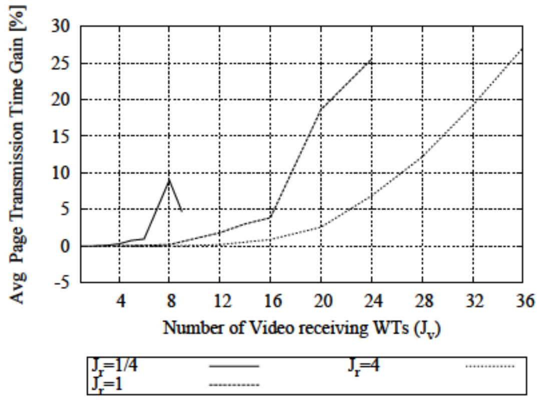


Figure 7.20: Web-page transmission time gain achieved with adaptive VQM compared to the simple VQM for different J_r

the absolute values for the DIV, shown exemplarily for $J_r = 1/4$ in Figure 7.23, indicate only a small quality decrease with both semantic VQM, which does not affect the DIV significantly. Since in this part of the performance study no DIV values larger than 20% are considered, basically no gains can be found. However, a very slight gain shows up for higher J_v with $J_r = 4$.

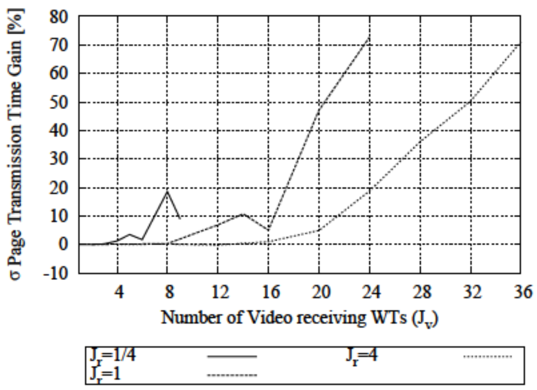


Figure 7.21: Gain for the standard deviation of the Web-page transmission time achieved with adaptive VQM compared to the simple VQM for different J_r

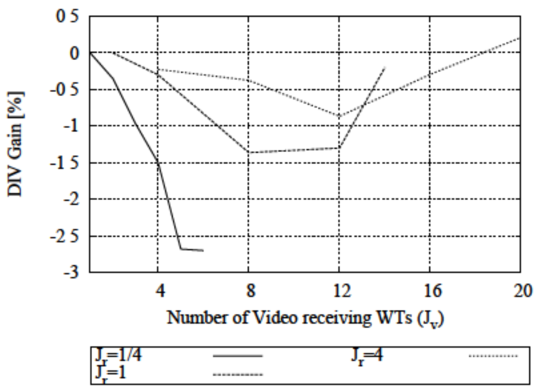


Figure 7.22: DIV gain achieved with adaptive VQM compared to the simple VQM for different J_r

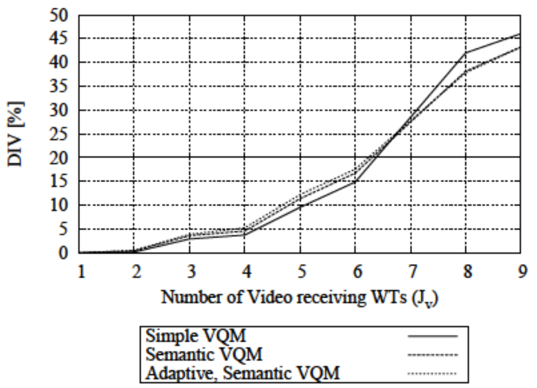


Figure 7.23: DIV for different video queue management methods and dynamic OFDM-FDMA for $J_r = 1/4$

7.4.2 Comparison of the semantic VQM to the adaptive VQM

In this section the gain achieved with the adaptive VQM versus the gain with the semantic VQM is shown for the HTTP metrics and the DIV. One result from Section 7.3.1 is, that the simple VQM is outperformed by the further developed semantic VQM. Therefore, the gains presented in this section are generally lower than the results in comparison to the simple VQM. However, still a high benefit is shown for the HTTP metrics. For the HTTP bit-rate (Figure 7.24) gains of up to 20% are achieved. The difference of 14% between the maximum achieved gains for $J_r = 1/4$ and $J_r = 4$ shows, that the achieved benefit depends on the proportion of video data in the queues. The gain achieved for $J_r = 1$ is with 1% slightly higher than for $J_r = 4$. In contrast to Figure 7.19 no decrease for rising J_v is visible for $J_r = 1$ if the adaptive VQM is compared to the plain semantic approach.

