MSc. Thesis

Optimisation of packet scheduling in wireless networks

Johann Huibert Schouten

March 25, 2009

Exam committee

Prof. dr. R.J. Boucherie
Prof. dr. ir. F.J.A.M. van Houten
Dr. J.C.W. van Ommeren
Ir. R.J. Mantel
Prof. dr. J.L. van den Berg (TNO ICT)
Dr. R. Litjens MSc. (TNO ICT)

University of Twente

Faculty of Electrical Engineering, Mathematics and Computer Science
Group of Stochastic Operations Research

Faculty of Engineering Technology
Group of Design, Production and Management

Summary

Self-organisation is foreseen as a key feature of future wireless communications networks in order to decrease operational cost and to exploit resources to their fullest potential. Self-organisation of wireless networks comes in three forms: self-optimisation, self-configuration and self-healing. In this report we will focus on self-optimisation, i.e. the adjustment of operational algorithms and parameters in order to adapt to changes in the network, traffic and channel conditions.

There are multiple capacity allocation mechanisms in the operations of wireless networks, which can be self-optimised. The main difference between these capacity allocation mechanisms is that they each perform a different task and act at their own operational timescale. In this research we will focus on the packet scheduling algorithm that acts on a timescale of milliseconds, i.e. at the packet level. Before self-optimisation methods for packet scheduling can be designed, the potential for self-optimisation of packet scheduling must be investigated, which is the focus of this thesis.

The research objective of this study is to investigate the sensitivity of the optimal parameter settings for packet scheduling algorithms in wireless networks, loaded with voice and data traffic, with respect to variations in network, traffic and channel conditions. We investigate this sensitivity by means of a mathematical analysis and a simulation model. In the mathematical model we use time scale decomposition between the packet and flow level. Scheduling algorithms operate at the packet level at which the user population is relatively stable. The flow level dynamics, due to flow transfer initiations and completions, causes the number of ongoing flow transfers to vary over time, typically at a timescale of seconds. Our modelling approach captures the packet level and flow level behaviour and basically consists of two steps. In the first step we consider the packet level and derive mathematical expressions for the long-term rate that users can expect in a given system state indicated by the number of active flows. In the second step we incorporate these rates in a Markov chain model to analyse flow level performance for the users of the mobile wireless network.

Since the analytical model that we introduce is limited in the set of sensitivity analysis that can be carried out with it, we introduce a simulation model with which we are able to incorporate more details and analyse more network, traffic and channel conditions. Furthermore, the simulation model replicates the behaviour of a network, with interaction between base stations, whereas the mathematical analysis is limited to a single cell scenario. The simulation model is a more realistic model of network, traffic and propagation characteristics. We investigate the flow level performance of users in this network for different parameters of the packet scheduling algorithm we describe in this thesis. This way we determine whether the optimal parameter setting for the packet scheduling algorithm depends on changes in the network, traffic and channel conditions.

Both the numerical results from the mathematical model and the results from the simulation study show that the optimal parameter settings of the considered packet scheduler are largely insensitive to changes in the considered network, traffic and channel conditions. We therefore conclude that there is no potential for self-optimisation of packet scheduling algorithms based on the investigated network, traffic and channel conditions.

Preface

The eleven months of research for this thesis I did at TNO ICT in Delft have been a great experience for me. During this project I had the opportunity to broaden my knowledge about operations research and I have learned how research is conducted in a professional environment. Within the context of my project I gained insight in wireless networking, in particular the issue of packet scheduling playing an important role in the operations of these networks. Performing literature studies, building a simulation model in Delphi programming code and performing more fundamental analysis have been a great challenge. I can look back at the project with satisfaction, knowing that I achieved my goals and produced some nice results.

I would like to thank Prof. dr. J.L. van den Berg and Dr. R. Litjens MSc. from TNO ICT for their great help and their critical questions to keep me focussed. Our weekly discussions were crucial for the progress of the project. Furthermore I would like to thank Prof. dr. R.J. Boucherie and Ir. R.J. Mantel from the University of Twente, for their personal approach and their coaching in the last years of my master, which kept me motivated. Their feedback in the project, as well as their support in combing a masters programme in both Mechanical Engineering as well as Applied Mathematics, have been great. I would like to thank my family and friends, especially my parents, for their help and good times during my studies.

Amsterdam, March 2009

Contents

Chapter 1

Introduction

1.1 Background

Wireless telecommunication networks are an essential part of modern life. Although large parts of Africa are not connected to fresh water supply, the people that live in the smallest villages do have mobile phone coverage. Mobile devices are becoming more sophisticated every day. Mobile wireless internet and applications for watching a video or interacting with players during an online game are becoming part of every day life. In order to allow people to use all these applications with increasing bandwidth needs, developments in mobile wireless networks are essential. While GSM networks were mainly focussed on wireless telephone conversations, mobile wireless networks of the future will be used by various applications, each with their own service requirements.

In order to provide users of the mobile network with the service they request, the capacity of the network must be allocated. Several methods to allocate capacity within mobile wireless networks exist. Each of these capacity allocation methods has its own parameter settings. But while networks and their usage are getting more complex, the optimisation of these parameters is still conducted by manual intervention of network operators. However, to allow increased service to users of mobile networks at competitive costs, wireless communications networks of the future will have automated optimisation of the parameters in their network capacity allocation mechanisms. This adaptation of parameters to the current state of the wireless network, self-optimisation, is one of the trends for self-organising, or more automated, wireless networks. Self-organising networks have automated installation of new base station in the network (self-configuration), triggers to detect failure on which they act without intervention of humans (self-healing) and continuous optimisation of their operational settings (self-optimisation). To cope with increasing demand at competitive cost, standardisation body 3rd Generation Partnership Project (3GPP) [1] and operators lobby Next Generation Mobile Networks (NGMN) [2] pursue a significant degree of self-organisation in future wireless networks. In order to implement self-organisation in wireless communications networks, studies such as those carried out by Van den Berg et al. [10] visualise how future wireless communications networks will use self-organisation methods. Self-optimisation, self-configuration and self-healing are foreseen as promising opportunities to automate wireless network planning and operation.

In self-optimisation active base stations may continuously adjust their operational algorithms and parameters due to changes in network, traffic and channel conditions. In this thesis we will investigate the packet scheduler, one of the operational algorithms to allocate capacity. The packet scheduler handles the allocation of the transmit power of base stations to users in the network, acting at millisecond scale. The packet scheduling algorithm is foreseen as promising for use in self-optimisation of wireless networks.

1.2 Previous research

As indicated in the previous section the standardisation body 3rd Generation Partnership Project (3GPP) [1] and operators lobby Next Generation Mobile Networks (NGMN) [2] pursue a significant degree of self-organisation in future wireless networks. Ideas on how to implement self-organisation in these networks are currently developed and described by e.g. Van den Berg et al. [10]. Research on actual methods and algorithms on how to implement self-organisation in networks is emerging and described by e.g. Lei et al. [25], Delli Priscoli et al. [34] and Senouci et al. [37] who describe self-optimisation of call admission control.

Research on self-optimisation of packet scheduling is hardly available. Necker [31] compares various different scheduling schemes, investigating service differentiation of traffic classes. Packet schedulers in wireless networks come in different forms. In fading environments, channel-aware scheduling strategies, such as the Proportional Fair algorithm are attractive to exploit channel quality fluctuations while providing user fairness [6, 23]. This algorithm is mostly investigated in networks consisting of data users. Shakkottai and Stolyar [35] investigate scheduling algorithms for a mixture of real-time and non-real-time traffic. Furthermore the recent work by Elsayed and Khattab [19, 24] describes an interesting and pragmatic scheduling principle, the Channel-Aware Earliest Deadline Due, that incorporates both channel-awareness as due date aspects to allow for this mixture of real-time and non-real-time traffic.

Performance of channel-aware scheduling algorithms has mostly been investigated at the packet level for static user populations, including packet-scale dynamics [5, 35], but often assuming infinite backlogs [4, 16, 30]. This assumption of a static user population is reasonable because of the separation in time scales. Scheduling algorithms operate at the packet level on which the user population evolves relatively slowly. Most analytical studies focus on the performance of schedulers without taking into account the impact of the flow level dynamics. Bonald, Proutière and Borst have studied the performance of wireless data networks that do include this flow level. Interesting references here are [11, 12, 13, 14, 15]. In these papers they provide analytical performance evaluation of wireless data networks.

Furthermore there are simulation studies taking the actual flow level characteristics into account. Litjens and Van den Berg extensively studied the flow level performance in HSDPA networks [9, 28] with analytical and simulation models. In their papers they investigate the influence of environment and traffic aspects on the flow level performance.

1.3 Research objectives

To allow for increased service to users of mobile networks at competitive costs, wireless communications networks of the future will have to become more automated. In order to decrease operational cost and to exploit resources to the fullest potential, self-organistion in wireless networks is foreseen as inevitable. Before these future wireless networks will be operational, research on self-organisation methods is needed. As we will describe in detail in Chapter 2 self-organisation of wireless networks comes in three forms; self-optimisation, self-configuration and self-healing. In this report we will focus on the adjustment of operational algorithms and parameters due to changes in the network-, traffic- and channel conditions: self-optimisation.

As we will describe in Chapter 2 there are multiple levels in wireless communications networks at which capacity of the network is defined and allocated. On each of these levels we can implement self-optimisation methodologies. The main difference between these capacity allocation mechanisms is that they each act at their own operational timescale. In this research we will focus on the packet scheduling algorithm that acts on a timescale of milliseconds at the packet level of the traffic management.

Before we can actually develop algorithms that self-optimise the operational parameters in wireless networks we discussed, the potential for self-optimisation of the parameters from these capacity allocation methods has to be investigated. There are multiple time scales on which different capacity allocation mechanisms act. We have to determine which of these mechanisms are crucial in cost-effective and efficient allocation of capacity. Upon determination of the potential of self-optimisation for each of the capacity allocation methods, actual operational self-optimisation can be developed.

In this study we determine the potential for self-optimisation of the packet scheduling algorithm. This algorithm acts on a timescale of milliseconds at the packet level of the traffic management. Therefore we define our research objective as:

Investigate the sensitivity of the optimal parameter settings for packet scheduling algorithms in wireless networks, loaded with voice and data traffic, with respect to variations in network, traffic and channel conditions, in order to determine the potential for self-optimisation of packet scheduling algorithms.

In this objective we identify 3 main points:

- To investigate the performance of packet scheduling algorithms, we define so called Quality of Sevice (QoS) measures for each of the user types we consider (voice and data users).
- To determine the sensitivity of the optimal parameter settings on the QoS measures for each user type under variations in network, traffic and channel conditions. We change network, traffic and channel conditions and measure the difference in QoS for the users of the network.
- To determine the possible advantage for self-optimisation of packet scheduling algorithms, we quantify the differences in traffic capacity (the supportable traffic load given predetermined QoS targets) of the network under changing network, traffic and channel conditions. We are mainly interested whether the optimal packet scheduling algorithm is dependent on one of the conditions we change.

Upon determining the possible advantage for self-optimisation of packet scheduling algorithms, we can define the gain of using self-optimisation techniques to reduce operational cost and to exploit resources to the fullest potential. This (potential) gain will help in satisfying customer requirement of mobile wireless networks of the future.

1.4 Structure

We will approach the research objective by elaborating on wireless communications networks in more detail in Chapter 2. We will give an overview on packet scheduling within wireless networks in Chapter 3. We then define a mathematical model in Chapter 4 and a simulation model in Chapter 5 to investigate the sensitivity of packet scheduling algorithms for variations in network, traffic and channel conditions. We will describe the numerical results for the studies performed in Chapter 6 and conclude this thesis with some conclusions and recommendations in Chapter 7.

Chapter 2

Wireless communication networks

In this chapter we will introduce the fundamentals of wireless communication networks. We will describe the methods for capacity allocation in these wireless networks and discuss the general idea behind self-organisation of wireless networks.

2.1 Fundamentals of wireless communication networks

This section highlights some basic aspects which are fundamental to the development and operations of wireless communication networks. Spectrum is the fundamental resource for wireless communications. In order to allow more than one user on a wireless channel and use this spectral resource effectively multiple access to this resource is necessary. Finally we will introduce the concept of propagation of radio signals, essential in understanding wireless communication networks.

2.1.1 Spectrum

Spectral resources are rather intangible yet fundamental to wireless communication. Spectrum is an expensive resource. When the UMTS spectrum became available in The Netherlands in 2000 five companies paid a total of 2.7 billion Euro for the licences to operate on the frequencies until 2016. Nations are in charge of the wireless spectrum in their own country, though the need for cost-efficient use of the resource led to European as well as international agreements. Technologies in the spectrum, e.g. GSM and UMTS and the future technologies Wimax and LTE each use part of the available spectrum. In this spectrum the uplink frequency is used to sent information from a mobile device (the user) to a base station, using dedicated or shared channels. On the downlink frequency information is sent from a base station to the user.

2.1.2 Multiple access

In order to use the expensive resource effectively and efficiently, wireless networks use techniques to allow for multiple access of the spectrum to its users. Allowing users access to the resource will allow them to be served at their requested service. There are three main technologies to allow multiplexing of users to the resource for efficient use.

- Frequency multiplexing slices the spectrum in small frequency pairs, that are used to support dedicated channels between the transmitter and receiver.
- Time multiplexing partitions the available channel into time slots, where each channel has multiple time slots that each can serve one user.
- Code multiplexing uses channelisation (or spreading) codes to grant access to a shared channel. More codes allows the user to send information at a higher bit rate.

In this research we will use a combination of time multiplexing and code multiplexing, which we will describe in more detail in Chapter 3.

2.1.3 Propagation

Essential in wireless communication is the degradation of the wireless signal, the radio signal, as it propagates from the transmitting base station to the intended receiving user. Propagation of this signal can be described by three mutually independent multiplicative propagation phenomena that affect the transmitted signal at different scales of time and magnitude. We will describe these phenomena of path loss, shadowing and multipath propagation below.

Path loss

At the largest time scale, the degree of attenuation is predominantly determined by the transmission path length, the antenna heights and the carrier frequency. In generic analyses, the effects of attenuation are usually modelled by assuming an average attenuation which increases with distance according to a power law. Models developed by Okumura [32] and Hata [21] describe this phenomenon of path loss. For wireless communications the Okumura-Hata method is most commonly used, since this model has continuously been updated in a.o. the Cost 231 project.

Shadowing

Shadowing is a medium scale effect which occurs whenever there is an obstruction in the direct path from the transmitter to the receiver. Such obstructions are mountains or buildings. Shadowing is denoted as slow fading, since the landscape between the transmitter and receiver changes slowly, the shadowed areas are large. Shadowing effects are observed experimentally by Egli in 1957 [18], which states that the local mean power is distributed lognormally around the attenuation based area mean power. By 'lognormal' is meant that the local-mean power expressed in logarithmic values, such as dB, has a normal (i.e., Gaussian) distribution. Egli's experimental observation is generally used to model shadowing effects in wireless networks.

Multipath fading

In wireless telecommunication, multipath is the propagation phenomenon that results in radio signals reaching the receiving antenna by two or more paths. Causes of multipath include atmospheric ducting, ionospheric reflection and refraction, and reflection from water bodies and terrestrial objects such as mountains and buildings. The effects of multipath include constructive and destructive interference, and phase shifting of the signal. As such multipath fading effects vary over very short (typically half-wavelength) distances, the term fast fading is often used. The standard statistical model of this gives a distribution known as the Rayleigh distribution [26]. Rayleigh fading with a strong line of sight content (i.e. with less constructive and destructive interference), is said to be Ricean fading [20, 26, 33, 41]. In this study we use the Rayleigh and Ricean fading models as they are most commonly used and capture the essence of multipath fading.

2.2 Capacity allocation in wireless communication networks

In order to allocate the scarce resource of the spectrum to the intended users of the wireless communication network some form of capacity allocation is essential. In Figure 2.1 the different capacity allocation mechanisms, as described by Litjens [27], are depicted.

In this figure we see that capacity allocation acts at four levels, each acting at their own time scale. Although each of the mechanisms carries out a well-defined task, interdependencies exist. We will describe the mechanisms at network planning level and traffic management level. A more

detailed description on packet scheduling can be found in Chapter 3, the other capacity allocation mechanisms are described in more detail by Litjens [27].

Network planning

At the network and cell level, the network planning mechanisms set the capacity and thus establishes the operational framework for the traffic management mechanisms that operate at the call and burst or packet level. At this level the locations of base stations, the directions of antennas and the planning that assigns frequencies to base stations is determined. Radio resource reservation policies prevent dropping of existing calls, e.g. by reserving a fraction of the system capacity specifically for handover calls.

Traffic management

The objective of traffic management is to exploit these resources most efficiently while providing adequate quality of service (QoS). Conversely the traffic management mechanisms determine the amount of resources services of various types consume, which is an essential input for network planning. QoS requirements differ per traffic class, while voice users need to receive packets within a threshold of 100 ms [22] in order to allow for a normal phone conversation, data users are more interested in the throughput of the network measured in seconds.

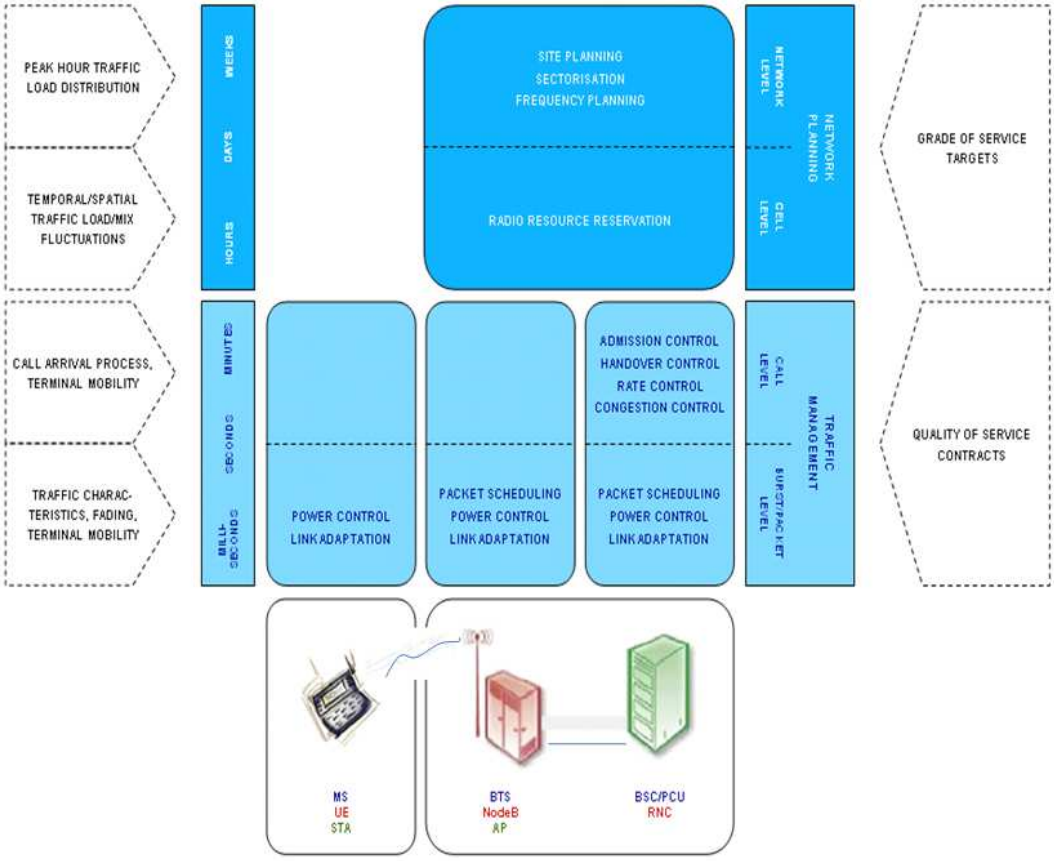


Figure 2.1: Capacity allocation in wireless communication networks: a layered overview of the principal mechanisms and the corresponding time scales [27].

In this research we will focus on the packet scheduling algorithm that acts on a timescale of milliseconds at the packet level of the traffic management. Each service in the wireless network effectively sends and receives small packets of information over the wireless link. Packet scheduling handles the transmission of packets on a per packet basis. Depending on the quality of the wireless channel as well as on possible deadlines for the information packets that need to be sent, the packet scheduler decides which users get access to the wireless channel. Scheduling mechanisms multiplexes the data flows in order to grant the users access. The goal of the packet scheduler is to fulfil such diverse and potentially contradictory aims as resource efficiency optimisation, fairness and QoS differentiation. We will describe the different types of packet schedulers in more detail in Chapter 3.