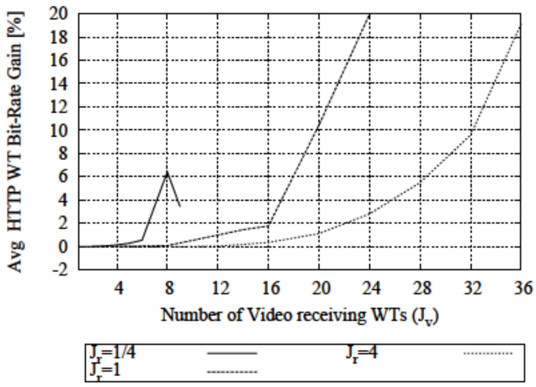


Figure 7.24: HTTP bit-rate gain achieved with adaptive VQM compared to the semantic VQM for different J_r

As for the TCP throughput the semantic VQM is clearly outperformed by the adaptive approach for the Web-page transmission time. For the mean the maximum gain is 22% for $J_r = 1$ and 54% for the standard deviation for $J_r = 4$. With $J_r = 1/4$ for both metrics a decrease shows up for higher J_v, while a slight decrease for $J_r = 1$ is visible only for the standard deviation if $J_v > 20$.

Although the rising gain for higher J_v is a clear benefit for the HTTP metrics this is not true for the DIV since here the absolute results are too high for an acceptable video quality. The comparison of the DIV gains versus the semantic and the adaptive VQM in Figure 7.27 is interesting, since here the effect of the adaptive filtering is shown separately. For low J_v this leads to a DIV decrease of up to 1.5% for the adaptive strategy ($J_r = J_v = 4$). Due to the small absolute results for the DIV, achieved for smaller J_v, the decrease in this range has no perceptible effect to the video. Also small gains can be found indicating that the semantic VQM is slightly outperformed by the adaptive approach. This is especially the case for larger portions of video streams in the cell ($J_r = 4$) and higher J_v, where a gain up to 1.6% can be found.

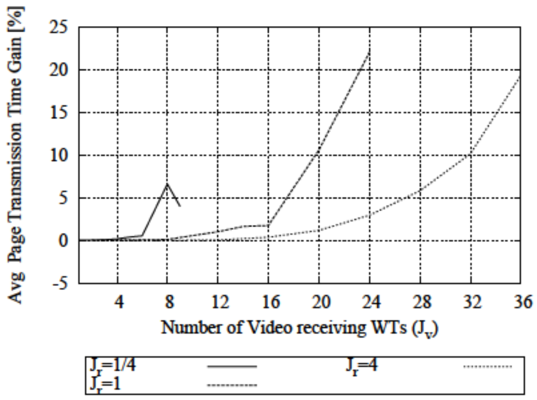


Figure 7.25: Web-page transmission time gain achieved with adaptive VQM compared to the semantic VQM for different J_r

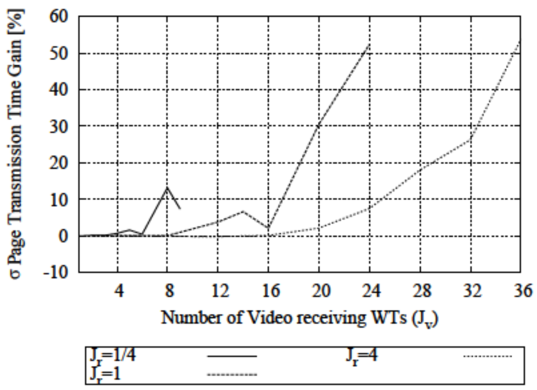


Figure 7.26: Gain for the standard deviation of the Web-page transmission time achieved with adaptive VQM compared to the semantic VQM for different J_r

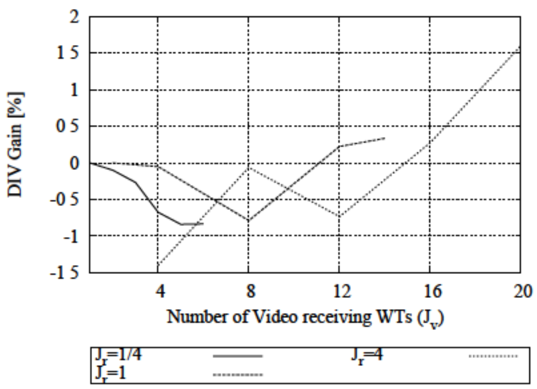


Figure 7.27: DIV gain achieved with adaptive VQM compared to the semantic VQM for different J_r

7.4.3 Interpretation

In this section the performance of the adaptive VQM in combination with dynamic OFDM-FDMA is investigated for varying amounts of video and HTTP receiving WTs in the cell. For the analyzed scenarios it is shown that the combination of adaptive VQM and dynamic OFDM-FDMA clearly outperforms the simple and the semantic VQM. The highest gains are typically achieved for high amounts of WTs if more than 50% of the heterogeneous traffic is streaming video. For these high-load scenarios the adaptive filtering provides an additional acceleration of the HTTP transmission. The maximum gains achieved with the adaptive VQM are summarized in Figure 7.28. The highest benefit is obtained for the standard deviation of the Web-page transmission time, where the simple VQM is outperformed by up to 73% and the plain semantic strategy by up to 54%.

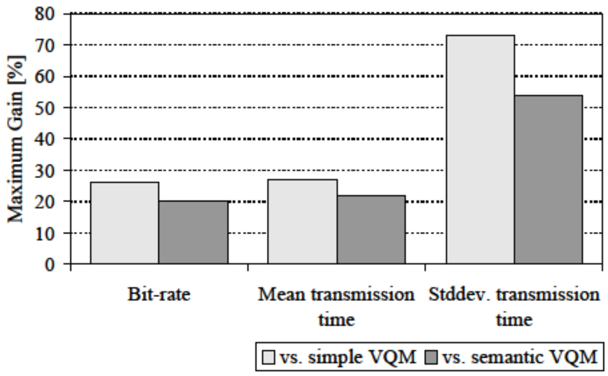


Figure 7.28: Maximum gains achieved for the WWW traffic metrics with the adaptive VQM in comparison to the simple and to the semantic VQM for scenarios with varying J_h and J_v

In general the benefit is higher, the higher the proportion of video data compared to the Web-traffic is. While this is more critical for $J_r < 1$ in the considered scenarios the VQM is not too sensitive. This is shown by the results obtained for the worst considered case with $J_r = 1/4$, i.e. only 25% of WTs in the cell are receiving video streams. Here the benefit which is achieved for the bit-rate by the adaptive VQM versus the simple scheme decreases for up to 21% in comparison to the scenarios with higher amounts of video traffic. However, even for low proportions of video streams in the queues benefits of up to 19% can be achieved (Figure 7.21). Although in general the gain rises for the HTTP metrics with rising numbers of WTs, in the scenarios for $J_r \leq 1$ the benefit starts to decrease from a specific J_v. Since as discussed in Section 6.2 rising traffic load stands for the disproportion of the demanded and the available resources of the transmission system, this decrease can be explained with basically the same argumentation than with static OFDM-FDMA in Section 7.3.4: Due to the high traffic load, caused by many WTs in the cell the rising queuing delays cause more video packets to hit their deadline. These packets are removed in the basic algorithm of the adaptive VQM and can not be considered in the further filtering process anymore. However, for higher J_r the increase of traffic load does not necessarily lead to drastic gain decreases. As shown for the standard deviation of the Web-page transmission time for $J_r = 1$ also slightly higher gains are possible for scenarios with lower percentages of video WTs.

The influence of the adaptive filtering to the video quality is represented by the DIV rate. In Section 7.4.1 the DIV rate resulting from the adaptive VQM is compared to the rate achieved with the simple VQM. As for fixed J_h in Section 7.3.1 a slight decrease shows up for small numbers of video streams in the cell. This is exemplarily shown for $J_r = 1/4$ in Figure 7.23, whereby the maximal decrease of 3% results from the employment of both semantic VQM approaches. Since it occurs in a range, where the absolute DIV rate is low, it causes no perceptible influence to the video quality. As for the HTTP transmission metrics the DIV gain starts to slope from a specific J_v. The comparison of the DIV achieved with adaptive VQM versus the semantic VQM has to be considered in particular. Since the semantic weighting is done by both VQM here only the influence of the adaptive filtering is shown. As presented in Figure 7.27 for low numbers of video streams the filter leads to a DIV decrease of up to 1%. As discussed above this does not cause drastic distortions to the video stream and is clearly justified by the benefit the adaptive VQM achieves for the HTTP transmission metrics of up to 54%.