2.3 Self-organisation of future wireless communication networks

In current wireless networks the optimal settings for capacity allocation mechanisms are determined via studies performed by network operators. Their parameters are optimised via off-line studies, not including real-time changes of the network in the parameter-settings. Network operators set targets for the different QoS measures for the users and tune the parameters they can adapt in the capacity allocation mechanisms in order to satisfy these QoS targets most cost-efficient. Planning and optimisation of capacity allocation mechanisms are labour intensive tasks and, due to the increasing complexity of mobile wireless networks, these costs are increasing.

New applications that require connections of high quality, arise in mobile wireless networks. Both the diversity and the use of these applications is growing. Planning, optimisation and operations of these complex networks require significantly increasing effort. In order to stay competitive in the highly competitive and price oriented telecommunications market, there is a growing trend towards decreasing the operational expenditure for the network operator, whilst increasing capacity, coverage and service quality.

To achieve this goal, wireless communication networks of the future will exhibit a significant degree of self-organisation, as recognised by standardisation body 3rd Generation Partnership Project (3GPP) [1] and operators lobby Next Generation Mobile Networks (NGMN) [2]. The objectives of introducing self-organisation into wireless networks are twofold. The first objective is to reduce operational and capital expenditure (OPEX and CAPEX) by minimising human involvement in network operational tasks. The second objective is to optimise the network capacity, coverage and service quality.

In order to implement self-organisation in wireless communication networks, studies such as those carried out by Van den Berg et al. [10] visualise how future wireless communication networks will use self-organisation methods. Self-optimisation, self-configuration and self-healing are foreseen as promising opportunities to automate wireless network planning and operation, which results in the operational process illustrated in Figure 2.2. Base stations will be added to an existing network and configure themselves via plug-and-play: self-configuration. Once active, base stations may continuously adjust their operational algorithms and parameters due to changes in network, traffic and channel conditions: self-optimisation. Due to automatic optimisation the required quality can be delivered as efficiently as possible. Incidental triggers, such as failure of a base station, will lead to adjustments to the parameter settings of neighbouring base stations in order to limit the consequences for the quality and coverage: self-healing.

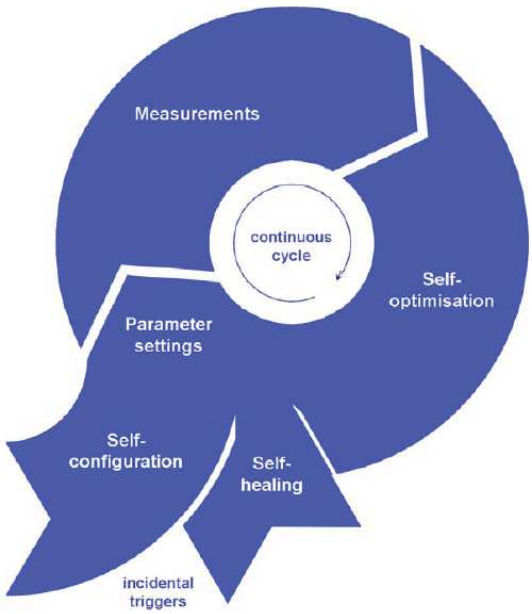


Figure 2.2: Self-organisation in future wireless networks, consisting of self-optimisation, self-configuration and self-healing, as envisioned by Van den Berg et al. [10].

Chapter 3

Packet scheduling in wireless networks

As we have discussed in Chapter 2 the quality of a wireless link between the transmitter and the intended receiver is typically dependent on the time and place of the user within the network. The variable character of the channel quality gives opportunity for the optimal exploitation of these variations. Efficient packet scheduling algorithms allow for this optimal exploration in order to provide for QoS for the different users of the wireless network.

In order to introduce different types of packet schedulers we will start this chapter with a general discussion on the fundamentals of a packet scheduling decision. We discuss the changes in the channel quality that result in changes of the achievable rate at which users of the wireless communications network can be served. We will discuss the different packet scheduling algorithms that can be found in literature. We will end this chapter with a discussion on the packet scheduling algorithm that we will use in this research. This packet scheduling algorithm needs parameters that we can adjust. Different settings of parameter-values will lead to differences in QoS for the users of the wireless network. This will allow us to determine the potential for self-optimisation of packet scheduling.

3.1 The fundamentals of packet scheduling on a wireless link

In wireless communication networks one base station typically provides service for more than one user at a time. Users either have a dedicated channel, such as in GSM or UMTS networks, or a shared channel, such as GPRS, HSDPA and LTE. The general idea behind this shared channel is indicated in Figure 3.1. In this figure we see that dedicated channels will have continue access to (part of the total) resource. In a shared channel the resource is shared, where the multi-user variable channel scheduling problem arises. In this thesis we will constrain ourself to the downlink frequency, on which information is sent from a base station to an intended user. Packet schedulers determine which of the users allocated to the base station are granted service. In order to allocate bandwidth to a user, the network operator has the possibility to use time-multiplexing and to allow for parallel transmission within time intervals. We will describe each of these multiplexing techniques below and denote how we interpret them in this research.

Time multiplexing shares the wireless channel in time instances. The scheduler assigns the bandwidth of the base station on each time interval called the Transmission Time Interval, TTI. These TTIs are small, e.g. 2 ms in a HSDPA network. In Figure 3.1b we see how time is divided into these time intervals.

Parallel transmission assigns the resource, the wireless channel, to multiple users at the same time interval. Assigning more resource to a user will result that this user can be served at a higher bit rate. Higher bit rates allow for faster transmission of the information, thereby increasing the QoS for the user. One way to allow for parallel transmission is by assigning channelisation codes to the users. When a higher number of codes is assigned to a user, this user can be served at a higher bit rate. For technical reasons, there is a limitation to serve up to 4 users in the same time interval [40]. Figure 3.1 shows the general idea behind parallel transmission in shared channels, allowing multiple users on the wireless channel in the same time interval.

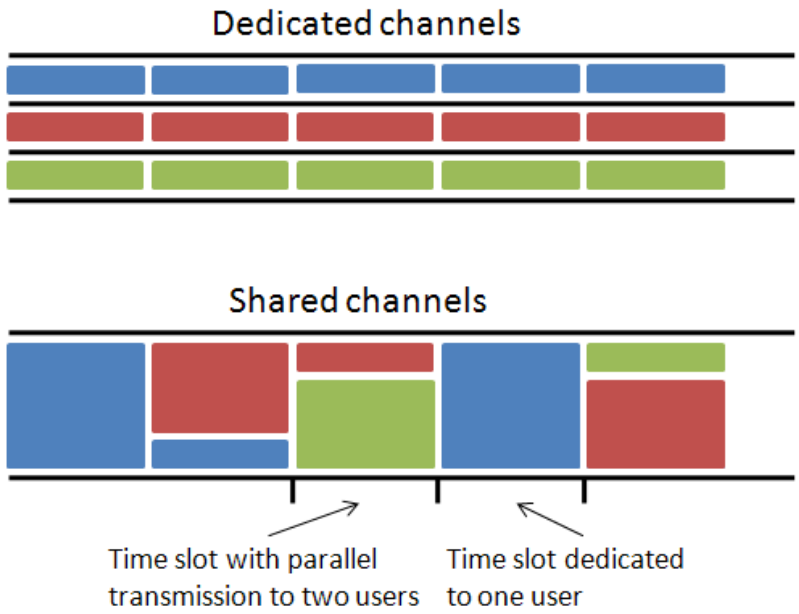


Figure 3.1: The downlink channel of a wireless telecommunication cell serves multiple users. The resource can be allocated in dedicated channels, or in shared channels. Shared channels can allocate the resource on time intervals, within time intervals the channel can allow for parallel transmission.

3.2 Scheduling algorithms

In order to decide which user (or users) have to be served by a base station in a given time interval, packet scheduling algorithms have to discriminate between the users in service. Discrimination is based on the required QoS. We distinguish two main aspects to base a scheduling decision, the queue characteristics of the packets waiting for transmission and the channel quality of the users. As we already denoted in Chapter 2 the channel quality of each user is subjected to different propagation effects, which typically change over time. In this section we will introduce the concepts of queue characteristics that may influence the scheduling decision as well as the channel quality concept and how and why this channel quality changes over time.

3.2.1 Queue characteristics

Queue characteristics are of interest for users that have delay sensitive service. For instance voice users need their packets to be sent within a time limit, e.g. 100 ms [22], in order to allow for adequate quality of this real-time service. The queue characteristics are defined by the first packet

to be served in the queue, the head-of-line packet. This packet has characteristics defined by

$W_i(t)$ = Delay of the first packet of user i at time t
T_i = Maximum delay or due date of the first packet of user i

The delay based scheduling component takes the delay of a packet, its deadline or a combination thereof into account when making the scheduling decision. Basically we can develop any kind of delay based scheduling. Well known methods for taking delay into account are delay based, $W_i(t)$, due date based, $\frac{W_i(t)}{T_i - W_i(t)}$, or based on the exponential rule, $\exp\left[\frac{a_i W_i(t) - \overline{aW}}{1+\sqrt{\overline{aW}}}\right]$, described by a.o. Shakkottai and Stolyar [35].

3.2.2 Channel quality variation

Channel quality varies randomly in time and independently for different users. The variations in channel quality are due to different, varying interference levels observed by different users as well as variable fading of the signal received by a user. We define the channel quality of user i at time t by its Signal-to-Interference-and-Noise-Ratio (SINR), denoted $\xi_i(t)$. For a given instantaneous SINR, it is possible to find the corresponding rate at which the user can be served by the base station. This relation, described by Shannon [36], is given by $r_{SH}(W, \xi_i) \equiv W \log_2(1 + \xi_i)$ (in kbits/sec) where W is the available bandwidth (in kHz). Note that differences in the channel quality lead to different rates at which the user can be served. Precisely this variation is the reason to incorporate channel quality variation in the scheduling decision. Serving users at time instants when their achievable rate is higher will lead to increased capacity of the network, resulting in better QoS with similar network characteristics. Now assume user i is assigned to a base station k and that there are $S\backslash\{k\}$ other interfering base stations transmitting signal to other users in the network. $\xi_i(t)$ is defined by

$$\xi_i(t) = \frac{(P_{tr,k}(t) - P_{pilot})\, G_k(t)}{\sum_{j \in S\backslash\{k\}} P_{tr,j}(t) G_j(t) + \omega P_{tr,k}(t) G_k(t) + N} \tag{3.1}$$

where $P_{tr,j}(t)$ and $P_{tr,k}(t)$ are the transmitting powers of base station j, respectively k at time t. This power has a minimum value P_{pilot}, the power of the pilot channel that is necessary to communicate with the (potential) users. The maximum transmitting power is denoted by P_{max}. In the numerator we substract the pilot power from the transmitting power to obtain the power that is available to transmit information over the traffic channel.

$G_j(t)$ is the gain that the signal transmitted by base station j has when reaching user i at time t. This gain is due to the path loss, shadowing, antenna gains and multipath fading. N is the thermal noise, which is existent in every network. We describe how this gain can be calculated in Chapter 5. ω is the orthogonality factor that determines the amount of intra-cell interference. Furthermore define $\widehat{\xi_i}$ as the average channel SINR for user i and $\widehat{r_i}$ as the average rate for user i.

Channel adaptation identifies how the rate of the current channel state, $r_i(t)$, is taken into consideration in the scheduling decision. In order to benefit from the stochastic behavior of the channel variations and to incorporate this phenomenon three main forms of channel adaptation scheduling are known, which are described in various papers, o.a. Van den Berg et al. [9].

Round Robin (RR)

The basic principle of channel-oblivious RR scheduling is that each base station cyclically assigns its channel to the different flows it serves.

Proportional Fair (PF)

The design principle of the channel aware PF scheme is to exploit the channel variations for throughput enhancement, while still providing fairness and avoiding starvation of flows with relatively poor link quality. To this end, in the considered scheme the base station assigns its channel to flow

$$i^* = \arg\max_i \frac{r_i(t)}{\widehat{r_i(t)}} \tag{3.2}$$

where $\widehat{r_i}$ denotes the average rate of user i. In actual live networks and in simulation based literature $\widehat{r_i}$ is calculated over a specified historical period by

$$\widehat{r_i}(t) = (1-\alpha)\,\widehat{r_i}(t-1) + \alpha r_i(t) \tag{3.3}$$

with $\alpha \in [0,1]$ the associated smoothing parameter and $\widehat{r_i}(t_0) = r_i(t_0)$ the assumed initial value at the flow's generation time t_0. It is readily verified that for $\alpha = 0$, respectively $\alpha = 1$ the PF scheduler is equivalent to the maximum SINR (see below), respectively RR, scheduler.

In analytical literature we schedule on SINR, thus by

$$i^* = \arg\max_i \frac{\xi_i(t)}{\widehat{\xi_i}} \tag{3.4}$$

where we assume $\widehat{\xi_i}$ to be known and constant over time. We use the expression of the Shannon-rate to determine the rate in the analysis. We will further discuss this approach in Chapter 4.

Maximum SINR

The channel aware Maximum SINR scheme exploits the channel variations for throughput enhancement even further than the PF scheme, though without providing fairness and by allowing starvation of flows with relatively poor link quality. In the considered scheme the base station assigns its channel to the flow with the highest instantaneous rate.

$$i^* = \arg\max_i r_i(t) \tag{3.5}$$

Again we schedule on SINR, $\xi_i(t)$, in analytical literature. Key difference between the Round Robin and Proportional Fair scheme on the one side and Maximum SINR scheduling on the other side is the inherent fairness property of the first two schemes. Flows obtain equal long term access to the base stations resources, regardless of their respective average SINRs.

3.3 Packet schedulers in literature

Packet schedulers can be identified on their channel adaptive and their delay based component. Besides this, schedulers are differentiated in homogeneous schedulers, which consider no differentiation between traffic types, and heterogeneous schedulers, which incorporate multiple classes of users thereby allowing differentiation between traffic types. Often the classes are based on the type of service of the users in the network. In order to allow for differentiation of the classes of users, some form of relative importance for the user is set by γ_i. Often users within the same class are given the same value for γ_i. Table 3.1 lists the schedulers we found in literature and their scheduling decision maximization formulas. For each scheduler we indicate the usage of channel adaptation, the incorporation of delay aspects and whether the scheduler is homogeneous or heterogeneous.

Scheduler	Maximization formula	Channel adaptive	Delay aspect	Type
RR	-	No	No	Homogenous
Max SINR	$r_i(t)$	Max SINR	No	Homogenous
PF	$\frac{r_i(t)}{\widehat{r_i(t)}}$	PF	No	Homogenous
FIFO	$W_i(t)$	No	Delay	Homogenous
EDD	$\frac{W_i(t)}{T_i - W_i(t)}$	No	Due Date	Homogenous
CD-EDD	$\frac{r_i(t)}{\widehat{r_i}} \left(\frac{W_i(t)}{T_i - W_i(t)} \right)$	PF	Due Date	Homogenous
LWDF	$\gamma_i W_i(t)$	No	Delay	Heterogenous
M-LWDF	$\gamma_i r_i(t) W_i(t)$	Max SINR	Delay	Heterogenous
Exp rule	$\gamma_i \frac{r_i(t)}{\widehat{r_i}} \exp \left[\frac{a_i W_i(t) - \overline{aW}}{1+\sqrt{\overline{aW}}} \right]$	PF	Delay	Heterogenous
CD-EDD	$\gamma_i \frac{r_i(t)}{\widehat{r_i}} \frac{W_i(t)}{T_i - W_i(t)}$	PF	Due Date	Heterogenous

Table 3.1: The schedulers we found in literature and their scheduling decision maximization formulas. For each scheduler we indicated the usage of channel adaptation, the incorporation of delay aspects and whether the scheduler is homogeneous or heterogeneous.

Hierarchical scheduling

It is possible to schedule different classes of users heterogenously with an hierarchical scheduling method, as described by Necker [31]. The hierarchical scheduler subdivides the scheduling decision to two levels. The first level, the traffic class scheduler, makes the high-level scheduling decision that determines which traffic class is scheduled. This can be accomplished with a Weighted Round Robin mechanism, which assigns a weighted share of the available resource to each user class, or a Static Prioritisation mechanism, which allows absolute priority of one traffic class over the others or by some other form of (weighted) priority of one traffic class over the other. Within each traffic class, the traffic class scheduler makes a scheduling decision based on one a scheduler that is composed of a channel adapive and/or delay based component.

3.4 Scope of the research

In this thesis we will distinguish between a network loaded with data traffic only and a network loaded with both data and voice traffic. As indicated in Chapter 2 we will simulate a network loaded with data traffic and a network loaded with both data and voice traffic. We will develop a mathematical analysis as well, which can be used for a network loaded with data traffic only. We will distinguish between the schedulers considered in the mathematical analysis and the simulation study. We will describe the mathematical analysis in Chapter 4 and the corresponding numerical results in Chapter 6.

Mathematical analysis

In the mathematical analysis, we consider a network in use by data users only. We therefore will consider the Round Robin, Proportional Fair and maximum SINR scheduler. As discussed above, we will take the analytical expression into account and base our scheduling decision on the SINR values, where we assume the average SINR $\widehat{\xi_i}$ to be known. We will describe the mathematical analysis in Chapter 4 and the corresponding numerical results in Chapter 6.

Simulation study

The network in the simulation study is loaded with both voice and data traffic. Therefore we want the scheduling mechanism we consider to be both 'channel sensitive' and 'delay sensitive'. Voice users have a deadline for their packets of typically 100 ms [22]. In order to satisfy the deadline of a packet, we should incorporate this deadline in the scheduling decision. Furthermore we incorporate channel sensitivity to exploit channel variations. In the simulation study we will therefore base our scheduling decision on an adjusted version of the CD-EDD scheduler, cf. Table 3.1

$$\arg\max_i = \frac{r_i(t)}{\widehat{r_i}} \left(1 + \frac{W_i(t)}{T_i - W_i(t)}\right)^{\zeta} \tag{3.6}$$

where $\frac{W_i(t)}{T_i - W_i(t)} = 0$ for data users. Furthermore recall from Equation (3.3) that $\widehat{r_i}(t) = (1-\alpha)\,\widehat{r_i}(t-1) + \alpha r_i(t)$.

The objective of our research is to investigate the sensitivity of the optimal packet scheduling parameters in wireless networks with respect to variations in network, traffic and channel conditions. Therefore we chose for the channel quality sensitive Proportional Fair expression. To determine the influence of the deadline component of the scheduler, we initiated the parameter ζ, which gives relative importance of the 'deadline sensitivity' to the 'channel sensitivity'. In Chapter 5 we will introduce the model simulation used in our simulation study. In Chapter 6 we will define the simulation scenarios and the parameter values for α and ζ that we consider.

Chapter 4

Mathematical modelling and analysis

In this chapter we develop a mathematical model to analyse the performance of different packet scheduling schemes in different fading environments in a wireless network. As indicated in Chapter 3 we will analyse the Round Robin, the Proportional Fair and the maximum SINR scheduler. Previous analytical research on scheduling mechanisms that incorporate multipath fading conducted by Berggren and Janti [7] and Litjens and Berggren [8, 29] assume a Rayleigh fading model, modelled by a distribution of $\xi_i/\widehat{\xi_i} \sim Exp(1)$*. The variance of this distribution characterises a high degree of multipath fading. We will generalise this fading model, making it possible to adjust the distribution of* $\xi_i/\widehat{\xi_i}$

4.1 Introduction

The goal of this chapter is to compare the performance of different scheduling schemes in different fading environments using an analytical approach. We want to compare the different scheduling schemes based on their performance at flow level. We incorporate the flow level dynamics due to flow (file) transfer completions and initiations by the users at random time instants, which causes the number of ongoing flow transfers to vary over time. The differences in fading environment and scheduling mechanisms lead to differences in the flow transfer completions. In this chapter we will develop an analytical model that allow us to calculate the rate with which these flow transfer completions occur, based on the fading environment and scheduling mechanisms.

We aim at quantifying the performance measure throughput, expressing how the performance depends on the users average channel quality, the fading environment and the scheduling mechanism under consideration. Furthermore we are interested in the performance throughput of the users which have the worst channel quality. In a mobile network the throughput that a typical user can experience is dependent on its distance to the base station, antenna gains, shadowing aspects and multipath fading. Users that are located further from the base station typically have a channel of worse quality than users located close to the base station and therefore typically receive a lower throughput. These users are called 'cell edge' users and we will quantify their performance throughput in this chapter. We will introduce a model that categorizes users based on their average channel quality, $\widehat{\xi_i}$, in L types.