Chapter 8

Conclusions and future work

In this thesis the performance of the dynamic OFDM-FDMA scheduling schemes for a heterogeneous traffic scenario was studied. Therefore, the simultaneous transmission of MPEG-4 coded VBR video streams and Web-pages was considered. In Chapter 5 three different combinations of dynamic OFDM-FDMA optimization schemes were investigated. Static OFDM-FDMA is the trivial case, where no adaption is performed. Semi-dynamic OFDM-FDMA stands for the dynamic assignment of subcarriers according to the measured channel states. The third case, dynamic OFDM-FDMA, considers the state-dependent assignment in combination with a dynamic multiple access scheme. This dynamic FDMA scheme enables the adaption to the varying application traffic load. It was shown, that the dynamic OFDM-FDMA approach clearly outperforms the static and the semi-dynamic case for the transmission of Web-pages and MPEG-4 video. With dynamic OFDM-FDMA the average transmission time of a Web-page is accelerated for up to 100% while up to 215% gain is obtained for its standard deviation. Furthermore, the simulation results show, that with dynamic OFDM-FDMA the quality of the transmitted video is increased for up to 37% in comparison to the static approach. A further comparison of the results for the three OFDM-FDMA schemes leads to the conclusion that the main part of the benefit is achieved with the dynamic subcarrier assignment. The gain starts to decrease for the dynamic multiple access scheme if typically the half of the maximum number of video receiving WTs per cell is reached.

For scenarios with high load of heterogeneous traffic the adaptive VQM was developed. It was proposed in Chapter 6 of this thesis. Considering application layer information this scheduling policy performs the load-adaptive filtering of video streams on top of the employed OFDM-FDMA schemes. Compared to the simple FIFO approach the adaptive VQM leads to performance gains of up to 33% for the mean and up to 83% for the standard deviation of the Web-page transmission time in addition to the gains achieved with dynamic OFDM-FDMA. This was found for scenarios with a high load of video traffic per cell.

Even though these scenarios may occur with best-effort networks as the Internet they are not characteristic. Most of the users will give up watching the video stream due to the bad quality. However, even with lower proportions of video on the over-all traffic load still acceptable results can be achieved (Section 7.4).

No performance increase for the transmission of Web-pages can be achieved if the adaptive VQM is applied on top of static or semi-dynamic OFDM-FDMA. This is also found in case of semantic VQM, proposed in [35]. A further result of the performance study in Chapter 7 is, that the achieved benefit for the HTTP transmission continuously rises for increasing amounts of video streams in the cell.

High traffic load leads in general to a drastic distortion of the transmitted video streams. Although this can be not prevented by the adaptive VQM, it is shown that its application does not lead to perceptible quality decreases.

The investigations show further, that the semantic VQM is clearly outperformed by the adaptive approach. For HTTP the comparison of the gains achieved with both schemes leads to gains for the adaptive VQM of up to 20% for TCP throughput and up to 54% for the standard deviation of the Web-page transmission time. Thus, the adaptive removal of video packets is more advantageous than the reactive scheduling strategy of the semantic VQM.

Beside the performance, the clear modular structure of the adaptive VQM has to be considered: The strict separation of the load estimation and the adaptive filter enables the adoption for other coding technologies than MPEG-4 and simplifies the development of new filter algorithms.

Due to the achieved performance increase and its generic design the adaptive VQM seems to be worth to be further investigated. For this purpose the following aspects show up in this thesis:

- Algorithm complexity:
  This important factor and the resulting cost have to be considered for the deployment of the adaptive VQM in transmission systems.

- Performance study for lowered algorithm complexity:
  The design of the adaptive VQM already considers the trade-off between calculation complexity and performance gain. With larger execution intervals the calculation overhead can be lowered by taking drawbacks for the adaption into account. It has to be examined how the adaptive VQM performs in this case.

- Further investigation of the contention of TCP and UDP for the capacities freed by the adaptive VQM:
  Although this topic was raised in the performance study of the adaptive VQM in combination with static and semi-dynamic OFDM-FDMA a more detailed study seems to be reasonable. Its results may lead to an extension of the VQM which directly avoids TCP retransmission timeouts.

- Improvement of the adaptive filtering algorithm:
  Aiming at additional performance gains – especially for lower percentages of video traffic in the cell – or the decrease of the algorithm complexity, for the development of new filtering algorithms two directions show up: At first the actual packet-drop filter can be enhanced in order to consider further

semantic features of the video streams. For example, considering the macro-block layer of MPEG a chess board like removal of macro-blocks could be implemented. Secondly, new filter approaches, e.g. re-quantization- and slicing filter could lead to better results.

- Optimization of the considered parameters:
  This includes the semantic weights, which are only heuristics taken from [35] as well as the parameter α chosen for the load estimation. The ideal solution would be a self-optimizing algorithm, which adapts those parameters to the actual scenario.

Common for all investigated optimization schemes is, that additional complexity is added to the system. Thus, it has to be analyzed whether the achieved performance gain legitimates the complexity increase, considering technical and finally economical metrics.

Appendix A

Scenario independent parameters

In the tables all the applied parameters are shown, which were left constant during the whole simulation. The list starts with the parameters for the radio channel and ends with those for the application layer.

Channel Modeling Parameters	Value
Maximum Speed of the WTs (v_{max})	1 m/s
Delay spread standard deviation ($\sigma\Delta\tau$)	0.15 μs
Path loss reference loss (10 log K)	46.7 dB
Path loss exponent (α)	2.4
Shadowing standard deviation (σ)	5.8

Table A.1: Values of radio channel modeling parameters for an indoor environment as defined in [1]

Physical Layer Parameters	Value
Total bandwidth (B)	16.25 MHz in the 5.2 GHz band
Number of subcarriers (S)	48
Maximum transmit power/subcarrier (P_{tx})	0.2 mW (-7 dBm)
Subcarrier spacing	312.5 kHz
Thermal noise power level (n_0)	-117 dBm
Symbol time (T_s)	4 μs
SEP threshold (P_s)	0.01

Table A.2: Values of the physical layer parameters according to the IEEE 802.11a WLAN standard [28]

MAC Frame Element	Duration
MAC frame (T_f)	2ms (500 Symbols)
Downlink- and Uplink phase	1ms (250 Symbols)
Downlink phase payload field [23]	0.968ms (242 Symbols)

Table A.3: Timing structure of a single MAC frame according to HIPERLAN/2 [11] and [23]

Session Level Parameter	Distribution Function	Distribution F. Parameters
Session-inter-time	Exponential	μ (variable)
Pages per session	Log normal	$\mu = 25.807$ pages/session
		$\sigma = 78.752$ pages/session
Page Level Parameters	Distribution Function	Distribution F. Parameters
Page-inter-time	Gamma	$\mu = 35.286$ s
		$\sigma = 147.39$ s
Page Size	Pareto	$\alpha = 1.7584$
		$\beta = 30458$ Bytes

Table A.4: Distribution functions and values of the WWW traffic model as defined in [54] for the corporate environment

MPEG-4 Video Parameters	Value
GOP	12-GOP: IBBPBBPBBPBB
Mean bit-rate	951 KBps
Frame size	352×288 Pixel (CIF)
Frame-rate	25 frames/s
Video length	4500 frames, i.e. 180 s

Table A.5: Applied MPEG-4 video parameters

Appendix B

Acronyms

16-QAM 16 Quadrature Amplitude Modulation

64-QAM 64 Quadrature Amplitude Modulation

256-QAM 256 Quadrature Amplitude Modulation

ADA Advanced Dynamic Algorithm

BPSK Binary Phase Shift Keying

CDMA Code Division Multiple Access

CPU Central Processing Unit

CIF Common Intermediate Format

CNR Channel Gain-to-Noise Ratio

DFT Discrete Fourier Transform

DIV Distortion In Interval

FBFQ Frame-Based Fair Queueing

FDMA Frequency Division Multiple Access

FFT Fast Fourier Transform

FIFO First-In-First-Out

GOP Group of Pictures

HTTP Hypertext Transfer Protocol

HTML Hypertext Mark-up Language

ICI Inter-carrier Interference

IEEE Institute of Electrical and Electronics Engineers, Inc.