In the next section we will describe the mathematical model. Section 4.3 describes the calculations of the long-term average throughputs for users in the network. We use these long-term average throughputs in the Markoc chain model we define in Section 4.4. In Section 4.5 we define the QoS measures that we use to determine the performance of the different packet scheduling mechanisms in the different fading environments considered. The numerical results of the mathematical model

described in this chapter are discussed in Chapter 6.

4.2 Mathematical modelling

We will consider a single cell scenario with users generating elastic traffic flows. Our modelling and analysis approach is based on time scale decomposition and consists basically of two steps. The first step takes the details of the scheduler's behaviour into account in a given state of the system, i.e. the number of users and their average channel quality measured in SINR. The step determines the user's average throughput. In the second step these throughputs and the rates at which new users become active are used to create a continuous-time Markov chain describing the system behaviour at flow level. From the steady-state distribution of the Markov chain the performance measures, such as the average throughput for a user of an arbitrary type, can be calculated. Our approach to capture packet level dynamics by analysis and flow level dynamics by Monte Carlo simulation is similar to the approach described by Dimitrova et al. [17]. The main difference is that this research focusses on the performance of the downlink, whereas the paper by Dimitrova et al. has a clear focus on the (enhanced) uplink.

We describe the modelling assumptions underlying the presented analysis. At system level, we consider the downlink of a single cell. Calls arriving at the considered base station are split in L types. The average SINR for users of type i is assumed to be known and is denoted by $\widehat{\xi_i}$.

A number of assumptions are made at the user level. Calls are generated according to spatially uniform Poisson arrival processes with rate λ. For the performance of calls it matters of which type they are. As a direct consequence of the uniformity assumption, the probability q_i that generated call is of type i is $1/L$, so that the call arrival rate of type i users is $\lambda_i = \lambda q_i$. Calls are characterised by a file that needs to be downloaded, whose size is exponentially distributed with mean F (in kbits). As no user mobility is considered, users keep their positions in the cell during the file transmission. The bit rate at which a call is served depends on the average SINR of the user type and the scheduling method, as we will see in the perfomance analysis section 4.3.

At a given time, the system state $\underline{n} = (n_1, n_2, .., n_L)$ is described by the number of calls n_i of type i, $i = 1, .., L$. The instantaneous SINR of a flow type i is modelled as a random variable denoted ξ_i with average $\widehat{\xi_i}$. The variability of the instantaneous SINR will be modelled using three fittings of the distribution for $\xi_i/\widehat{\xi_i}$.

1. For the first fitting we consider that the variability of the instantaneous SINR is caused by small-scale Rayleigh fading, which is reflected by modelling $\xi_i/\widehat{\xi_i} \sim \text{Exp}(1)$ as is done in literature by e.g. Berggren and Janti [7] and Litjens and Berggren [8, 29].

2. For the second fitting we model the variability as caused by the Ricean fading model [20, 26, 33, 41] and determine a general distribution, we use an Erlang(β, k) distribution. We assume an identical distribution for all types of users, so $\xi_i/\widehat{\xi_i} \sim \text{Erlang}(\beta, k)$.

3. For the third fitting we model the variability by fitting an individual distribution, we use an Erlang(β_i, k_i) distribution on each of the types individually, so $\xi_i/\widehat{\xi_i} \sim \text{Erlang}(\beta_i, k_i)$.

For a given instantaneous SINR, a link adaptation scheme is assumed to provide a suitable transmission rate when the flow is scheduled for service, viz. $r_{SH}(W, \xi_i) \equiv W \log_2(1 + \xi_i)$ [36], where W denotes the available bandwidth. The resulting instantaneous rates lead to long-term average throughput values for users of type i in a given state $\underline{n}$. We use these long-term average throughput values to calculate the transition rates in the Markov chain model, as will be discussed in Section 4.4. The resulting Markov Jump queue can then be analysed or simulated to obtain the flow level performance for each of the types of users in the network. Figure 4.1 shows the resulting transition state diagram of the continuous-time Markov chain describing the system at flow level for the case $L = 2$.

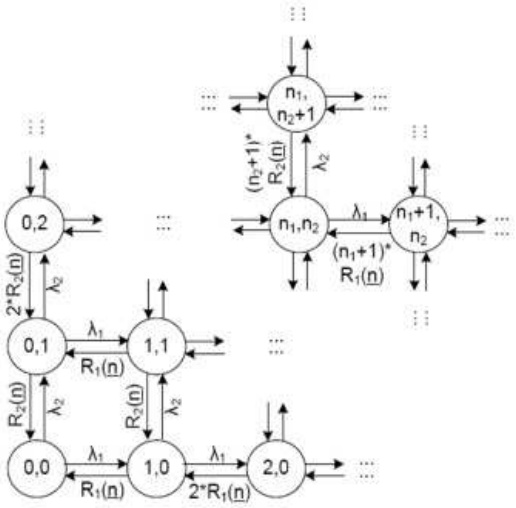


Figure 4.1: Transition state diagram of the continuous-time Markov chain describing the system at flow level for the case $L = 2$.

From Figure 4.1 we see that the long-term average throughput values for users of type i, $R_i(\underline{n})$ are of essential influence in the determination of the steady state of the Markov Chain model. We will determine the influence of the different scheduling schemes and the different fading environments by determining the associated long-term average throughput values based on the scheduling scheme and the fading environment under consideration. As discussed we will determine the influence of the different fading environments by fitting an Erlang distribution on to $\xi_i/\widehat{\xi_i}$. We consider three fittings in our analysis, $\xi_i/\widehat{\xi_i} \sim \text{Exp}(1)$, $\xi_i/\widehat{\xi_i} \sim \text{Erlang}(\beta, k$ and $\xi_i/\widehat{\xi_i} \sim \text{Erlang}(\beta_i, k_i)$. In the next section we will determine the long-term average throughput $R_i(\underline{n})$ for each of these fittings.

4.3 Performance analysis

As discussed in the previous section the modelling approach consist of two steps. The first step is to take the details of the scheduler's behaviour into account in a given state of the system to determine the user's average throughput. In this section we describe this first step and determine the long-term average throughput rate of calls of type i, $R_i(\underline{n})$, based on the system state and scheduling scheme. These rates will be input for the transition rates of the Markov model jump queue. We distinct the three fittings as identified in Section 4.2. For each of these fittings we consider the access selection based on scheduling via Round Robin, Proportional Fair and Maximum SINR, as described in Chapter 3. For the analysis of the Proportional Fair scheduler we assume the idealised version of the Proportional Fair scheduler, where we assume the average SINR $\widehat{\xi_i}$ to be known.

In the second step discussed in Section 4.2, we use these throughputs and the rates at which new users become active are used to create a continuous-time Markov chain describing the system behaviour at flow level. From the steady-state distribution of the Markov chain the performance measures, such as the average throughput for a user of an arbitrary type, can be calculated.

We will now describe how to determine the long-term average throughput rate of calls of type i, $R_i(\underline{n})$, based on the system state and scheduling scheme for the three fittings we consider in our analysis.

4.3.1 Fitting 1: Modelling $\xi_i/\widehat{\xi_i} \sim \mathbf{Exp}(1)$

We will determine the long-term average rate for calls of type i, $R_i(\underline{n})$ under the assumption that $\xi_i/\widehat{\xi_i}$ is distributed according to Exponential(1). We consider the scheduling schemes Round Robin, Proportional Fair and Maximum SINR.

Round robin access selection

Define the auxiliary function

$$\begin{aligned} H(k,\widehat{\xi_i}) &\equiv \int_0^\infty r_{SH}(W,\xi_i)\frac{1}{\widehat{\xi_i}}e^{-\frac{\xi_i k}{\widehat{\xi_i}}}d\xi_i \\ H(k,\widehat{\xi_i}) &= W\frac{e^{k/\widehat{\xi_i}}}{k\ln 2}\Gamma\left(\frac{k}{\widehat{\xi_i}}\right) \end{aligned} \tag{4.1}$$

where $\Gamma(x) = \int_x^\infty t^{-1}e^{-t}dt$ is an incomplete Gamma function, applying substitution and partial integration techniques to obtain the final expression. For $k = 1$, $H(1,\widehat{\xi_i})$ is readily seen to express the long-term average throughput experienced by an isolated flow with average SINR $\widehat{\xi_i}$, which is obtained by conditioning on the exponentially distributed instantaneous SINR. Given a presence of $N = \sum_{i=1}^{L} n_i$ flows, the fraction of time a single flow of type i is served is equal to $1/N$ due to the nature of the RR scheme. The long-term average throughput experienced by a flow of type i is then

$$R_i(\underline{n}) = \frac{1}{N}H(1,\widehat{\xi_i}) \tag{4.2}$$

Note that the long-term average throughput experienced by a flow of type i is only dependent on the total amount of users present in the system.

Proportional Fair access selection

As introduced above, the basic principle of PF access selection is to schedule a flow based on the ratio $\xi_i/\widehat{\xi_i}$ of the instantaneous and average SINRs in order to enhance throughput while preserving fairness. Given the instantaneous SINR values, the probability that a given flow of type i is served by the RA is equal to

$$\begin{aligned} Pr\left(flow\ i\ is\ served|\xi_i\right) &= Pr\left(\frac{\frac{\xi_i}{\widehat{\xi_i}} \geq \frac{\xi_j}{\widehat{\xi_j}}, \forall j \neq i}\right) \\ Pr\left(flow\ i\ is\ served|\xi_i\right) &= \left(1 - e^{-\frac{\xi_i}{\widehat{\xi_i}}}\right)^{N-1} \end{aligned} \tag{4.3}$$

which is noted to be increasing in ξ_i and not dependent on the SINRs of the competing calls. Deconditioning Expression (4.3) with respect to the experienced SINR ξ_i we find

$$\begin{aligned} Pr\left\{flow\ i\ is\ served\right\} &= \int_0^\infty \left(1 - e^{-\frac{\xi_i}{\widehat{\xi_i}}}\right)^{N-1} \frac{1}{\widehat{\xi_i}} e^{-\frac{\xi_i}{\widehat{\xi_i}}} d\xi_i \\ Pr\left\{flow\ i\ is\ served\right\} &= \left[\frac{1}{N}\left(1 - e^{-\frac{\xi_i}{\widehat{\xi_i}}}\right)^{N}\right]_0^\infty \\ Pr\left\{flow\ i\ is\ served\right\} &= \frac{1}{N} \end{aligned} \tag{4.4}$$

expressing the above mentioned fairness property, i.e. the PF scheme indeed establishes fair access to the resources, regardless of a flow's average SINR, which is also immediately clear from the

fact that access selection is based on the ratio $\xi_i/\widehat{\xi_i}$, which is probabilistically identical for all flows.

Under the PF scheme, the long term average throughput of a flow i is then

$$\begin{aligned} R_i(\underline{n}) &= \int_0^\infty r_{SH}(W,\xi_i)\left(1-e^{-\frac{\xi_i}{\widehat{\xi_i}}}\right)^{N-1}\frac{1}{\widehat{\xi_i}}e^{-\frac{\xi_i}{\widehat{\xi_i}}}d\xi_i \\ R_i(\underline{n}) &= \sum_{k=0}^{N-1}\binom{N-1}{k}(-1)^k\int_0^\infty r_{SH}(W,\xi_i)\frac{1}{\widehat{\xi_i}}e^{-\frac{\xi_i}{\widehat{\xi_i}}}d\xi_i \\ R_i(\underline{n}) &= \frac{1}{N}\sum_{k=1}^{N}\binom{N}{k}(-1)^{k+1}H\left(k,\widehat{\xi_i}\right) \end{aligned} \tag{4.5}$$

applying Newton's binomium with $H(\cdot)$ again as defined in Expression (4.1). It can be verified that Expression (4.5) is larger than Expression (4.2), basically since $Pr(flow\ i\ is\ served|\xi_i)$ is increasing in ξ_i while overall access is fair, i.e. a flow is more likely to be granted its fair share of access time at instances where it experiences a more favourable radio link quality and thus a higher potential bit rate. This establishes the multi-user diversity gain that is achieved by the PF scheme. Again, note that the long-term average throughput experienced by a flow of type i is only dependent on the total number of users present in the system.

Maximum SINR access selection

As introduced above, the basic principle of Maximum SINR access selection is to schedule a flow based on its instantaneous SINR ξ_i. We know that $\xi_i/\widehat{\xi_i} \sim \text{Exp}(1)$, so $\xi_i \sim \text{Exp}(1/\widehat{\xi_i})$. Given the instantaneous SINR values, the probability that a given flow of type i is served by the RA is equal to

$$\begin{aligned} Pr\left(flow\ i\ is\ served|\xi_i\right) &= Pr\left(\xi_i \geq \xi_j, \forall j \neq i\right) \\ Pr\left(flow\ i\ is\ served|\xi_i\right) &= \left(1-e^{-\frac{\xi_i}{\xi_1}}\right)^{n_1}\left(1-e^{-\frac{\xi_i}{\xi_2}}\right)^{n_2}... \\ &\quad ...\left(1-e^{-\frac{\xi_i}{\xi_i}}\right)^{n_i-1}...\left(1-e^{-\frac{\xi_i}{\xi_J}}\right)^{n_J} \end{aligned} \tag{4.6}$$

which is noted to be increasing in ξ_i and though not independent on the SINRs of the competing calls. Deconditioning Expression (4.6) with respect to the experienced SINR ξ_i we find

$$\begin{aligned} Pr\left\{flow\ i\ is\ served\right\} &= \int_0^\infty \left(1-e^{-\frac{\xi_i}{\xi_1}}\right)^{n_1}\left(1-e^{-\frac{\xi_i}{\xi_2}}\right)^{n_2}...\left(1-e^{-\frac{\xi_i}{\xi_{n_i}}}\right)^{n_i-1} \\ &\quad ...\left(1-e^{-\frac{\xi_i}{\xi_{n_J}}}\right)^{n_J}\frac{1}{\widehat{\xi_i}}e^{-\frac{\xi_i}{\xi_i}}d\xi_i \end{aligned} \tag{4.7}$$

Under the Maximum SINR scheme, the long term average throughput of flow i is then

$$\begin{aligned} R_i(\underline{n}) &= \int_0^\infty r_{SH}(W,\xi_i)\left(1-e^{-\frac{\xi_i}{\xi_1}}\right)^{n_1}\left(1-e^{-\frac{\xi_i}{\xi_2}}\right)^{n_2}...\left(1-e^{-\frac{\xi_i}{\xi_i}}\right)^{n_i-1} \\ &\quad ...\left(1-e^{-\frac{\xi_i}{\xi_J}}\right)^{n_J}\frac{1}{\widehat{\xi_i}}e^{-\frac{\xi_i}{\xi_i}}d\xi_i \end{aligned} \tag{4.8}$$

The long-term average throughput experienced by a flow of type i is thus dependent on the total state $(\underline{n})$ of users present in the system.

4.3.2 Fitting 2: Modelling $\xi_i/\widehat{\xi_i} \sim$ Erlang(β, k)

We will determine the long-term average rate for calls of type i, $R_i(\underline{n})$ under the assumption that $\xi_i/\widehat{\xi_i}$ is distributed according to Erlang(β, k). We consider the scheduling schemes Round Robin, Proportional Fair and Maximum SINR.

Round robin access selection

Given a presence of $N = \sum_{i=1}^{L} n_i$ flows, the fraction of time a single flow of type i is served is equal to $1/N$ due to the nature of the RR scheme. The long-term average throughput experienced by a flow of type i is then

$$R_i(\underline{n}) = \frac{1}{N}\int_0^\infty r_{SH}(W,\xi_i)\frac{\left(\frac{\beta}{\widehat{\xi_i}}\right)^k \xi_i^{k-1}}{(k-1)!}e^{-\beta\frac{\xi_i}{\widehat{\xi_i}}}d\xi_i \tag{4.9}$$

Note that the long-term average throughput experienced by a flow of type i is only dependent on the total amount of users present in the system.

Proportional Fair access selection

As introduced above, the basic principle of PF access selection is to schedule a flow based on the ratio $\xi_i/\widehat{\xi_i}$ of the instantaneous and average SINRs in order to enhance throughput while preserving fairness. Assuming identical probability density functions for $\xi_i/\widehat{\xi_i}$ for all types of users, the probability that a given flow of type i is served by the RA is equal to

$$\begin{aligned} Pr\left(flow\ i\ is\ served|\xi_i\right) &= Pr\left(\frac{\xi_i}{\widehat{\xi_i}} \geq \frac{\xi_j}{\widehat{\xi_j}}, \forall j \neq i\right) \\ Pr\left(flow\ i\ is\ served|\xi_i\right) &= \left(F(\frac{\xi_i}{\widehat{\xi_i}})\right)^{N-1} \end{aligned} \tag{4.10}$$

Which gives us for the distribution $\xi_i/\widehat{\xi_i} \sim$ Erlang(β, k)

$$Pr\left(flow\ i\ is\ served|\xi_i\right) = \left(1 - \sum_{j=0}^{k-1} e^{-\beta\frac{\xi_i}{\widehat{\xi_i}}}\frac{\left(\beta\frac{\xi_i}{\widehat{\xi_i}}\right)^j}{j!}\right)^{N-1} \tag{4.11}$$

which is noted to be increasing in ξ_i and not dependent on the SINRs of the competing calls. Deconditioning Expression (4.11) with respect to the experienced SINR ξ_i we find

$$\begin{aligned} Pr\left\{flow\ i\ is\ served\right\} &= \int_0^\infty \left(1 - \sum_{j=0}^{k-1} e^{-\beta\frac{\xi_i}{\widehat{\xi_i}}}\frac{\left(\beta\frac{\xi_i}{\widehat{\xi_i}}\right)^j}{j!}\right)^{N-1} \frac{\left(\frac{\beta}{\widehat{\xi_i}}\right)^k \xi_i^{k-1}}{(k-1)!}e^{-\beta\frac{\xi_i}{\widehat{\xi_i}}}d\xi_i \\ Pr\left\{flow\ i\ is\ served\right\} &= \frac{1}{N} \end{aligned} \tag{4.12}$$

Under the PF scheme, the long term average throughput of a flow i under assumption of $\xi_i/\widehat{\xi_i} \sim$ Erlang(β, k) is then

$$R_i(\underline{n}) = \int_0^\infty r_{SH}(W,\xi_i)\left(1 - \sum_{j=0}^{k-1} e^{-\beta\frac{\xi_i}{\widehat{\xi_i}}}\frac{\left(\beta\frac{\xi_i}{\widehat{\xi_i}}\right)^j}{j!}\right)^{N-1} \frac{\left(\frac{\beta}{\widehat{\xi_i}}\right)^k \xi_i^{k-1}}{(k-1)!}e^{-\beta\frac{\xi_i}{\widehat{\xi_i}}}d\xi_i \tag{4.13}$$

The long-term average throughput experienced by a flow of type i is only dependent on the total amount of users present in the system.