IETF Internet Engineering Task Force

IP Internet Protocol

ISI Inter-Symbol Interference

ITU International Telecommunication Union

IWFQ Idealized Weighted Fair Queueing

LAN Local Area Networks

MAC Medium Access Control

MCM Multi Carrier Modulation

MPEG Moving Picture Experts Group

MOS Mean Opinion Score

MSS Maximum Segment Size

MTU Maximum Transmission Unit

OFDM Orthogonal Frequency Division Multiplexing

PSNR Peak Signal-to-Noise Ratio

QoS Quality of Service

QPSK Quadrature Phase Shift Keying

RTT Round-Trip-Time

SEP Symbol Error Probability

SCM Single Carrier Modulation

SNR Signal-to-Noise Ratio

TCP Transmission Control Protocol

TDMA Time Division Multiple Access

UDP User Datagram Protocol

UWB Ultra Wideband

VBR Variable Bit-Rate

VQM Video Queue Management

WAN Wide Area Network

WLAN Wireless Local Area Networks

WTs Wireless Terminals

WWW World-Wide-Web

Bibliography

[1] A. Aguiar and J. Gross. Wireless channel models. Technical Report TKN-03-007, Telecommunication Networks Group, April 2003.

[2] A. Bakne and B. R. Badrinath. I-TCP: indirect TCP for mobile hosts. In *Proc. Fifteenth Int. Conference on Distributed Computer Systems, IEEE*, pages 136–143, 1995.

[3] H. Balakrishnan, S. Seshan, and R. H. Katz. Improving reliable transport and handoff performance in cellular wireless networks. In *Proc. ACM Mobile Computing and Networking Conf.*, pages 2–11, 1995.

[4] P. Barford and M. E. Crovella. Generating represantive web workloads for network and server performance evaluation. In *Proc. of the ACM SIGMETRICS '98*, 1998.

[5] T. Berners-Lee, R. Fielding, and H. Frystyk. Hypertext transfer protocol – HTTP/1.0. RFC 1945, IETF, May 1996.

[6] P. Bhagwat, P. Bhattacharya, A. Krishna, and S. Tripathi. Enhancing throughput over wireless LANs using channel state dependent packet scheduling. In *IEEE INFOCOM*, 1994.

[7] J. Cavers. *Mobile Channel Characteristics*. Kluwer Academic, 2000.

[8] R. W. Chang. Synthesis of band-limited orthogonal signals for multichannel data transmission. *Bell Systems Technical Journal*, 45, December 1966.

[9] ETSI. Broadband radio access networks (BRAN); HIPERLAN type 2; system overview. Technical Report TR 101 683, February 2000.

[10] ETSI. Digital video broadcasting (DVB); OFDM modulation for microwave digital terrestrial television. EN 301 701, August 2000.

[11] ETSI. *Broadband Radio Access Networks (BRAN); HIPERLAN Type 2; Data Link Control (DLC) Layer; Part 1: Basic Data Transport Functions*, TS 101-761-1 edition, December 2001.

[12] ETSI. *Broadband Radio Access Networks (BRAN); HIPERLAN Type 2; Physical (PHY) Layer*, TS 101-475 edition, December 2001.

[13] ETSI. Radio broadcasting systems; digital audio broadcasting (DAB) to mobile, portable and fixed receivers. EN 300 401, May 2001.

[14] R. Fielding, J. Gettys, J. Mogul, and H. Frystyk et al. Hypertext transfer protocol – HTTP/1.1. RFC 2068, IETF, January 1997.

[15] F. Fitzek, P. Seeling, M. Reisslein, and M. Rossi et al. Investigation of the GoP structure for H.26L video streams. Technical Report acticom-02-004, acticom mobile networks, December 2002.

[16] S. Floyd and T. Henderson. The NewReno modification to TCP's fast recovery algorithm. RFC 2582, IETF, April 1999.

[17] C. Fraleigh, S. Moon, B. Lyles, and C. Cotton et al. Packet-level traffic measurements from the sprint IP backbone. *IEEE Network*, November 2003.

[18] A. Francini and F. M. Chiussi. A weighted fair queueing scheduler with decoupled bandwidth and delay guarantees for the support of voice traffic. In *Proc. of Internet Performance Symposium (IPS)*, November 2001.

[19] J. Gross, H. Karl, F. Fitzek, and A. Wolisz. Comparison of heuristic and optimal subcarrier assignment algorithms. In *Proc. of the 2003 International Conference on Wireless Networks (ICWN'03)*, June 2003.