Maximum SINR access selection

As introduced above, the basic principle of Maximum SINR access selection is to schedule a flow based on its instantaneous SINR ξ_i. We know that $\xi_i/\widehat{\xi_i} \sim \text{Erlang}(\beta, k)$, so $\xi_i \sim \text{Erlang}(\beta/\widehat{\xi_i}, k)$. Assuming identical probability density functions for $\xi_i/\widehat{\xi_i}$ for all types of users, the probability that a given flow of type i is served by the RA is equal to

$$Pr\left(flow\ i\ is\ served|\xi_i\right) \quad = \quad Pr\left(\xi_i \geq \xi_j, \forall j \neq i\right)$$

$$Pr\left(flow\ i\ is\ served|\xi_i\right) = \left(1 - \sum_{j=0}^{k-1} e^{-\beta\frac{\xi_i}{\xi_1}} \frac{\left(\beta\frac{\xi_i}{\xi_1}\right)^j}{j!}\right)^{n_1} \left(1 - \sum_{j=0}^{k-1} e^{-\beta\frac{\xi_i}{\xi_2}} \frac{\left(\beta\frac{\xi_i}{\xi_2}\right)^j}{j!}\right)^{n_2} ... \\ ... \left(1 - \sum_{j=0}^{k-1} e^{-\beta\frac{\xi_i}{\xi_i}} \frac{\left(\beta\frac{\xi_i}{\xi_i}\right)^j}{j!}\right)^{n_i-1} ... \left(1 - \sum_{j=0}^{k-1} e^{-\beta\frac{\xi_i}{\xi_J}} \frac{\left(\beta\frac{\xi_i}{\xi_J}\right)^j}{j!}\right)^{n_J} \quad (4.14)$$

which is noted to be increasing in ξ_i, though not independent on the SINRs of the competing calls. Deconditioning Expression (4.14) with respect to the experienced SINR ξ_i we find

$$Pr\left\{flow\ i\ is\ served\right\} = \int_0^\infty \left(1 - \sum_{j=0}^{k-1} e^{-\beta\frac{\xi_i}{\xi_1}} \frac{\left(\beta\frac{\xi_i}{\xi_1}\right)^j}{j!}\right)^{n_1} \left(1 - \sum_{j=0}^{k-1} e^{-\beta\frac{\xi_i}{\xi_2}} \frac{\left(\beta\frac{\xi_i}{\xi_2}\right)^j}{j!}\right)^{n_2} ... \\ ... \left(1 - \sum_{j=0}^{k-1} e^{-\beta\frac{\xi_i}{\xi_i}} \frac{\left(\beta\frac{\xi_i}{\xi_i}\right)^j}{j!}\right)^{n_i-1} ... \left(1 - \sum_{j=0}^{k-1} e^{-\beta\frac{\xi_i}{\xi_J}} \frac{\left(\beta\frac{\xi_i}{\xi_J}\right)^j}{j!}\right)^{n_J} \frac{\left(\frac{\beta}{\xi_i}\right)^k \xi_i^{k-1}}{(k-1)!} e^{-\beta\frac{\xi_i}{\xi_i}} d\xi_i \quad (4.15)$$

Under the Maximum SINR scheme, the long term average throughput of flow i under assumption of $\xi_i/\widehat{\xi_i} \sim \text{Erlang}(\beta, k)$ is then

$$R_i(\underline{n}) = \int_0^\infty r_{SH}(W, \xi_i) \left(1 - \sum_{j=0}^{k-1} e^{-\beta\frac{\xi_i}{\xi_1}} \frac{\left(\beta\frac{\xi_i}{\xi_1}\right)^j}{j!}\right)^{n_1} \left(1 - \sum_{j=0}^{k-1} e^{-\beta\frac{\xi_i}{\xi_2}} \frac{\left(\beta\frac{\xi_i}{\xi_2}\right)^j}{j!}\right)^{n_2} ... \\ ... \left(1 - \sum_{j=0}^{k-1} e^{-\beta\frac{\xi_i}{\xi_i}} \frac{\left(\beta\frac{\xi_i}{\xi_i}\right)^j}{j!}\right)^{n_i-1} ... \left(1 - \sum_{j=0}^{k-1} e^{-\beta\frac{\xi_i}{\xi_J}} \frac{\left(\beta\frac{\xi_i}{\xi_J}\right)^j}{j!}\right)^{n_J} \frac{\left(\frac{\beta}{\xi_i}\right)^k \xi_i^{k-1}}{(k-1)!} e^{-\beta\frac{\xi_i}{\xi_i}} d\xi_i \quad (4.16)$$

The long-term average throughput experienced by a flow of type i is thus dependent on the total state $(\underline{n})$ of users present in the system.

4.3.3 Fitting 3: Modelling $\xi_i/\widehat{\xi_i} \sim$ Erlang(β_i, k_i)

We will determine the long-term average rate for calls of type i, $R_i(\underline{n})$ under the assumption that $\xi_i/\widehat{\xi_i}$ is distributed according to Erlang(β_i, k_i). We consider the scheduling schemes Round Robin, Proportional Fair and Maximum SINR.

Round robin access selection

Given a presence of $N = \sum_{i=1}^{L} n_i$ flows, the fraction of time a single flow of type i is served is equal to $1/N$ due to the nature of the RR scheme. The long-term average throughput experienced

by a flow of type i is then

$$R_i(\underline{n}) = \frac{1}{N}\int_0^\infty r_{SH}(W,\xi_i)\frac{\left(\frac{\beta_i}{\widehat{\xi_i}}\right)^{k_i}\xi_i^{k_i-1}}{(k_i-1)!}e^{-\beta_i\frac{\xi_i}{\widehat{\xi_i}}}d\xi_i \tag{4.17}$$

Note that the long-term average throughput experienced by a flow of type i is only dependent on the total amount of users present in the system.

Proportional Fair access selection

As introduced above, the basic principle of PF access selection is to schedule a flow based on the ratio $\xi_i/\widehat{\xi_i}$ of the instantaneous and average SINRs in order to enhance throughput while preserving fairness. Assuming identical probability density functions for $\xi_i/\widehat{\xi_i}$ for all types of users, the probability that a given flow of type i is served by the RA is equal to

$$Pr\left(flow\ i\ is\ served|\xi_i\right) = Pr\left(\frac{\xi_i}{\widehat{\xi_i}} \geq \frac{\xi_j}{\widehat{\xi_j}}, \forall j \neq i\right)$$

$$Pr\left(flow\ i\ is\ served|\xi_i\right) = \left(F_1(\frac{\xi_i}{\widehat{\xi_i}})\right)^{n_1}\left(F_2(\frac{\xi_i}{\widehat{\xi_i}})\right)^{n_2}...\left(F_i(\frac{\xi_i}{\widehat{\xi_i}})\right)^{n_i-1}...\left(F_J(\frac{\xi_i}{\widehat{\xi_i}})\right)^{n_J} \tag{4.18}$$

Which gives us for the distribution $\xi_i/\widehat{\xi_i} \sim \text{Erlang}(\beta_i, k_i)$

$$\begin{aligned}Pr\left(flow\ i\ is\ served|\xi_i\right) = &\left(1-\sum_{j=0}^{k_1-1}e^{-\beta_1\frac{\xi_i}{\xi_i}}\frac{\left(\beta_1\frac{\xi_i}{\widehat{\xi_i}}\right)^j}{j!}\right)^{n_1}\left(1-\sum_{j=0}^{k_2-1}e^{-\beta_2\frac{\xi_i}{\xi_i}}\frac{\left(\beta_2\frac{\xi_i}{\widehat{\xi_i}}\right)^j}{j!}\right)^{n_2}...\\ &...\left(1-\sum_{j=0}^{k_i-1}e^{-\beta_i\frac{\xi_i}{\xi_i}}\frac{\left(\beta_i\frac{\xi_i}{\widehat{\xi_i}}\right)^j}{j!}\right)^{n_i-1}...\left(1-\sum_{j=0}^{k_J-1}e^{-\beta_J\frac{\xi_i}{\xi_i}}\frac{\left(\beta_J\frac{\xi_i}{\widehat{\xi_i}}\right)^j}{j!}\right)^{n_J}\end{aligned} \tag{4.19}$$

which is noted to be increasing in ξ_i and not independent on the SINRs of the competing calls. Deconditioning Expression (4.19) with respect to the experienced SINR ξ_i we find

$$\begin{aligned}Pr\left\{flow\ i\ is\ served\right\} = &\int_0^\infty\left(1-\sum_{j=0}^{k_1-1}e^{-\beta_1\frac{\xi_i}{\xi_i}}\frac{\left(\beta_1\frac{\xi_i}{\widehat{\xi_i}}\right)^j}{j!}\right)^{n_1}...\\ &...\left(1-\sum_{j=0}^{k_2-1}e^{-\beta_2\frac{\xi_i}{\xi_i}}\frac{\left(\beta_2\frac{\xi_i}{\widehat{\xi_i}}\right)^j}{j!}\right)^{n_2}...\left(1-\sum_{j=0}^{k_i-1}e^{-\beta_i\frac{\xi_i}{\xi_i}}\frac{\left(\beta_i\frac{\xi_i}{\widehat{\xi_i}}\right)^j}{j!}\right)^{n_i-1}...\\ &...\left(1-\sum_{j=0}^{k_J-1}e^{-\beta_J\frac{\xi_i}{\xi_i}}\frac{\left(\beta_J\frac{\xi_i}{\widehat{\xi_i}}\right)^j}{j!}\right)^{n_J}\frac{\left(\frac{\beta_i}{\widehat{\xi_i}}\right)^k\xi_i^{k-1}}{(k-1)!}e^{-\beta_i\frac{\xi_i}{\xi_i}}d\xi_i\end{aligned} \tag{4.20}$$

Under the PF scheme, the long term average throughput of a flow i under assumption of $\xi_i/\widehat{\xi_i} \sim$ Erlang(β_i, k_i) is then

$$R_i(\underline{n}) = \int_0^\infty r_{SH}(W,\xi_i)\left(1-\sum_{j=0}^{k_1-1}e^{-\beta_1\frac{\xi_i}{\xi_i}}\frac{\left(\beta_1\frac{\xi_i}{\widehat{\xi_i}}\right)^j}{j!}\right)^{n_1}\left(1-\sum_{j=0}^{k_2-1}e^{-\beta_2\frac{\xi_i}{\xi_i}}\frac{\left(\beta_2\frac{\xi_i}{\widehat{\xi_i}}\right)^j}{j!}\right)^{n_2}...$$

$$\ldots \left(1-\sum_{j=0}^{k_i-1} e^{-\beta_i \frac{\xi_i}{\gtrless \xi_i}} \frac{\left(\beta_i \frac{\xi_i}{\gtrless \xi_i}\right)^j}{j!}\right)^{n_i-1} \ldots \left(1-\sum_{j=0}^{k_J-1} e^{-\beta_J \frac{\xi_i}{\gtrless \xi_i}} \frac{\left(\beta_J \frac{\xi_i}{\gtrless \xi_i}\right)^j}{j!}\right)^{n_J} \frac{\left(\frac{\beta_i}{\gtrless \xi_i}\right)^k \xi_i^{k-1}}{(k-1)!} e^{-\beta_i \frac{\xi_i}{\gtrless \xi_i}} d\xi_i \quad (4.21)$$

The long-term average throughput experienced by a flow of type i is thus dependent on the total state $(\underline{n})$ of users present in the system.

Maximum SINR access selection

As introduced above, the basic principle of Maximum SINR access selection is to schedule a flow based on its instantaneous SINR ξ_i. We know that $\xi_i/\widehat{\xi_i} \sim \text{Erlang}(\beta_i, k_i)$, so $\xi_i \sim \text{Erlang}(\beta_i/\widehat{\xi_i}, k_i)$. Assuming independent probability density functions for $\xi_i/\widehat{\xi_i}$ for all types of users, the probability that a given flow of type i is served by the RA is equal to

$$Pr\left(flow\ i\ is\ served|\xi_i\right) = Pr\left(\xi_i \geq \xi_j, \forall j \neq i\right)$$

$$Pr\left(flow\ i\ is\ served|\xi_i\right) = \left(1-\sum_{j=0}^{k_1-1} e^{-\beta_1 \frac{\xi_i}{\gtrless \xi_1}} \frac{\left(\beta_1 \frac{\xi_i}{\gtrless \xi_1}\right)^j}{j!}\right)^{n_1} \left(1-\sum_{j=0}^{k_2-1} e^{-\beta_2 \frac{\xi_i}{\gtrless \xi_2}} \frac{\left(\beta_2 \frac{\xi_i}{\gtrless \xi_2}\right)^j}{j!}\right)^{n_2} \ldots$$
$$\ldots \left(1-\sum_{j=0}^{k_i-1} e^{-\beta_i \frac{\xi_i}{\gtrless \xi_i}} \frac{\left(\beta_i \frac{\xi_i}{\gtrless \xi_i}\right)^j}{j!}\right)^{n_i-1} \ldots \left(1-\sum_{j=0}^{k_J-1} e^{-\beta_J \frac{\xi_i}{\gtrless \xi_J}} \frac{\left(\beta_J \frac{\xi_i}{\gtrless \xi_J}\right)^j}{j!}\right)^{n_J} \quad (4.22)$$

which is noted to be increasing in ξ_i, though not independent on the SINRs of the competing calls. Deconditioning Expression (4.22) with respect to the experienced SINR ξ_i we find

$$Pr\left\{flow\ i\ is\ served\right\} = \int_0^\infty \left(1-\sum_{j=0}^{k_1-1} e^{-\beta_1 \frac{\xi_i}{\gtrless \xi_1}} \frac{\left(\beta_1 \frac{\xi_i}{\gtrless \xi_1}\right)^j}{j!}\right)^{n_1} \ldots$$
$$\ldots \left(1-\sum_{j=0}^{k_2-1} e^{-\beta_2 \frac{\xi_i}{\gtrless \xi_2}} \frac{\left(\beta_2 \frac{\xi_i}{\gtrless \xi_2}\right)^j}{j!}\right)^{n_2} \ldots \left(1-\sum_{j=0}^{k_i-1} e^{-\beta_i \frac{\xi_i}{\gtrless \xi_i}} \frac{\left(\beta_i \frac{\xi_i}{\gtrless \xi_i}\right)^j}{j!}\right)^{n_i-1} \ldots$$
$$\ldots \left(1-\sum_{j=0}^{k_J-1} e^{\beta_J - \frac{\xi_i}{\gtrless \xi_J}} \frac{\left(\beta_J \frac{\xi_i}{\gtrless \xi_J}\right)^j}{j!}\right)^{n_J} \frac{\left(\frac{\beta_i}{\gtrless \xi_i}\right)^{k_i} \xi_i^{k_i-1}}{(k_i-1)!} e^{-\beta_i \frac{\xi_i}{\gtrless \xi_i}} d\xi_i \quad (4.23)$$

Under the Maximum SINR scheme, the long term average throughput of flow i under assumption of $\xi_i/\widehat{\xi_i} \sim \text{Erlang}(\beta_i, k_i)$ is then

$$R_i(\underline{n}) = \int_0^\infty r_{SH}(W, \xi_i) \left(1-\sum_{j=0}^{k_1-1} e^{-\beta_1 \frac{\xi_i}{\gtrless \xi_1}} \frac{\left(\beta_1 \frac{\xi_i}{\gtrless \xi_1}\right)^j}{j!}\right)^{n_1} \ldots$$
$$\ldots \left(1-\sum_{j=0}^{k_2-1} e^{-\beta_2 \frac{\xi_i}{\gtrless \xi_2}} \frac{\left(\beta_2 \frac{\xi_i}{\gtrless \xi_2}\right)^j}{j!}\right)^{n_2} \ldots \left(1-\sum_{j=0}^{k_i-1} e^{-\beta_i \frac{\xi_i}{\gtrless \xi_i}} \frac{\left(\beta_i \frac{\xi_i}{\gtrless \xi_i}\right)^j}{j!}\right)^{n_i-1} \ldots$$
$$\ldots \left(1-\sum_{j=0}^{k_J-1} e^{\beta_J - \frac{\xi_i}{\gtrless \xi_J}} \frac{\left(\beta_J \frac{\xi_i}{\gtrless \xi_J}\right)^j}{j!}\right)^{n_J} \frac{\left(\frac{\beta_i}{\gtrless \xi_i}\right)^{k_i} \xi_i^{k_i-1}}{(k_i-1)!} e^{-\beta_i \frac{\xi_i}{\gtrless \xi_i}} d\xi_i \quad (4.24)$$

The long-term average throughput experienced by a flow of type i is thus dependent on the total state ($\underline{n}$) of users present in the system.

4.4 Markov Chain Modelling

Now that the packet level analysis is completed we can introduce flow level dynamics. This is done in the second step of the analysis with the creation of a continuous-time Markov chain model describing the dynamics of call initiations and completions in the cell. The states in the Markov model are given by $\underline{n} = (n_1, n_2, .., n_L)$, i.e. the distribution of the calls over the different zones in the cell. Hence the Markov model itself has L dimensions, each of the L dimensions is unlimited in the number of admissible calls. The transition rates of the Markov model are as follows, cf. Figure 4.1

$$\begin{array}{llll} \underline{n} \rightarrow (n_1, .., n_i + 1, .., n_L) & \text{at rate} & \lambda_i & \text{call arrival} \\ \underline{n} \rightarrow (n_1, .., n_i - 1, .., n_L) & \text{at rate} & \frac{n_i}{F} R_i(\underline{n}) & \text{call completion} \end{array}$$

Due to the complexity of the resulting Markov model (transition rates are dependent on the full state) an analytical solution is not feasible. When no closed-form expressions are available, standard techniques for deriving the steady-state distribution can be used, e.g. numerical solution of the balance equation or simulation of the Markov chain. Using numerical simulation we can easily derive the desired performance measures from the steady-state distribution of the Markov model. We will describe the numerical results of this mathematical model and the conclusions that we can draw from these results in Chapter 6.

4.5 QoS Measures

In order to evaluate the numerical results of the Markov Jump queue we defined in the previous section, we have to define the measures on which we determine the QoS. We will distinguish between measures taken from all users that receive service in the network and users which we will denote 'cell edge' users. We define a 'cell edge' user as a with the worst average channel quality, thus a user of type L. With this measure 10% of the users is labeled as 'cell edge' users, since we consider $L = 10$ types of users. We define the following QoS measures for the data users.

1. Average throughput for all users.
2. Throughput for users of type L, the 'cell edge' users.

Operators typically optimise a network by setting minimal requirements for the throughput for 'cell edge' users. We measure throughput for each user as the file size divided by the time needed to sent this file, expressed in kbits/sec.

Chapter 5

Simulation study

In this chapter we will describe the simulation model we used to investigate the influence of packet scheduling decisions on the QoS for users in the network. The network model, traffic model, service rate model and propagation model are described. They are combined with selected scheduling algorithms as described in Chapter 3. The resulting simulation model is implemented in Delphi programming code in order to perform numerical studies. With these numerical studies we want to determine the influence of the different environment characteristics identified in Chapter 3 and the different packet scheduling algorithms on the QoS for the users of the network. The numerical results for the simulation study are presented in Chapter 6 and show the QoS for users by simulation of a network loaded with data traffic only and a network loaded with voice and data traffic.

5.1 Introduction

The goal of this chapter is to compare the performance impact of different parameter settings for the selected scheduling algorithm as described in Chapter 3 for changes in the network, traffic and channel conditions using a simulation model. Different network, traffic and channel conditions and different parameter settings for selected scheduling algorithm lead to differences in the experienced service quality for users of the network. The simulator of the wireless network that we define in this research comprises specific models for the network, traffic, service rate and propagation as described in Sections 5.2, 5.3, 5.4 and 5.5, respectively. As a high level overview of the applied simulator, the call and packet level operations are described below.

At the call level we initiate and terminate calls. Upon call generation the geographic location of the call, as well as its service requirement is determined. We determine when the information, that needs to be transmitted from the base station to the user, arrives and how large these information packets are. These arriving packets are placed in a buffer for each user. Furthermore, upon call generation the propagation characteristics of the call are modelled and the serving base station is selected. Calls that have received their requested service are terminated. Upon termination we determine the QoS that this user has received in accordance with the QoS measures as described in Section 5.6.

At the packet level we model the network in so-called transmission time intervals of 2 ms. In each time interval the path gains, as described in Section 5.5, are updated. Based on these updated path gains the instant achievable service rate, $r_i(t)$, is calculated. We use the packet scheduler, as described in Section 3.4, to make the scheduling decision for each base station. Users that are scheduled receive service according to their instant bit rate, $r_i(t)$, until their buffer is emptied. We detect whether transmission of packets that empty the buffer lead to call termination, which is then processed at the call level.

5.2 Network model

The network topology will be characterized by 12 x 3 cells in hexagonal layout as is shown in Figure 5.1a. In this model we define an area consisting of hexagons. Each of the 12 sites has 3 directed antennas, resulting in a total of 36 base stations. Users are generated by the traffic model described in Section 5.3. Users are placed at a position in the network and they are allocated to one of these 36 base station cells. This allocation is based on the propagation characteristics of a user, as described in Section 5.5. The area uses wraparound in so-called 'doughnut-style', which means that the network area is wrapped around from top to bottom and left to right. The lines in Figure 5.1a denote the allocated base station of each user. This figure clearly visualises the wraparound principle. Figure 5.1 denotes how each site is divided in 3 base stations, serving different users.

We have to define some parameters and aspects of the network to fully determine the network model. We will set these parameters to match a suburban scenario. These network settings are typically found in suburban areas. Urban areas are characterised by a higher density of base stations, with less inter-site distance, while country side areas are characterised by lower density of base stations, with a greater inter-site distance. In our simulation model, we define the following parameters.

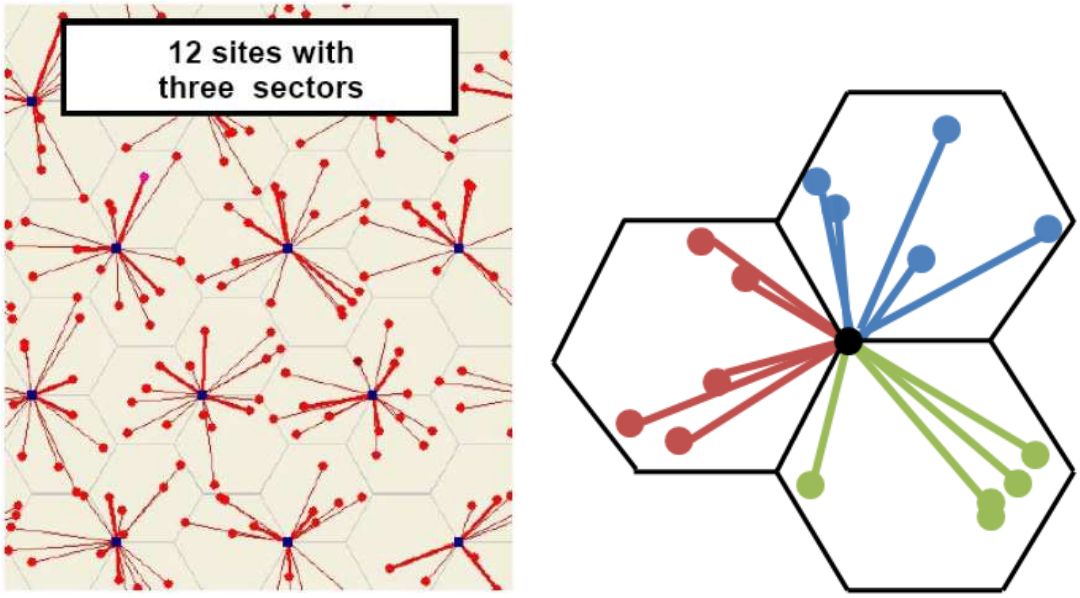
1. Inter-site distance = 1.224 km.
2. Pilot transmitting power of the base station P_{pilot} = 5 Watt. This is the part of the transmitting power a base station is sending to communicate with all possible users. This part of the transmitting power is not available for sending information to users of the base stations and is also being sent when there are no users allocated to the base station. This pilot power includes other signalling channels in the network.
3. Maximum transmitting power of the base station P_{max} = 20 Watt is the maximum power at which a base station can transmit. The difference between the maximum transmitting power and the pilot power can be used to sent information to the users served by the base station.
4. Noise level in the network, N = 9.5499E-14 Watt.
5. Bandwidth of the network, W = 5 MHz.

5.3 Traffic model

Traffic model data users

Users requiring data service arrive to the network characterised by a Poisson process. The inter-arrival times of the users is exponentially distributed with parameter λ_{data}, the aggregate Poisson session arrival rate. By varying λ_{data} we can effectively vary the load on the network. We measure λ_{data} in $calls/sec/cell$.

The general traffic model used in this simulation is denoted in Figure 5.2. Arriving data traffic is characterised by a size which has a hyperexponential probability density function, characterised by a mean and squared coefficient of variation (SCOV). The data packets are relatively large and fill a buffer at the base station with information to be transmitted to the user. We assume that the size of the packets that are transmitted to the user can be of arbitrary size. When the user is scheduled, the base station will sent a packet from the buffer with a size that is in accordance with the channel quality, as we will describe in section 5.4. The user leaves the network when the buffer at the base station becomes empty.



(a) (b)

Figure 5.1: (a) The network layout. (b) The sectorisation of a site in 3 base stations (cells). Each cell in the network is served by a directional antenna.

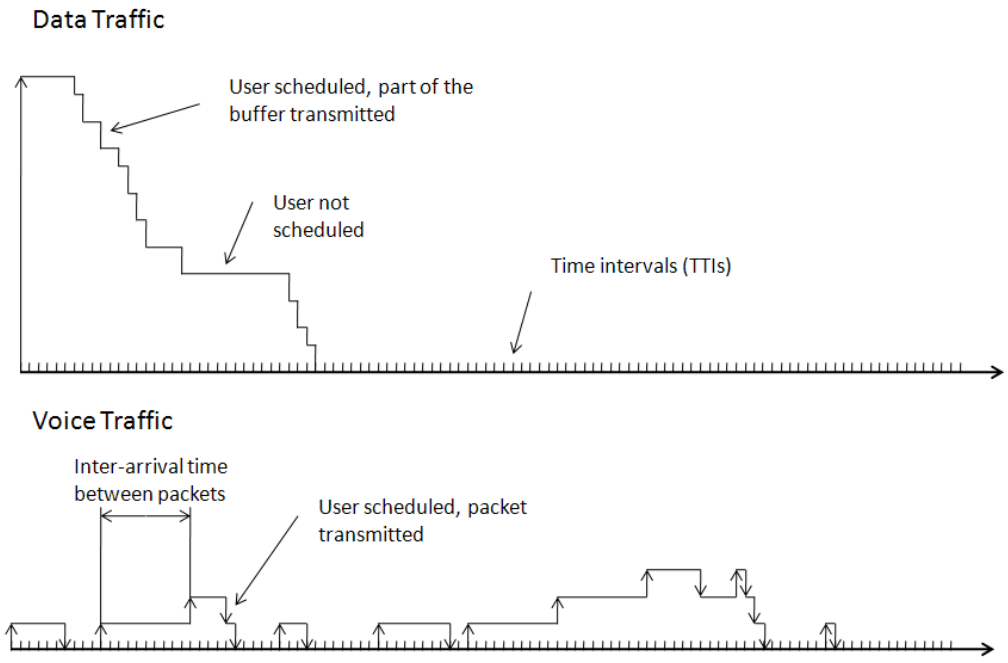


Figure 5.2: The traffic model.

Traffic model voice users

Users requiring voice service arrive to the network characterised by a Poisson process. The inter-arrival times of the users is exponentially distributed with parameter λ_{voice}, the aggregate Poisson session arrival rate. By varying λ_{voice} we can effectively vary the load on the network. We measure λ_{voice} in $calls/sec/cell$.

Voice traffic is characterised by deterministic inter-arrival times of the voice packets. We assume that voice users need a bandwidth of 12.2 kbits/sec and that the inter arrival time between voice packets is 20 ms. This results in deterministic voice packet sizes of 244 bits. The length of a voice call is further characterised by a geometric distribution for the number of packets that arrive in a session with an average of 500 packets. This is equivalent with an average voice call length of 10 seconds. This way we model the voice calls as talk spurts of normal phone conversations. Figure 5.2 shows this voice traffic model. Voice packets are delay-sensitive and need to be sent to the intended user within 100 ms. Packets that violate this deadline are removed from the buffer (not shown in the figure).

5.4 Service rate model

An essential aspect of wireless communications is the degradation of radio signals as they propagate from a transmitter, the base station, to the intended receiver, the user. Since wireless telecommunication is subject to radio propagation, we have to include a model that replicates this phenomenon. In this research we include the propagation aspects of path loss, multipath fading, shadowing and antenna gain. These propagation aspects will determine the loss that the signal of a user is subject to. To see the influence of propagation on the service a user can expect we will show how the rate at which a user can be served (if scheduled) is dependent on its Signal-to-Interference-and-Noise-Ratio (SINR) and how this SINR is based on the different propagation aspects.

A user that is scheduled to receive service by its allocated base station will be served at a rate based on the relation between rate and SINR as described by Shannon [36], with a maximum of 14400 kbits/sec (the maximum service rate in HSDPA). We denote the rate (in kbits/sec) that user i can be served at time t, if scheduled, by $R_i(t)$.

$$R_i(t) = \min\left(14400, W \log_2(1 + \xi_i(t))\right) \tag{5.1}$$

where W denotes the bandwith of the base station in the network and $\xi_i(t)$ is the SINR of user i at time t. Assume user i is allocated to base station k. We define $\xi_i(t)$ in accordance with Equation (3.1) introduced in Chapter 3.

In the remainder of this section we will describe the calculation of the transmitting power $P_{tr,j}(t)$. In the next section we will discuss the propagation model and describe how $G_j(t)$ is calculated. It is obvious that the packet sizes that can be achieved this way are a multiplication of the service rate and the length of the time interval.

Calculating the transmitting power $P_{tr,j}(t)$

When calculating the transmit power of the base stations in the network $P_{tr,j}(t)$ we have to distinguish between the simulation of a network used by data users alone and the simulation of the network used by voice and data users.

For the network loaded with data traffic, we assume that a data user uses all available transmitting power at a base station when scheduled. Typically data traffic is sufficiently large to allow this. We assume that multiplexing is only possible in time, i.e. there is a limitation of up to one

user in the same time interval. Therefore we set $P_{tr,j} = P_{max}$ if base station j has users in service and $P_{tr,j} = P_{pilot}$ if base station j has no users in service. When scheduled, the buffer of the user decreases at the rate calculated with Equation (5.1), until the buffer becomes empty and the user ends his service.

For the network loaded with voice and data traffic, we see that voice users have smaller buffers since the packets they sent are smaller. Therefore they request lower rates than the rates that a base station that transmits at maximum power can achieve. As we discussed in Chapter 3 a network that allows parallel transmission can serve up to 4 users in the same time interval. In this situation part of the available transmit power of a base station is assigned to one user. We calculate the needed power of this user to empty his buffer from Equations (5.1) and (3.1) by determining a needed rate to empty the buffer in Equation (5.1) and the corresponding $P_{tr,k}(t) - P_{pilot}$ in Equation (3.1). The resulting power $P_{max} - P_{tr,k}(t) - P_{pilot}$ can now be used to serve other users at this base station in the same TTI, with a limitation of up to 4 users due to technical reasons.

Users are scheduled based on the scheduling algorithm. This algorithm does take the maximum achievable rate $r_i(t)$ into account which is based on transmitting at full power. This rate is calculated with Equation (5.1) with $\xi_i(t)$ as obtained in Equation (3.1) by setting $P_{tr,k}(t) = P_{max} - P_{pilot}$ in the numerator.

5.5 Propagation model

In this section we will describe how the propagation environment is modelled in the simulation model. The propagation environment determines how the gain $G_j(t)$, used in Equation (3.1), is calculated. The gain of the signal that reaches a user can be subdivided in static gain due to pathloss, shadowing and antenna characteristics and in non-static gain due to multipath fading. The static gain static depends solely on the position of a user in the network. As we assume no mobility of the users, the static gain does not change over time in this simulation model. In simulations with mobility this static gain is time dependent. The non-static gain due to multipath fading is time-varying. The total gain for user i relative to base station k at time t can now be calculated by

$$G_{i,k}(t) = G_{Static,i,k} \cdot G_{Non-static,i,k}(t) \tag{5.2}$$

When determining which base station will serve a user, we determine which base station has the best $G_{Static,i,k}$. This way we do not incorporate the stochastically changing non-static gain in this decision.

Calculating the static gain

We will calculate the static gain in dB and then convert this gain in dB to a linear gain. The static gain is composed of the following components.

1. Gain due to path loss. A signal that is sent from an antenna to a receiver is subject to path loss. This loss increases over the distance the signal is sent over. We initiate each user at a location in the network, as we do each base station. We can then simply calculate the distance between the base station and the user. Since se assume no mobility of the user, this distance remains equal throughout the stay of the user in the network. We will base the gain due to path loss on the Okumura-Hata model [21] [32] and will calculate the loss by $Loss = 136.54 + 35.74 \log_{10}(d) - 12.0$ (in dB), where d is the distance of a user to the base station. This is the loss for a suburban environment. Since $Gain(dB) = -Loss(dB)$ this gives us for the gain due to path loss (in dB).

$$G_{OH,i,k} = -\left(136.54 + 35.74 \log_{10}(d_{i,k}) - 12.0\right) \tag{5.3}$$

2. Shadowing is loss that occurs whenever there is an obstruction in the direct path from the user to the receiver. We model the amount of shadowing with

$$G_{Shadowing,i,k} = \sqrt{0.7}u_i + \sqrt{0.3}v_k \tag{5.4}$$

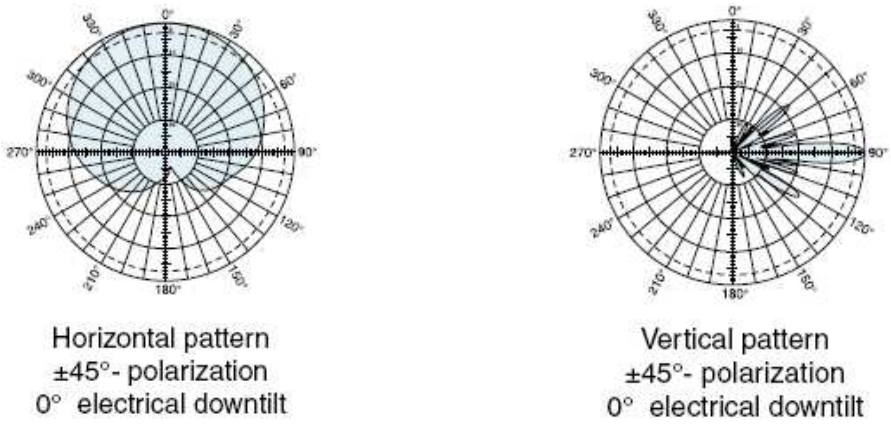
with u_i and $v_k \sim N(0, 9^2)$ [39]. With these parameters the shadowing effects on the radio links from a given user to different sites has a 70% correlation.

3. The antenna gain that the signal is subjected to can be subdivided in three components. The main lobe antenna gain is the gain experienced when sending a signal in line-of-sight. The horizontal and vertical antenna gains are incurred when users are not in line-of-sight of the antenna. Thus both a horizontal and vertical gain are considered here. The values for the mail lobe antenna gain, vertical and horizontal antenna gain are based on measurements from the antenna Kathrein, model 741 989 [3].

   - Antenna Gain Mainlobe is the overall loss from the antenna use.

$$G_{MainLobe} = 11.5dB \tag{5.5}$$

   - The Horizontal Antenna Gain is determined from the horizontal angle between a user and the direction of the antenna of the base station. Since we know the place of the user and the base station in the network, we can easily calculate this angle. For each angle the corresponding gain, $G_{AntennaHorizontal,i,k}$, can be found in Appendix A. We interpolate linearly between angles which are not integer. Figure 5.3a visualises the horizontal antenna gain.
   - The Vertical Antenna Gain is determined from the vertical angle between a user and the direction of the antenna of the base station. Antennas are considered to be directed at an angle of 3 degrees down tilt. We assume that base stations have a hight of 30 meter from the earth and users have a hight of 1.5 meters from the earth. Since we initiated users at a position in the network, we can now calculate the vertical angle between each user and antenna. For each angle the corresponding gain, $G_{AntennaVertical,i,k}$, can be found in Appendix B. We interpolate linearly between angles which are not integer. Figure 5.3b visualises the vertical antenna gain.



(a) (b)

Figure 5.3: (a) The horizontal pattern of the antenna Kathrein, model 741 989. This pattern determines the horizontal gain. (b) The vertical pattern of the antenna Kathrein, model 741 989. This pattern determines the vertical gain.

4. The Indoor Penetration Loss models the loss that the signal experiences when penetrating through walls of buildings. We assume a fixed penetration loss for all users.

$$G_{IndoorPenetrationLoss} = -17.6dB \tag{5.6}$$

We need a minimum static gain to guarantee that there is indeed path loss for each user, as even users located close to the base station experience some path loss. We will set the minimum *Loss* for each user to 30 dB, equivalent to a gain of -30.0 dB. This gives us for the total static gain that user j experiences

$$\begin{aligned} G_{Static,i,k} = min\,(G_{OH,i,k} + G_{Shadowing,i,k} + G_{MainLobe} \\ +G_{AntennaHorizontal,i,k} + G_{AntennaVertical,i,k} + G_{IndoorPenetrationLoss}, -30.0) \end{aligned} \tag{5.7}$$

Now all we have to do is convert this gain in dB into linear gain. This is easily done with

$$G_{Static,i,k}(linear) = 10^{\frac{G_{Static,i,k}(dB)}{10}} \tag{5.8}$$

Calculating the non-static gain

The non-static gain consists of the multipath fading that user i experiences on the power received from base station k we use the Ricean fading model [20][26][33][41]. In this model parameter K is a measure for the degree of multipath fading. $K = 0$ models a high degree of multipath fading, i.e. much interference between waves sent from the same the base station. The Ricean fading model with $K = 0$ is equivalent to the Rayleigh fading model. For $K = \infty$ we model a situation where the gain due to multipath fading is constant over time and equal to 1. This is a situation without multipath fading, also indicated as pure line-of-sight. The intermediate values of parameter K give varying multipath fading from high degree of fading ($K = 0$) to no fading ($K = \infty$). The non-static (linear) gain described by the Ricean fading model is calculated as follows

$$G_{Non-static,i,k}(t) \quad = \quad r_1(t)^2 + r_2(t)^2 \tag{5.9}$$

with

$$\begin{aligned} r_1(t) &= \sqrt{\frac{K}{K+1}} \cos\left(\phi_{k,i,0} + \frac{2\pi vt}{\lambda}\cos\beta_{k,i}\right) + \sqrt{\frac{1}{(K+1)n_w}}\left[\sum_{l=1}^{n_w}\cos\upsilon(t)\right] \\ r_2(t) &= \sqrt{\frac{K}{K+1}} \sin\left(\phi_{k,i,0} + \frac{2\pi vt}{\lambda}\cos\beta_{k,i}\right) + \sqrt{\frac{1}{(K+1)n_w}}\left[\sum_{l=1}^{n_w}\sin\upsilon(t)\right] \\ \upsilon(t) &= \phi_{i,k,l} + \frac{2\pi vt}{\lambda}\cos\zeta_{i,k,l} \end{aligned}$$

Here the terminal velocity is denoted $v(t)$ (in m/s), λ is the wavelength, n_w is the aggregated number of waves, $\phi_{i,k,l}$ is the phase of wave l and $\zeta_{i,k,l}$ is the azimuth angle of wave l (i.e. $\zeta_{i,k,l} = \frac{\zeta_{i,k}(l-1)}{n_w}$), which is not in line-of-sight. $\beta_{i,k}$ is the angle of wave 0, which we assume to be in line-of-sight, i.e. in a direct path from the sender (base station) to the receiver (user). We take $\phi_{i,k,l}$, $\zeta_{i,k}$ and $\beta_{i,k} \sim$ Uniform $(0, 2\pi)$. We see from equation (5.9) that we can determine the amount of influence of the line-of-sight wave (wave 0) compared to the not-line-of-sight wave (wave 1 till n_w) by choosing parameter K.

5.6 QoS measures

In order to evaluate the numerical results of the simulations, we have to define the measures on which we determine the QoS. We will distinguish between measures taken from all users that receive service in the network and users which we will denote 'cell edge' users. We define a 'cell edge' user

as a user that has lower gain, i.e. more loss, than a user at a distance to the base station of 60% of the inter site distance in the direction of the main antenna lobe. This user has strong line-of-sight, resulting in relatively high antenna gain. We base this decision on the place-dependent static loss and do not take gain due to multipath fading into account. The multipath fading component is on average equal for all users, while the static component differentiates 'good' users close to the base station and 'worse' users further from the base station. With this measure approximately 10% of the users is labeled as 'cell-edge' users. We define the following QoS measures for data and voice users.

For data users

1. Average throughput for all users.
2. 10% Quantile of throughput for 'cell edge' users.

Operators typically optimise a network by setting minimal requirements for the 10% quantile of throughput for 'cell edge' users. We measure throughput for each user as the file size divided by the time needed to sent this file, expressed in kbits/sec.

For voice users

1. Average packet loss for all users.
2. 10% Quantile of the packet loss for 'cell edge' users.

Operators typically optimise a network by setting minimal requirements for the 10% quantile of the packet loss for 'cell edge' users. We measure packet loss for each user as the number of packets lost divided by the total number of voice packets that were sent to the user, resulting in a packet loss percentage.

Chapter 6

Numerical results

In this chapter we present the numerical results of our research. First we discuss the results of the mathematical model, which determines the QoS for the three scheduling algorithms, as denoted for the mathematical analysis in Section 3.4, for a single cell that is loaded with data traffic.

Furthermore we present the numerical results for the simulation study for different settings of the parameters of the packet scheduler, as denoted for the simulation study in Section 3.4, for a network loaded with data traffic and a network loaded with voice and data traffic as described in Chapter 5.

6.1 Numerical results of the mathematical model

In this section we discuss the numerical results that we obtained for the mathematical model in a network loaded with data traffic that we defined in Chapter 4. We start with a discussion on the scenarios, the determination of the input parameters in the model and sensitivities we investigate in this mathematical study. We then discuss the numerical results from the Markov Jump queue as described in Section 4.4. Finally we end with some conclusions we can draw from the results.