[20] J. Gross, J. Klaue, H. Karl, and A. Wolisz. Subcarrier allocation for variable bit rate video streams in wireless OFDM systems. In *Proc. of Vehicular Technology Conference (VTC Fall)*, Florida, USA, 2003.

[21] J. Gross, J. Klaue, H. Karl, and A. Wolisz. Cross-layer optimization of OFDM transmission systems for MPEG-4 video streaming. *Computer Communications*, 2004.

[22] J. Gross, S. Valentin, H. Karl, and A. Wolisz. A study on the impact of inband signaling and realistic channel knowledge for an example dynamic OFDM-FDMA system. *to appear in Computer Communications*, 2004.

[23] J. Gross, S. Valentin, and R. Vu. Dynamic OFDM-FDMA systems under realistic assumptions: On the influence of channel knowledge and inband signaling. Technical Report TKN-04-[to appear], Telecommunication Networks Group, Technische Universität Berlin, September 2004.

[24] R. Grünheid. *Vielfachzugriffsverfahren für die Multiträger-Übertragungstechnik.* Dissertation, TU Braunschweig, 2000.

[25] M. Hemy, U. Hengartner, P. Steenkiste, and T. Gross. MPEG system streams in best-effort networks. In *Proc. Packet Video '99*, New York, USA, April 1999.

[26] M. Hemy, P. Steenkiste, and T. Gross. Evaluation of adaptive filtering of MPEG system streams in IP networks. In *Proc. IDME*, 2000.

[27] International Engineering Consortium (IEC). *UMTS protocols and protocol testing*, 2003.

[28] IEEE. *Supplement to standard for telecommunications and information exchange between systems – LAN/MAN Specific requirements – Part 11: Wireless LAN MAC and PHY specifications: High speed physical layer in the 5-GHz band*, P802.11a/D7.0 edition, July 1999.

[29] IEEE. *IEEE Standard for Information technology – Telecommunications and information exchange between systems – LAN/MAN Specific requirements – Part 11: Wireless LAN MAC and PHY Layer specifications: – Amendment 4: Further higher-speed physical layer extension in the 2.4 GHz band*, Std 802.11g, 1999 (reaff 2003) edition, 2003.

[30] University of Southern California Information Sciences Institute. Internet protocol DARPA Internet program protocol specification. RFC 791, DARPA, September 1981.

[31] University of Southern California Information Sciences Institute. Transmission control protocol DARPA Internet program protocol specification. RFC 793, DARPA, September 1981.

[32] ITU. *ITU-T Recommendation P.930 – Principles of a reference impairment system for video*, August 1996.

[33] V. Jacobson. Berkeley TCP evolution from 4.3-Tahoe to 4.3-Reno. In *Proc. of the Eighteenth Internet Engineering Task Force*, page 365, September 1990.

[34] V. Jacobson and M. J. Karels. Congestion avoidance and control. In *Computer Communication Review*, volume 18, pages 314–329, August 1988.

[35] J. Klaue, J. Gross, H. Karl, and A. Wolisz. Semantic-aware link layer scheduling of MPEG-4 video streams in wireless systems. In *Proc. of Applications and Services in Wireless Networks (ASWN)*, Bern, Switzerland, July 2003.

[36] P. Klenner. Das OFDM Multiträgerverfahren. Universität Bremen, June 2004.

[37] B. Krongold, K. Ramchandran, and D. Jones. Computationally efficient optimal power alloction algorithms for multicarrier communication systems. In *Proc. of IEEE Int. Conference on Communications (ICC)*, pages 1018–1022, 1998.

[38] H. W. Kuhn. The hungarian method for the assignment problem. *Naval Research Logistics Quarterly 2*, 1955.

[39] T. Le-Ngoc, N. Damji, and Y. Xu. Dynamic resource allocation for multimedia services over OFDM downlink in cellular systems. In *Proc. of Spring Vehicular Technology Conference (VTC Spring)*, 2004.

[40] D. A. Menascé and V. A. F. Almeida. *Capacity Planning For Web Performance*. Prentice Hall, Inc., Upper Saddle River, NJ, USA, first edition, 1998.

[41] X. Meng, H. Yang, and S. Lu. Application-oriented multimedia scheduling over lossy wireless networks. In *Proc. of IEEE ICCCN'02*, October 2002.