6.1.1 Scenario

In this section we will determine the influence of changes in the propagation characteristics on the performance for data users for Round Robin, Proportional Fair and maximum SINR scheduling. We investigate whether changes in propagation lead to different optimal schedulers. In order to determine the influence of the environment characteristics, we analyse the performance of the packet scheduling algorithms in three fading environments.

- Scenario 1: A fading environment with a large degree of multipath fading, characterised by Ricean fading with parameter $K = 0$.
- Scenario 2: A fading environment with an average degree of multipath fading, characterised by Ricean fading with parameter $K = 10$.
- Scenario 3: A fading environment with no multipath fading, characterised by Ricean fading with parameter $K = \infty$.

As discussed in Chapter 4 we defined three fittings of the distribution for $\xi_i/\widehat{\xi_i}$ that we will use to determine the long-term average throughput.

- Fitting 1: $\xi_i/\widehat{\xi_i} \sim \text{Exp}(1)$
- Fitting 2: $\xi_i/\widehat{\xi_i} \sim \text{Erlang}(\lambda, k)$

- Fitting 3: $\xi_i/\widehat{\xi_i} \sim \text{Erlang}(\lambda_i, k_i)$

For each of the three scenarios we will determine the performance of the scheduling methods Round Robin, Proportional Fair and maximum SINR for each of the three possible fittings. This performance will be measured at an increasing load offered to the single cell in consideration. In Chapter 4 we have defined a model that consists of L types of users. We determine the performance of the different schedulers based on the average throughput over all of these L types and the average throughput for the user type with the worst average channel quality (the cell-edge users), $\widehat{\xi_i}$, as discussed in Chapter 4. From an operators perspective we determine the optimal scheduler as the scheduler that can handle the highest allowable load, while maintaining a QoS target for the throughput of 500 $kbits/sec$ for the cell-edge users.

We set the average file size at $F = 1000$ $kbit$ and we assume a network with $L = 10$ types of users. Before we can determine numerical results from simulation of the Markov Jump queue, we have to determine the parameters of the fitting distribution of $\xi_i/\widehat{\xi_i}$ for each of the fading environments, as we defined in Chapter 4.

Determining the parameters of the fitting distribution

We used the simulation model as discussed in Chapter 5 to measure realisations of ξ_i and the variance of ξ_i. We used these measurements to fit an Erlang distribution to each of the types of users (Class i) for each of the thee fading environments considered, the results of which are presented in Table 6.1, 6.2 and 6.3 respectively. To obtain these measurements we randomly generated 10000 users, which each were given 10000 multipath fading realisations. The users were then ordered in 10 groups, based on the average channel quality $\widehat{\xi_i}$.

In these tables we see that the average channel qualities, the SINRs ξ_i, are approximately the same for each of the three fading environments. Furthermore we see that the variance of $\xi_i/\widehat{\xi_i}$ is quite different for the three degrees of fading, as we expected. The average variance over the $L = 10$ classes for $K = 0$, $K = 10$ and $K = \infty$ is 0.76, 0.17 and 0.01 respectively.

We can now determine the associated parameters λ_i, k_i that we need for fitting 3 as indicated in Tables 6.1, 6.2 and 6.3. For fitting 2 the associated parameters for $K = 0$, $K = 10$ and $K = \infty$ are$\lambda = k = 1$, $\lambda = k = 6$ and $\lambda = k = 100$ respectively.

6.1.2 Results

The numerical results for the Markov Jump queue analysis of the mathematical model as described in Chapter 4 are shown in Figures 6.1, 6.2 and 6.3. We simulated 30000 jumps in the Markov Jump queue to come to these results. Figures 6.1, 6.2 and 6.3 represent the QoS measures we have indicated: the average throughput and the throughput for the edge users. The graphs show how the throughput for these measures develops for increasing load offered to the the base station. The load is measured by the average amount of $kbits$ offered per second to the base station we consider in this model. As we would expect the average throughput and the throughput for the edge users decreases with increasing load. As we know from classical queueing theory higher loads lead to longer waiting times for the users, resulting in a lower throughput.

From the results presented in Figures 6.1 and 6.2 we see that the optimal scheduler for flow level performance of data traffic does not depend on the fading environment. For the average throughput maximum SINR is the optimal scheduler, for the throughput for the edge users Proportional Fair is optimal. We will discuss the results for each of the fading environments in more detail below.

Class i	$\widehat{\xi_i}$	Variance ξ_i	Variance $\xi_i/\widehat{\xi_i}$	λ_i	k_i
1	3.1536	3.2479	0.3266	3	3
2	2.1489	2.2273	0.4823	2	2
3	1.6258	1.6064	0.6077	2	2
4	1.2653	1.1405	0.7124	2	2
5	1.0131	0.8047	0.7841	1	1
6	0.8117	0.5474	0.8309	1	1
7	0.6497	0.3738	0.8857	1	1
8	0.5185	0.2507	0.9323	1	1
9	0.4100	0.1620	0.9636	1	1
10	0.2850	0.0846	1.0409	1	1

Table 6.1: Associated moments for ξ_i for an environment with a high degree of multipath fading, characterised by Ricean fading with parameter $K = 0$

Class i	$\widehat{\xi_i}$	Variance ξ_i	Variance $\xi_i/\widehat{\xi_i}$	λ_i	k_i
1	3.5224	1.0250	0.0826	12	12
2	2.2822	0.6011	0.1154	9	9
3	1.6423	0.3911	0.1450	7	7
4	1.2280	0.2481	0.1645	6	6
5	0.9610	0.1658	0.1796	6	6
6	0.7643	0.1118	0.1914	5	5
7	0.6087	0.0707	0.1909	5	5
8	0.4834	0.0453	0.1941	5	5
9	0.3817	0.0286	0.1965	5	5
10	0.2670	0.0160	0.2246	5	5

Table 6.2: Associated moments for ξ_i for an environment with an average degree of multipath fading, characterised by Ricean fading with parameter $K = 10$

Class i	$\widehat{\xi_i}$	Variance ξ_i	Variance $\xi_i/\widehat{\xi_i}$	λ_i	k_i
1	3.6346	0.3006	0.0228	44	44
2	2.3205	0.0673	0.0125	80	80
3	1.6397	0.0235	0.0087	115	115
4	1.2136	0.0093	0.0063	140	140
5	0.9409	0.0037	0.0042	241	241
6	0.7437	0.0026	0.0048	212	212
7	0.5934	0.0015	0.0044	229	229
8	0.4719	0.0010	0.0043	230	230
9	0.3730	0.0007	0.0053	188	188
10	0.2621	0.0021	0.0308	32	32

Table 6.3: Associated moments for ξ_i for an environment with no multipath fading, characterised by Ricean fading with parameter $K = \infty$

Large degree of multipath fading, characterised by Ricean fading with $K = 0$

The flow level performance for a fading environment with a large degree of fading is shown in Figure 6.1. As we have discussed above the modelling assumption for the distribution of $\xi_i/\widehat{\xi_i}$ is identical for fittings 1 and 2 in this fading environment, since the Erlang$(1, 1)$ distribution is equal to the Exp(1) distribution. The difference in flow level performance between fitting 2 and fitting 3 is small, which shows that better actual representation of the variation of $\xi_i/\widehat{\xi_i}$ does not result in significant differences on flow level performance. For the average throughput maximum SINR is the optimal scheduler, for the throughput for the edge users Proportional Fair is optimal.

Average degree of multipath fading, characterised by Ricean fading with $K = 10$

For the environment with an average degree of multipath fading changing the from fitting 1, Exp(1), to fitting 2, Erlang$(6, 6)$ did result in differences in the flow level performance for the schedulers considered, as shown in Figure 6.2. The flow level performance for Round Robin scheduling improved. In the figure we see that for fitting 2 a higher maximum load is allowed while still satisfying a given QoS throughput target (both for the average throughput as the throughput for cell edge users). This difference is a direct result of the decrease in variation of $\xi_i/\widehat{\xi_i}$. Since the actual rate at which a user is served is dependent on the Shannon rate, via $R_i = W \log_2(1 + \xi_i)$, we note that the increase in achieved rates due to the decreased chance on low rates is higher than the decrease due to the decreased chance on high rates.

In this environment we see an increase in average throughput and throughput for the cell edge users for the Proportional Fair scheduler at low loads on the network and a decrease at high loads. This difference is due to the extend to which channel quality variations can be exploited. At lower loads the reduction in variability leads to higher throughput with the same reasoning as for the Round Robin case. For higher loads the reduction in exploitation of channel quality variations, due to the lower variability, is greater than the increase in achieved rates due to the decreased chance on low rates.

We furthermore see a small increase in average throughput for the maximum SINR scheduler. This difference is explained by the same reasoning as for Round Robin scheduler. The increase in achieved rates due to the decreased chance on low rates is higher than the decrease in achieved rates due to the decreased chance on low rates. The throughput for the cell edge users is considerably lower for the maximum SINR scheduler. For this user type the increase in achieved rates due to the decreased chance on low rates is considerably lower than the decrease in achieved rates due to the decreased chance on high rates.

Although scheduling based on maximum SINR does result in the highest average throughput for the environment with average degree of multipath fading, Proportional Fair scheduling results in the highest allowable load for a given throughput target for the cell-edge users, which is the threshold used by network operators.

No multipath fading, characterised by Ricean fading with $K = \infty$

The numerical results of the simulation of the Markov Jump queue for a situation with no multipath fading are shown in Figure 6.3. Since the calculations to obtain results in a situation with no multipath fading for the Round Robin, Proportional Fair and maximum SINR scheduler in fitting 3 are computationally too demanding, we have not included these in this research. As the parameters in the Erlang distribution become large, Matlab can no longer numerically determine the long term rate $R_i(\underline{n})$ as needed in the Markov Jump queue (in the calculation of $R_i(\underline{n})$ the factorial of k_i needs to be determined for the largest parameter $k_i = 241$, which is numerically not feasible). As we have seen for the scenarios with a large degree of multipath fading ($K = 0$) and an average degree of multipath fading ($K = 10$), the flow-level performance for fitting 2 and fitting 3 do not

differ significantly. Therfore, we conjecture that the differences in the situation with no multipath fading ($K = \infty$) the flow level performance for fitting 2 and fitting 3 do not differ significantly.

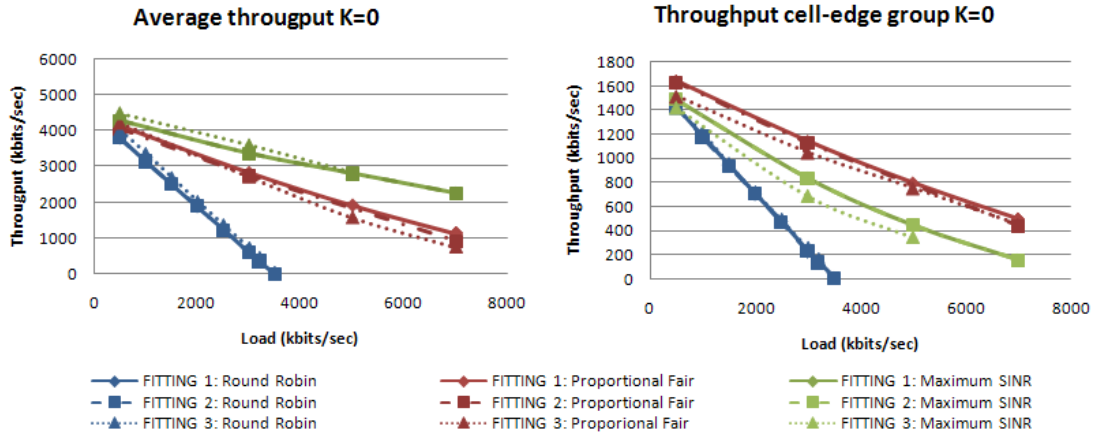


Figure 6.1: Results for the average flow level throughput (left) and for the throughput for cell edge users (right) for a wireless network in an environment with large degree of multipath fading, characterised by Ricean fading with parameter $K = 0$.

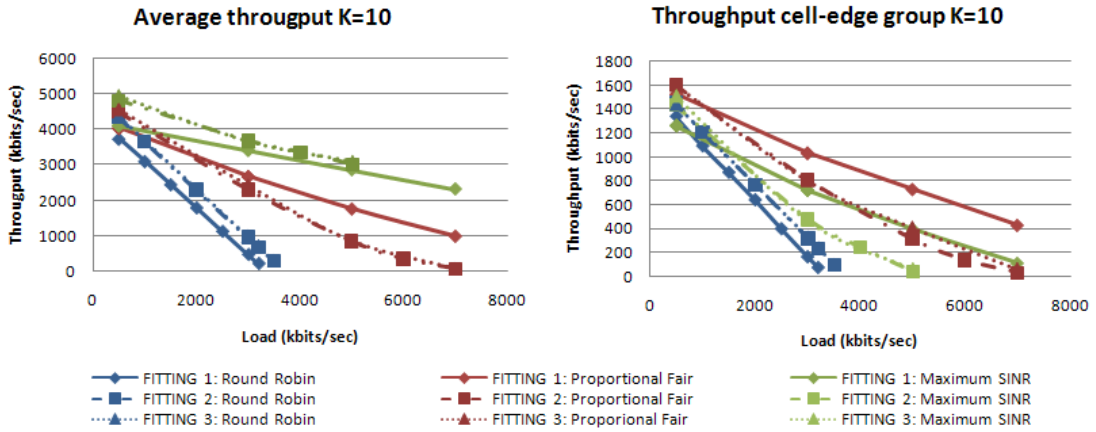


Figure 6.2: Results for the average flow level throughput (left) and for the throughput for cell edge users (right) for a wireless network in an environment with average degree of multipath fading, characterised by Ricean fading with parameter $K = 10$.

6.1.3 Conclusions

In the fading environments with large degree of multipath fading ($K = 0$) and average degree of multipath fading ($K = 10$) for the fittings 1, 2 and 3, as well as in the fading environments with no multipath fading ($K = \infty$) for the fittings 1 and 2, maximum SINR scheduling results in the highest allowable load for given average throughput target. For these scenarios Proportional Fair scheduling resulted in the highest allowable load for the throughput target of 500 kbits/sec for the cell edge users.

Since the optimal scheduling method does not change due to a change in the propagation environment, we see that the choice for the optimal scheduling method is independent with respect

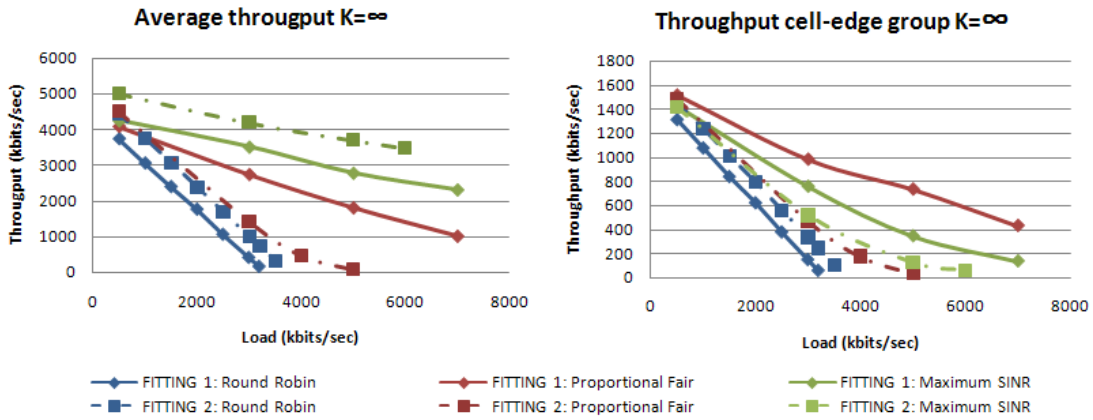


Figure 6.3: Results for the average flow level throughput (left) and for the throughput for cell edge users (right) for a wireless network in an environment with no multipath fading, characterised by Ricean fading with parameter $K = \infty$.

to the propagation environment. For the given QoS measure of a throughput target of 500 kbits/sec for the cell edge users, the scheduling method of Proportional Fair is optimal in the mathematical model.

To determine the sensitivity of optimal parameter settings for actual implementations of the packet scheduler with respect to the fading environment as well as other network, traffic and channel conditions, we will further investigate this relation in the numerical results obtained from the simulation study presented in the next two sections.

6.2 Numerical results of the simulation study: data only

In this section we discuss the numerical results that we obtained for the simulation study in a network loaded with data traffic. We start with a discussion on the simulation scenarios and sensitivities we investigate in this simulation study. We then discuss the numerical results from these simulations and end with some conclusions from these simulations.

6.2.1 Simulation scenario

In this section we will determine the influence of changes in the traffic characteristics and propagation characteristics to the performance for data users for Round Robin, Proportional Fair and maximum SINR scheduling. We investigate whether changes in traffic or propagation lead to different optimal schedulers. In order to determine the influence of the traffic and environment characteristics, we determine a default simulation scenario. We will examine the influence of each of the characteristics by changing this setting from the default scenario and determine the influence that this change has on the QoS measurements. Our default scenario (Scenario D1) is characterised by

- Traffic characteristics
  - The file size distribution is hyperexponential, with balanced means as described by Tijms [38].
  - The mean file size is $F = 1000$ kbits.
  - The squared coefficient of variation of the file size is $SCOV = 1$.
- Propagation characteristics

– The degree of multipath fading propagation is characterised by Ricean fading with $K = 0$. This is equivalent to Rayleigh fading.

To determine the difference in the influence of changing traffic and environment characteristics we will change one of the parameter settings defined above. We will look at the following traffic and environment characteristics settings.

- Scenario D2 - The degree of multipath fading propagation is characterised by Ricean fading with K= 10.
- Scenario D3 - The degree of multipath fading propagation is characterised by Ricean fading with K= ∞.
- Scenario D4 - The squared coefficient of variation of the file size is $SCOV = 16$.
- Scenario D5 - The mean file size is $F = 100$ kbits.
- Scenario D6 - The mean file size is $F = 10000$ kbits.

For each of these six scenarios (the default scenario and the 5 adjusted scenarios) we will determine the QoS measures described for the scheduling methods Round Robin, Proportional Fair with parameter $\alpha = 0.001$, $\alpha = 0.2$ and $\alpha = 0.9$ and Maximum SINR. We will determine the QoS measure dependent on an increasing load for the network, which we will quantify by the average load that is offered to each base station.

6.2.2 Results

The numerical results for the simulation of Scenario D1 till D5 are shown in Figures 6.4 until 6.9. We used a warmup of 5000 arrivals of data users after which we simulated 10000 arrivals of data users to determine the QoS measures of interest. Figures 6.4 until 6.9 represent the QoS measures we have indicated: the average throughput and the 10% quantile of throughput for the edge users. The graphs show how the throughput for these measures develops for increasing load offered to the network. The load is measured by the average amount of kbits offered per second to each base station in the network. As we would expect the average throughput and the 10% quantile of throughput for the edge users decreases with increasing load. As we know from classical queueing theory higher loads lead to longer waiting times for the users, resulting in a lower throughput. The absolute value for the QoS measures, as well as the difference in performance between the schedulers considered is dependent on the considered scenario. From the figures we can distill the influence of the fading environment as well as the file size distribution on the QoS measures.

The influence of the fading environment

From Figure 6.4, 6.5 and 6.6 we see that a lower degree of multipath fading, a higher value for K, results in higher throughput for both QoS measures. On the one hand a larger degree of multipath fading results in more multi-user diversity gain. On the other hand less multipath fading results in less volatile rates, thus in less low rates due to negative multipath fading influence. This last effect is strongest.

The differences between the schedulers are bigger for more heavily loaded cases. For the average throughput we see that maximum SINR scheduling has the highest allowable load on the network at any given throughput target. For the 10% quantile of throughput target of 500 kbits/sec for the edge users the proportional fair scheduler with parameter $\alpha = 0.001$ allows the highest load for scenario D1, D2, D4, D5 and D6. For scenario D3 the proportional fair scheduler with $\alpha = 0.2$ slightly outperforms the scheduler with parameter $\alpha = 0.001$. The difference between the schedulers is largest for scenario D1, with the highest degree of multipath fading. It should be noted that the difference between the schedulers is small.

The influence of the traffic characteristics: file size variability

By comparing the results from scenario D1 and D4, we see that the QoS measures for the different schedulers do not change significantly. For the case of low variability, SCOV of the file size = 1, as well as for the case of high variability, SCOV of the file size = 16, the schedulers for maximum SINR and PF with $\alpha = 0.001$ have similar performance. The average throughput is slightly higher for the maximum SINR scheduler.

The influence of the traffic characteristics: file size average

The differences between scenario D1, D5 and D6 are caused by changes in the file size average. From the numerical results obtained in scenario D1 and D6 we see that an increase on the file size average does not significantly change the throughput for either of the QoS measures. On the other hand we see that a decrease of the file size average, obtained by comparing scenario D1 and D5, does lead to lower throughput for each of the QoS measures. This phenomenon is caused by the so called one-on-one scheduling between users and available TTIs. A user with a small buffer of data that needs to be transmitted will still be assigned all the transmitting power of a base station for one TTI. In case file sizes are larger, the waste of transmitting power is small, since this waste only occurs in the last TTI a user is served. For a smaller file size averages there will be more waste, resulting in a lower system capacity. This leads to lower throughput values on each of the QoS measures. In practice this effect can be mitigated by allowing parallel transmission, as discussed in Chapter 3. For the changes in file size averages we again see that the schedulers for maximum SINR and PF with $\alpha = 0.001$ have similar performance.

6.2.3 Conclusions

The throughput values for each of the two QoS measures considered do not differ significantly. For the considered scenarios D1 till D6 maximum SINR scheduling and proportional fair scheduling with parameter $\alpha = 0.001$ result in the highest allowable load for the throughput target of 500 kbits/sec for the 10% quantile of the throughput for the cell edge users for all scenarios except D3. For this scenario $\alpha = 0.2$ has better performance. The maximum SINR scheduling method results in the highest average throughput values at given load.

The choice for the optimal scheduling method is relatively independent of the propagation environment characteristics considered. Although the scheduler with $\alpha = 0.2$ has slightly better performance in scenario D3 on the QoS target for the cell-edge users, the difference with the scheduler with $\alpha = 0.001$ was so small that it would not be interesting to exploit these differences from an operators perspective. The traffic characteristics, both the file size average as the file size SCOV, have identical optimal scheduling paramters. The scheduler with parameter $\alpha = 0.001$ resulted in the highest allowable load on the network for the given throughput target of 500 kbits/sec for the 10% quantile of the throughput for the cell edge users.

From these observations we can conclude that the optimal parameter for the packet scheduler does not change due to changes in the considered traffic and environment aspects.

6.3 Numerical results of the simulation study: voice & data

In this section we discuss the numerical results that we obtained for the simulation study in a network loaded with voice and data traffic. We start with a discussion on the simulation scenarios and sensitivities we investigate in this simulation study. We then discuss the numerical results from these simulations and end with some conclusions from these simulations.

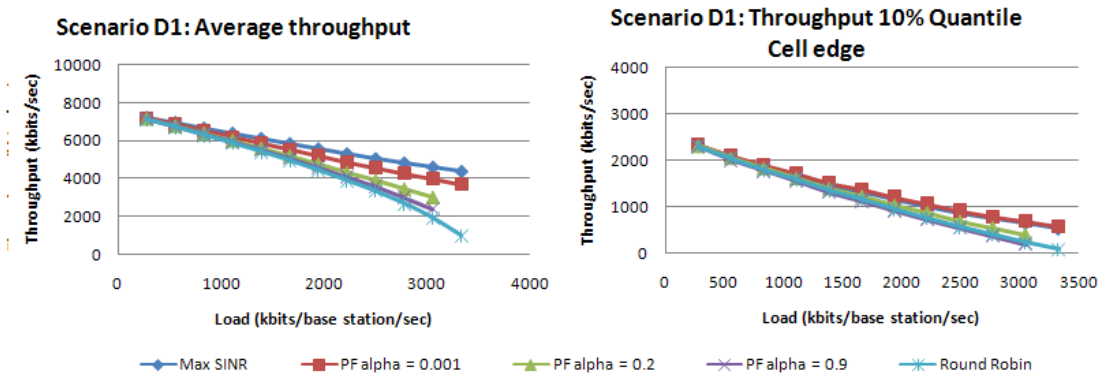


Figure 6.4: Simulation results Scenario D1, the default scenario. In this scenario the mean file size is $F = 1000$ kbits, the squard coefficient of variation of the file size is $SCOV = 1$ and the degree of multipath fading propagation is characterised by Ricean fading with K= 0.

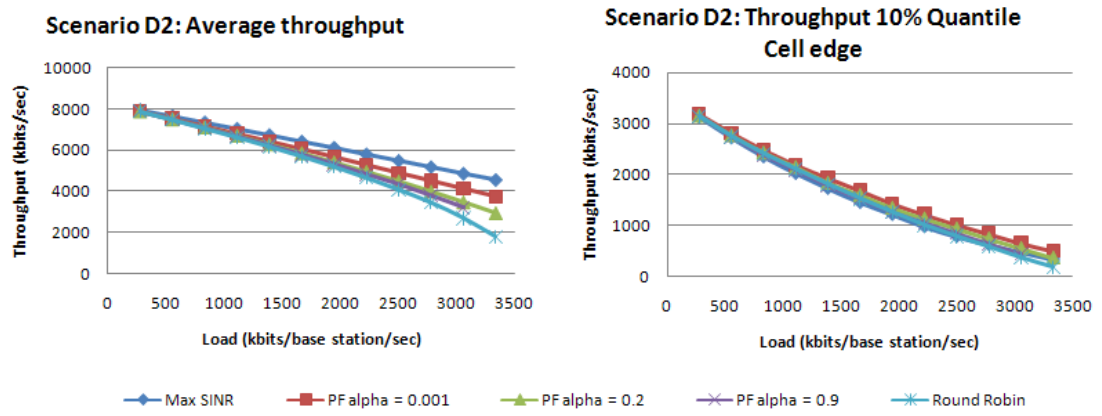


Figure 6.5: Simulation results Scenario D2. In this scenario the degree of multipath fading propagation is characterised by Ricean fading with K= 10.

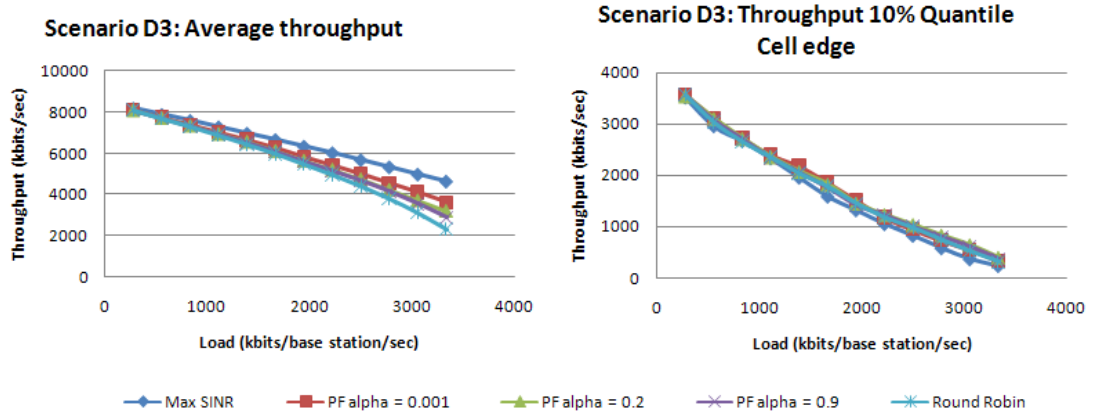


Figure 6.6: Simulation results Scenario D3. In this scenario the degree of multipath fading propagation is characterised by Ricean fading with K= ∞.

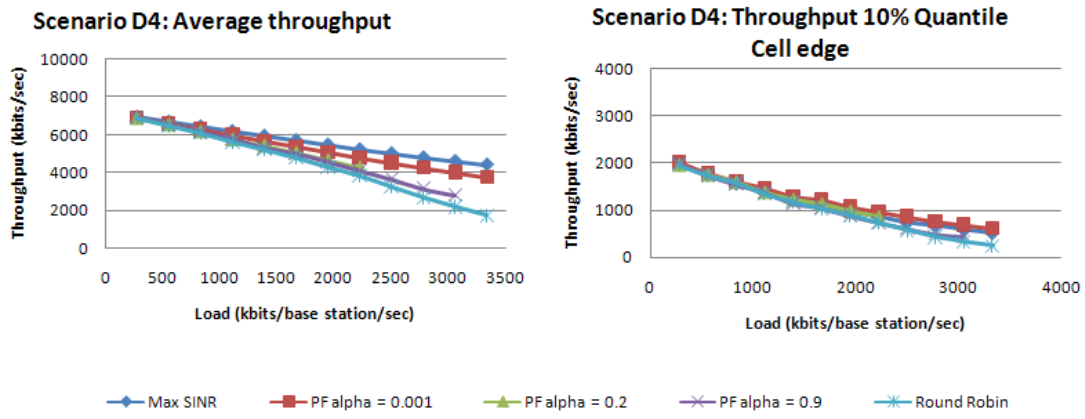


Figure 6.7: Simulation results Scenario D4. In this scenario the squared coefficient of variation of the file size is $SCOV = 16$.

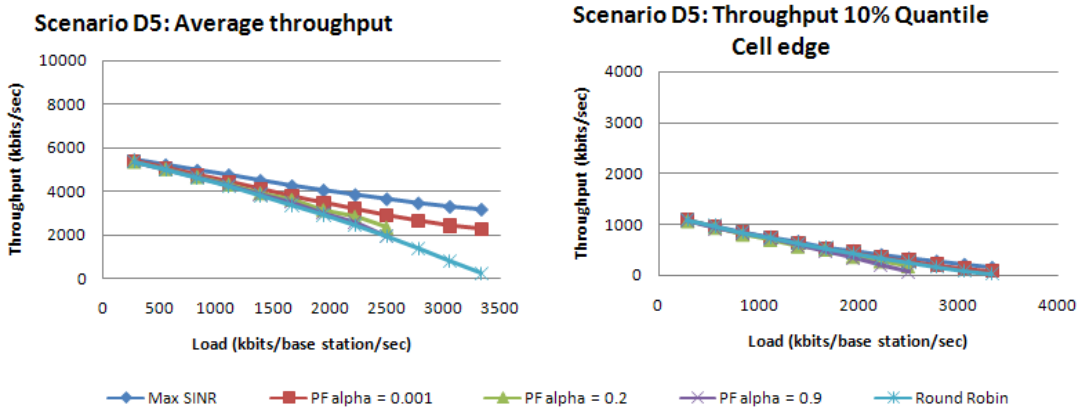


Figure 6.8: Simulation results Scenario D5. In this scenario the mean file size is $F = 100$ kbits.

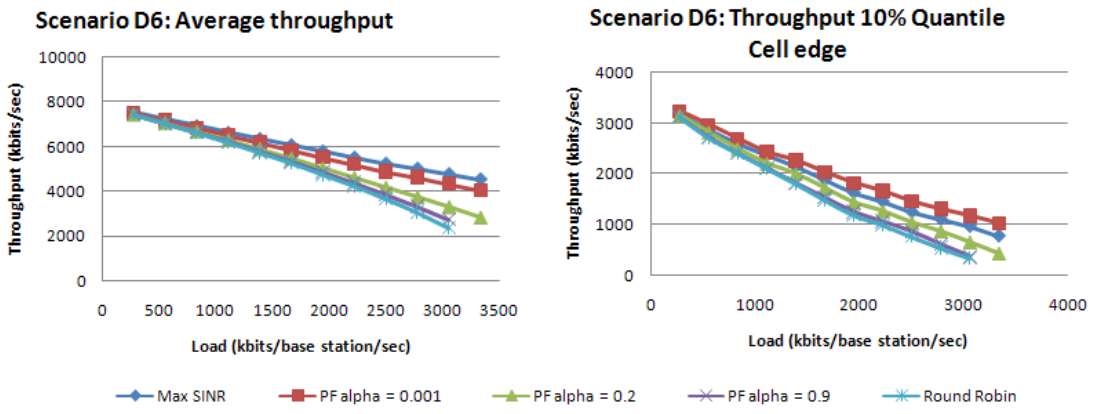


Figure 6.9: Simulation results Scenario D6. In this scenario the mean file size is $F = 10000$ kbits.

6.3.1 Simulation scenario

In order to determine the influence of the traffic mix on the QoS performance for the users, we determine the simulation scenario used in the simulation of a network used by voice and data users. We will examine the influence of the traffic mix and the offered load to the network for.

Our simulation scenario is characterised by

- Traffic characteristics
  - Data traffic
    - The file size distribution is hyperexponential, with balanced means as described by Tijms [38].
    - The mean file size is $F = 1000$ kbits.
    - The squared coefficient of variation of the file size is $SCOV = 4$.
  - Voice traffic
    - The number of voice packets for a voice call has a geometric distribution with a mean of 500. This is equivalent with an average length of a voice call of $L = 10$ seconds.
    - The needed bandwidth for a voice user is $B_{voice} = 12.2$ $kbits/sec$.
- Propagation characteristics
  - The degree of multipath fading propagation is characterised by Ricean fading with $K = 0$. This is equivalent to Rayleigh fading.
- Scheduling parameter
  - The parameter for updating the average rate $\widehat{r}_i$ is $\alpha = 0.001$. We chose this value since $\alpha = 0.001$ resulted in the highest allowable load, as we have seen in Section 6.2
- Service mix of the offered load
  - The amount of load offered to the network is a variable, as well as the fraction of traffic that is voice traffic. Offered load is measured in $kbits/cell/sec$. The offered load for voice and data users is dependent on the stochastic behaviour of the arrival process as described in Section 5.3. Based on the expected file size of for data traffic and the expected length of a call for voice traffic we can now calculate λ_{data} and λ_{voice}, measured in $calls/sec/cell$. Assume a fraction of voice traffic of β and a load of Y $kbits/cell/sec$.

    The number of data arrivals in a cell per second is thus

    $$\lambda_{data} = \frac{(1-\beta)Y}{F} \tag{6.1}$$

    The number of data arrivals in a cell per second is thus

    $$\lambda_{voice} = \frac{\beta Y}{B_{voice} L} \tag{6.2}$$

We identify the scheduling method $\arg\max_i = \frac{r_i(t)}{\widehat{r_i}} \left(1 + \frac{W_i(t)}{T_i - W_i(t)}\right)^{\zeta}$ as described in Chapter 3 with parameter $\zeta = 0$, $\zeta = 0.5$, $\zeta = 1$ and $\zeta = 10$. To determine how the different parameter settings perform, we set minimal QoS targets for both data users and voice users. For data users the 10% quantile of throughput for cell edge users the minimal QoS target is set at 500 kbits/sec. For voice users the 10% quantile of the packet loss for cell edge users the minimal QoS target is set at 2%. For various combinations of offered load and traffic mix it is indicated whether the voice target QoS, the data QoS or both are met. We do not allow blocking of offered load in this simulation.

6.3.2 Results

The numerical results for the simulation of a network used by voice and data users is shown in Figure 6.10. We used a warmup of 5000 arrivals of voice users and 5000 arrivals of data users after which we simulated 10000 arrivals of voice users and 10000 arrivals of voice users to determine the QoS measures of interest. Figure 6.10 shows whether or not the QoS targets for the data users and the voice users, as described in the previous section, were met. As we can see from this figure increasing load results in violation of the QoS targets. The admissible region, the combination of offered load and the fraction of load that is voice user, is largest for the schedulers with parameter $\zeta = 1$ and $\zeta = 10$. Figure 6.10 shows that the scheduler with parameter $\zeta = 1$ satisfies both QoS targets for a fraction of voice users of 0.2 at a load of 3000 kbits/sec and at a fraction of voice users of 0.4 at a load of 2500 kbits/sec, whereas the scheduler with parameter $\zeta = 10$ does violate the data QoS target for these combinations. We should note that the QoS target for voice users for these combinations are only just met for the scheduler with parameter $\zeta = 1$. Furthermore the QoS target for data users is only slightly violated for the scheduler with parameter $\zeta = 10$. From the results we see that the admissible region for given fraction of voice traffic is comparable for the schedulers with parameter $\zeta = 1$ and $\zeta = 10$.

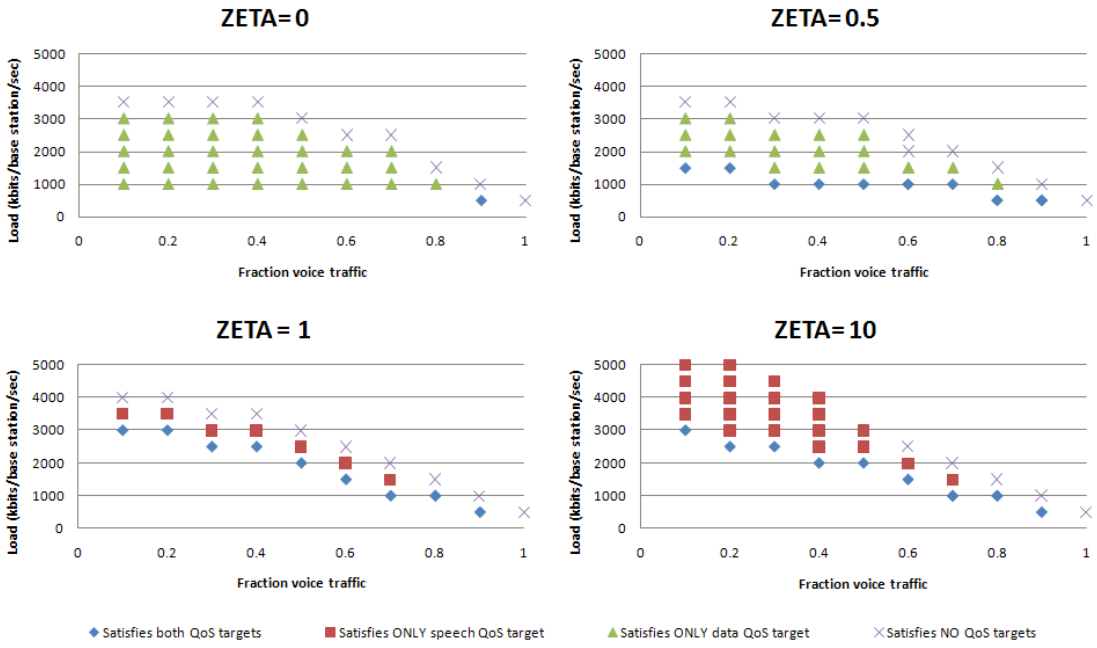


Figure 6.10: Simulation results for a network used by both voice and data users: The combinations of load and fraction of voice users that meet the QoS targets are indicated.

Now take a closer look at the performance of the schedulers for each of the QoS measures. Figure 6.11 shows the performance for the data users for variable load, but fixed fraction of voice users. The average throughput and the 10% quantile of the throughput for the cell edge users are shown. For the case of a fraction of 0.1 voice traffic load and the fraction of 0.3 voice traffic load the performance for the data users is comparable. However we do see that increasing the fraction of voice users leads to more differentiation of the results. The parameter $\zeta = 1$ is performing better than $\zeta = 10$.

On the contrary we see in Figure 6.12 the performance for voice users for variable load, but fixed fraction of voice users. From this figure we see that parameter $\zeta = 10$ is performing better than parameter $\zeta = 1$. Especially the performance for 10% quantile of the packet loss for the cell

edge users is considerable.

6.3.3 Conclusions

We have investigated the performance for the scheduler with parameters $\zeta = 0$, $\zeta = 0.5$, $\zeta = 1$ and $\zeta = 10$ for variations in service mix. For each fraction of voice traffic load, we investigated which total load on the cell is allowed without violating the QoS measures we defined. From Figure 6.10 we see that $\zeta = 1$ and $\zeta = 10$ have similar performance. When we take a closer look at the performance on each of the QoS measure, we see small differences in performance for the QoS measure for data users between the parameter values $\zeta = 1$ and $\zeta = 10$. However, we see that parameter $\zeta = 10$ has significantly lower packet loss for the voice users.

We should note here that voice traffic is typically not bandwidth intensive. As the packets that need to be transmitted are typically small, only part of the transmit power at the base station is needed to transmit this packet. The residual power can be used to schedule other users. Since the network can serve up to 4 users in each TTI, parallel service of voice and data users is possible. This will result in a situation where (relative) priority for the deadline-sensitive voice user will not result in large deterioration in the throughput of the data user.

From these observations we can conclude that the optimal parameter for the packet scheduler does not depend on the offered traffic mix.