[42] M. Miyabayashi, N. Wakamiya, M. Murata, and H. Miyahara. MPEG-TFRCP: video transfer with TCP-friendly rate control protocol. In *Proc. of IEEE Int. Conference on Communications (ICC)*, 2001.

[43] B. Moraru, F. Copaciu, G. Lazar, and V. Dobrota. Practical analysis of TCP implementations: Tahoe, Reno, NewReno. In *RoEduNet International Conference*, pages 125–130, June 2003.

[44] K. N. Ngan, C. W. Yap, and K. T. Tan. *Video Coding for Wireless Communication Systems*. Signal Processing and Communications. Marcel Dekker, Inc., 2070 Madison Avenue, NY, USA, 2001.

[45] J. Nieh. *The Design Implementation, and Evaluation of SMART: A Scheduler for Multimedia Applications*. Dissertation, Department of Electrical Engineering, Stanford University, June 1999.

[46] NTIA Office of Spectrum Management. US frequency allocation chart. Retreived September 3, 2004 from `http://www.ntia.doc.gov/osmhome/allochrt.html`.

[47] B. O'Hara and A. Petrick. *IEEE 802.11 Handbook: A designers companion*. IEEE Press, 1999.

[48] J.-R. Ohm. *Bildsignalverarbeitung für Multimedia-Systeme*. TU Berlin, 1999.

[49] P802.15. Wireless personal networks; P802.15.3a alt PHY selection criteria. Technical Report 03/031e5, IEEE, December 2002.

[50] S. Pietrzyk and G. J. M. Janssen. Multiuser subcarrier allocation for QoS provision in the OFDMA systems. In *Proc. of IEEE Vehicular Technology Conference Fall 2002*, September 2002.

[51] J. Postel. User datagram protocol. RFC 768, ISI, August 1980.

[52] J. Proakis. *Digital Communications*. McGraw-Hill, 1995.

[53] T. Rappaport. *Wireless Communications Principles & Practice*. Prentice Hall, Inc., Upper Saddle River, NJ, USA, 1999.

[54] A. Reyes-Lecuona, E. González-Parada, E. Casilari, and J. C. Casasola et al. A page-oriented WWW traffic model for wireless system simulations. In *Proc. ITC 16*, June 1999.

[55] M. J. Riley and I. E. G. Richardson. *Digital Video Communications.* Artech House, 1997.

[56] R. Srikant S. Lu, V. Bharghavan. Fair scheduling in wireless packet networks. In *Proc. of the ACM SIGCOMM '97 conference on Applications, technologies, architectures, and protocols for computer communication*, September 1997.

[57] W. Stallings. *Data and Computer Communications.* Prentice-Hall, Inc., Upper Saddle River, NJ, USA, sixth edition, 2000.

[58] W. R. Stevens. *TCP/IP Illustrated Volume 1.* Professional Computing Series. Addison-Wesley Longman, Inc., 1994.

[59] A. S. Tanenbaum. *Computer networks.* Prentice-Hall, Inc., Upper Saddle River, NJ, USA, third edition, 1996.

[60] D. Thie and J. Flohrer. Modellierung des Datenverkehrs von WWW-Clienten. In *Proc. MMB'99*, September 1999.

[61] A. Varga. OMNeT++ discrete event simulation system. Software is available on the Internet: `http://www.omnetpp.org` (verified September 7, 2004), June 2003. (Release 2.3).

[62] P. Vary, U. Heute, and W. Hess. *Digitale Sprachsignalverarbeitung.* B. G. Teubner, Stuttgart, Germany, 1998.

[63] K. Wehrle, J. Reber, D. Holzhausen, and V. Boehm et al. The OMNeT++ Internet protocol suite. Software is available on the Internet: `http://www.omnetpp.org` (verified September 7, 2004), December 2001.

[64] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch. Multiuser OFDM with adaptive subcarrier, bit, and power allocation. *IEEE Journal on Selected Areas in Communications, 17(10)*, October 1999.

[65] N. Yeadon, F. García, D. Hutchison, and D. Shepherd. Continuous media filters for heterogeneous internetworking. In *Proc. of SPIE Multimedia Computing and Networking (MMCN'96)*, San Jose, USA, January 1996.

Erklärung der Urheberschaft

Die selbstständige und eigenhändige Anfertigung der vorliegenden Magisterarbeit versichere ich an Eides statt.

Ort, Datum Unterschrift