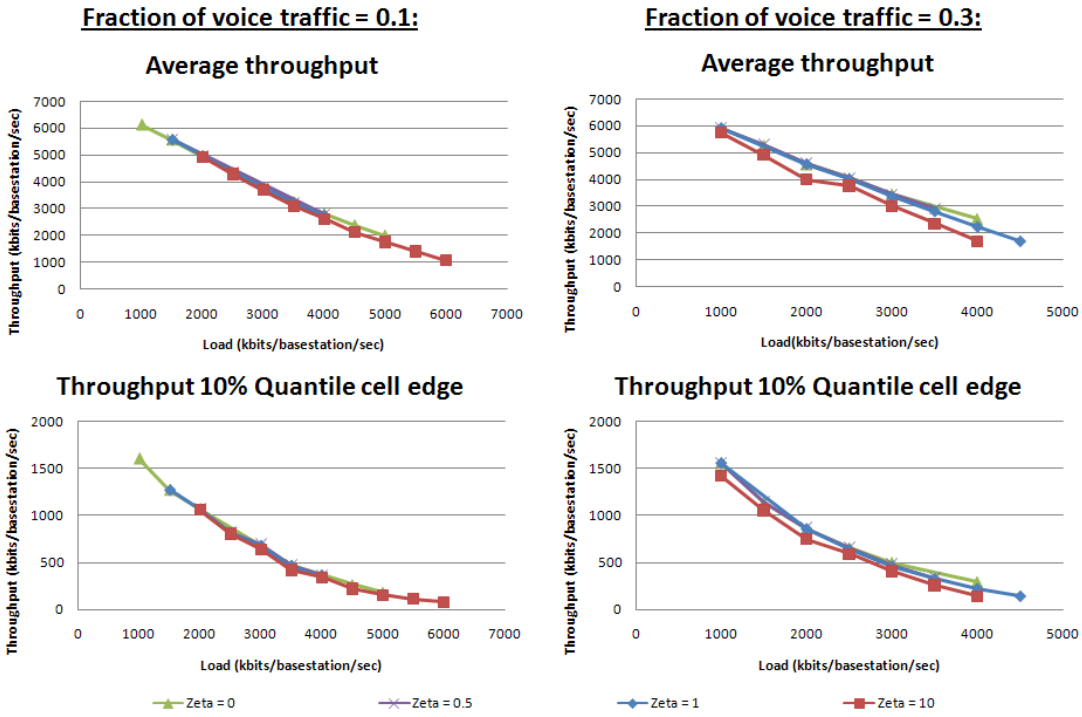


Figure 6.11: Simulation results for a network used by both voice and data users: The data performance for schedulers with parameters $\zeta = 0$, $\zeta = 0.5$, $\zeta = 1$ and $\zeta = 10$ is shown versus the aggregate cell load for a fraction of voice traffic load of 0.1 and 0.3.

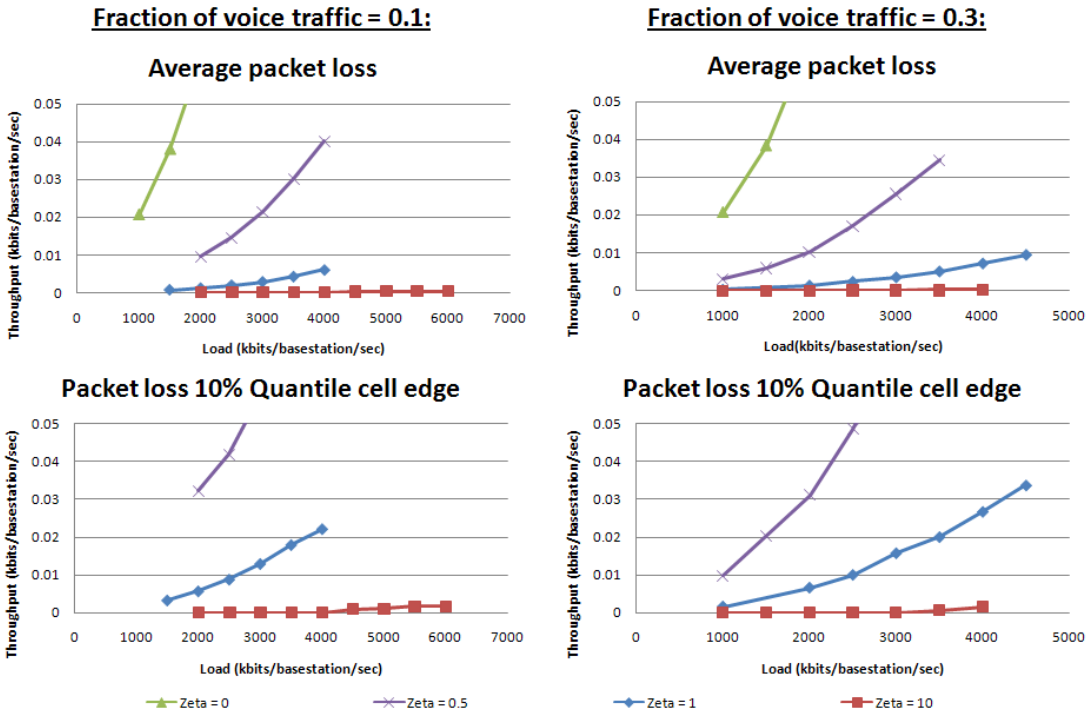


Figure 6.12: Simulation results for a network used by both voice and data users: The voice performance for schedulers with parameters $\zeta = 0$, $\zeta = 0.5$, $\zeta = 1$ and $\zeta = 10$ is shown versus the aggregate cell load for a fraction of voice traffic load of 0.1 and 0.3.

Chapter 7

Conclusions and recommendations

7.1 Conclusions

The research objective is to investigate the sensitivity of the optimal parameter settings for packet scheduling algorithms in wireless networks, loaded with voice and data traffic, with respect to variations in network, traffic and channel conditions. Both the numerical results from the mathematical model and the results from the simulation study show that the optimal parameter settings of the scheduler are largely insensitive to changes in the considered network, traffic and channel conditions. Therefore we conclude that there is no potential for self-optimisation of packet scheduling algorithms based on the investigated network, traffic and channel conditions.

Besides this general conclusion, we describe our findings in more detail. In this research the optimal scheduler in the mathematical model allows for the highest load, while maintaining a minimum throughput. From the numerical results of the mathematical model defined in Chapter 4 we have seen that the optimal scheduler for flow level performance of data traffic does not depend on the fading environment. While changing the fading environment from a situation with a large degree of multipath fading to an average degree of multipath fading, the Proportional Fair algorithm remained the optimal scheduler. We note that changes in the fading environment do result in different behaviour of the channel quality, measured by the signal-to-interference ratio (SINR), as well as different corresponding average throughput and throughput for cell edge users. However, the optimal scheduling principle remains unchanged.

Since the analytical model that we introduced has its limitations, we introduced a simulation model in Chapter 5 with which we were able to incorporate more details and analyse more network characteristics. In this research the optimal scheduler in the simulation model allows for the highest load, while maintaining a minimum throughput for data users or a maximum packet loss for voice users. As presented in Section 6.2, changes in the traffic and environment characteristics considered do not result in different optimal parameter settings for a network loaded with data traffic. For the traffic characteristics, changes in both the mean file size of data users and the variation of the file size did not influence optimal parameter settings. Nor did changes in the degree of multipath fading. For all considered scenarios the Proportional Fair algorithm with the same parameter is shown to be optimal or near optimal. When expanding the network to allow for a combination of voice traffic and data traffic in Section 6.3 we found that the optimal parameter setting does not depend on the offered traffic mix (i.e. the fraction of voice traffic in the offered load) either. For all service mixes considered the influence of the deadline within the scheduling decision should be set at an arbitrary high value, giving voice packets absolute priority over data packets. Since the voice packets that are transmitted are small compared to the size of the data packets, priority for voice packets does not result in significant degradation in the throughput for data traffic. Scheduling based on less influence of the deadline in the scheduling decision does

result in higher packet loss for voice traffic, particularly for relatively high traffic loads.

7.2 Recommendations

Given the results of this research we can recommend the network vendors and operators that they do not have to implement self-optimisation of the parameters of their packet scheduling algorithms based on changes in the network, traffic and channel conditions that we considered. These changes do not result in different optimal parameter settings of the packet scheduling algorithms.

Besides a recommendation to the network vendors and operators, there are recommendations for future research to be made.

The first recommendation is to include mobility as a changing factor in the traffic conditions. In this research we have limited ourselves to static users. Users that have mobility will have a changing geographical place in the network and as a direct result changing average channel qualities. Parameter settings that permit faster adaptation to the changed average channel quality might result in better performance, as the scheduling decision is based on the actual average channel quality instead of the previous average channel quality.

The second recommendation is to change the parameter settings of the packet scheduling algorithms based on on the actual number of users in the network and potentially even the actual channel condition and experienced performance. We would then adapt the parameters based on actual realisation of the statistical distribution, rather than on its average values. In this case optimisation of the scheduling parameter would no longer just be coupled to the average network, traffic and channel conditions, but to the actual number of users in the network that can vary on a smaller timescale, in the order of seconds.

The third and final recommendation is to investigate the potential for self-optimisation of capacity allocation mechanisms other than the packet scheduling algorithm considered in this research (see Chapter 2). Since this research field is relatively new, the potential for each of the capacity allocation mechanisms in mobile wireless networks should be investigated. Once the potential for each of the mechanisms is obtained, researchers can combine these mechanisms to develop actual self-optmisation mechanisms that can be implemented in future wireless networks.

Bibliography

[1] 3GPP TR 32.816 (2007), *Telecommunication management; study on management of Evolved Universal Terrestrial Radio Access Network (E-UTRAN) and Evolved Packet Core (EPC) (Release 8)*, v1.3.1.

[2] NGMN (2006), *Next generation mobile networks beyond HSPA & EVDO*, white paper, www.ngmn.org.

[3] Kathrein antenna type 741989 (2009), http://www.kathrein-scala.com/catalog/741989.pdf.

[4] R. Argawal, A. Bedekar, R.J. La and V. Subramanian (2001), *Class and channel condition based weigthed proportional fair scheduler*, Teletraffic Engineering in the Internet Era, Proceedings ITC-17, pp. 553-565.

[5] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, R. Vijayakumar and P. Whiting (2004), *Scheduling in a Queueing System with Asynchronously Varying Service Rates*, Probability in the Engineering and Informational Sciences, vol. 18, no. 2, pp. 191-217.

[6] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushayana and A. Viterbi (2000), *CDMA/HDR: a bandwidth-efficient high-speed wireless data service for nomadic users*, IEEE Communications Magazine, Vol. 38, no. 7, pp. 70-77.

[7] F. Berggren and R. Jantti (2003), *Multiuser Scheduling over Rayleigh Fading Channels*, Proceedings of Globecom '03 vol. 1, pp. 158-162.

[8] F. Berggren and R. Litjens (2006), *Performance analysis of access selection and transmit diversity in multi-access networks*, Proceedings of Mobicom '06, Los Angeles, USA, pp. 251-261.

[9] J.L. van den Berg, R. Litjens and J.F. Laverman (2004), *HSDPA flow level performance: the impact of key system and traffic aspects*, Proceedings of MSWiM '04, Venice, Italy, pp. 283-292.

[10] J.L. van den Berg, R. Litjens, A. Eisenblatter, M. Amirijoo, O. Linnell, C. Blondia, T. Krner, N. Scully, J. Oszmianski, L.C. Schmelz (2008), *Self-organisation in future mobile communication networks*, Submitted to ICT Mobile Summit, Stockholm, Sweden.

[11] T. Bonald, S. Borst, N. Hegde and A. Proutiere (2004), *Wireless data performance in multi-cell scenarios*, Proceedings of the joint international conference on Measurement and modeling of computer systems '04, pp. 378-388.

[12] T. Bonald and A. Proutiere (2003), *Wireless downlink data channels: user performance and cell dimensioning*, Proceedings of Mobicom '03, pp. 339-352.

[13] T. Bonald and A. Proutiere (2004), *Insensitive Bandwidth Sharing in Data Networks*, Queueing Systems, vol. 44, no. 1, pp 69-100.

[14] T. Bonald, S.C. Borst and A. Proutiere (2004), *How mobility impact the flow-level performance of wireless data systems*, Proceedings of Infocom '04, pp. 1872-1881.

[15] S.C. Borst (2003), *User-level performance of channel-aware scheduling algorithms in wireless data networks*, Proceedings of Infocom '03, pp. 321-331.

[16] S.C. Borst and P.A. Whiting (2001), *Dynamic rate control algorithms for HDR throughput optimization*, Proceedings of Infocom '01, pp. 976-985.

[17] D.C. Dimitrova, J.L. van den Berg, G. Heijenk and R. Litjens (2008), *Flow Level Performance Comparison of Packet Scheduling Schemes for UMTS EUL*, Lecture Notes in Computer Science, pp. 27-40.

[18] J.J. Egli (1957), *Radio propagation above 40 MC/s over irregular terrain*, Proceedings of the IRE, vol. 45, pp. 1383-1391.

[19] K.M.F. Elsayed and A.K.F. Khattab (2006), *Channel-Aware Earliest Deadline Due Fair Scheduling for Wireless Multimedia Networks*, Wireless Personal Communications, vol. 38, no. 2, pp. 233-252.

[20] E. Green (1990), *Radio link design for microcellular systems*, British Telecomub. Tech. J., vol. 8, pp. 85-96.

[21] M. Hata (1980), *Empirical formula for propagation loss in land mobile radio services*, IEEE Trans. Veh. Tech., vol. 29, No. 3, pp. 317-325.

[22] H. Holma and A. Toskala (2006), *HSDPA/HSUPA for UMTS*, John Wiley & Sons, Chichester, United Kingdom.

[23] A. Jalali, R. Padovani and R. Pankaj (2000), *Data throughpute of CDMA-HDR in high efficiency-high data rate personal communications wireless system*, Proceedings of VTC'00, pp. 1854-1858.

[24] A.K.F. Khattab, K.M.F. Elsayed (2004), *Channel-quality dependent earliest deadline due fair scheduling schemes for wireless multimedia networks*, Proceedings of the 7th ACM international symposium on modeling, analysis and simulation of wireless mobile systems '04, pp. 31-38.

[25] H. Lei, M. Yu. A. Zhao, Y. Chang and, D. Yang(2008), *Adaptive connection admission control algorithm for LTE systems*, Proceedings of VTC '08, pp. 2336-2340.

[26] J.M.G. Linnartz (1993), *Narrowband Land-Mobile Radio Networks.* Boston, MA: Artech House.

[27] R. Litjens (2003), *Capacity allocation in wireless communication networks: models and analyses*, Ph.D. thesis, University of Twente, The Netherlands, ISBN 90-9017132-0.

[28] R. Litjens (2005), *HSDPA flow level performance and the impact of terminal mobility*, Proceedings of WCNC '05, pp. 1657-1663.

[29] R. Litjens and F. Berggren (2007), *Assessment of Access Selection and Transmit Diversity in (Non) Cosited Multi-Access Networks*, Proceedings of VTC '07, Baltimore, USA, pp. 320-325.

[30] X. Liu, E.K.P Chong and N.B. Shroff (2003), *A framework for opportunistic scheduling.* Computer Networks, vol. 41, no. 4, pp. 451-474.

[31] M.C. Necker (2006), *A comparison of scheduling mechanisms for service class differentiation in HSDPA networks*, International Journal of Electronics and Communications, vol. 60, no. 2, pp. 136-141.

[32] Y. Okumura, E. Ohmori, T. Kawano and K. Fukuda (1968), *Field strength and its variability in VHF and UHF land-mobile service*, Rev. Elec. Comm. Lab., vol. 16, No. 9-10, pp. 825-873.

[33] R. Prasad and A. Kegel (1993), *Effects of Ricean faded and log-normal shadowed signals on spectrum efficiency in microcellular radio*, IEEE Trans. Veh. Technol., vol. VT-42, pp. 274-281.

[34] F. Delli Priscolli, V. Suraci, A. Di Giorgio, I. Calabrese and G. La Sala (2008), *A reinforcement learning admission control for wireless next generation networks*, ICT Mobile Summit '08.

[35] A. Shakkottai and A.L. Stolyar (2001), *Scheduling Algorithms for a Mixture of Real-Time and Non-Real-Time Data in HDR*, Proceedings of ITC '01, pp. 793-804.

[36] C.E. Shannon (1949), *Communication in the presence of boise*, Proceedings of the Institute of Radio Engineers, vol. 37, no. 1, pp. 10-21.

[37] S.M. Senouci, A.L. Beylot and G. Pujolle (2004), *Call admission control in cellular networks: A reinforcement learning solution*, International Journal of Network Management, Vol. 14, no. 2, pp. 89-103.

[38] H.C. Tijms (2003), *A first course in stochastic models.* Wiley, New York.

[39] A.J. Viterbi (1995), *CDMA Principles of spread spectrum communication* Addison Wesley Publising Company, Reading.

[40] B. Wang, K.I. Pedersen, T.E. Kolding and P.E. Mogensen (2005), *Performance of VoIP on HSDPA*, Proceedings of VTC '05, vol. 4, pp. 2335-2339.

[41] Y.D. Yao and A.V.H. Sheikh (1990), *Outage probability analysis for micro-cell mobile radio systems with co-channel interferences in Ricean/Rayleigh fading environment*, Electron. Lett., vol. 26, pp. 864-866.

Appendix A

Horizontal antenna gains

Angle	Gain	Angle	Gain	Angle	Gain	Angle	Gain	Angle	Gain	Angle	Gain	Angle	Gain	Angle	Gain
0	0	45	-2.8	90	-9.1	135	-14.3	180	-17.1	225	-14.9	270	-9.3	315	-3.1
1	0	46	-2.9	91	-9.2	136	-14.3	181	-17	226	-14.8	271	-9.2	316	-3
2	0	47	-3.1	92	-9.4	137	-14.4	182	-17	227	-14.7	272	-9	317	-2.8
3	0	48	-3.2	93	-9.6	138	-14.5	183	-17	228	-14.5	273	-8.8	318	-2.7
4	0	49	-3.3	94	-9.7	139	-14.6	184	-17	229	-14.5	274	-8.7	319	-2.6
5	0	50	-3.4	95	-9.9	140	-14.7	185	-17	230	-14.3	275	-8.6	320	-2.5
6	0	51	-3.5	96	-10.1	141	-14.8	186	-17	231	-14.3	276	-8.4	321	-2.4
7	-0.1	52	-3.7	97	-10.3	142	-14.9	187	-17	232	-14.1	277	-8.3	322	-2.3
8	-0.1	53	-3.8	98	-10.4	143	-15.1	188	-17	233	-14	278	-8.2	323	-2.2
9	-0.1	54	-4	99	-10.6	144	-15.2	189	-17	234	-13.9	279	-8	324	-2.1
10	-0.1	55	-4	100	-10.8	145	-15.3	190	-17.1	235	-13.8	280	-7.9	325	-1.9
11	-0.2	56	-4.2	101	-10.9	146	-15.5	191	-17.1	236	-13.7	281	-7.8	326	-1.8
12	-0.2	57	-4.3	102	-11.1	147	-15.6	192	-17.2	237	-13.5	282	-7.6	327	-1.7
13	-0.2	58	-4.4	103	-11.3	148	-15.8	193	-17.2	238	-13.4	283	-7.5	328	-1.6
14	-0.2	59	-4.5	104	-11.5	149	-16	194	-17.3	239	-13.4	284	-7.3	329	-1.5
15	-0.3	60	-4.7	105	-11.7	150	-16.2	195	-17.3	240	-13.3	285	-7.2	330	-1.5
16	-0.3	61	-4.9	106	-11.8	151	-16.4	196	-17.4	241	-13.1	286	-7.1	331	-1.4
17	-0.4	62	-4.9	107	-12	152	-16.5	197	-17.5	242	-13	287	-6.9	332	-1.3
18	-0.4	63	-5.1	108	-12.1	153	-16.8	198	-17.6	243	-12.9	288	-6.8	333	-1.3
19	-0.5	64	-5.3	109	-12.3	154	-17	199	-17.7	244	-12.8	289	-6.6	334	-1.2
20	-0.5	65	-5.3	110	-12.4	155	-17.2	200	-17.7	245	-12.7	290	-6.6	335	-1.1
21	-0.6	66	-5.5	111	-12.6	156	-17.4	201	-17.7	246	-12.6	291	-6.4	336	-1
22	-0.7	67	-5.6	112	-12.7	157	-17.6	202	-17.7	247	-12.4	292	-6.3	337	-0.9
23	-0.7	68	-5.8	113	-12.8	158	-17.9	203	-17.7	248	-12.3	293	-6.1	338	-0.9
24	-0.8	69	-6	114	-13	159	-18.1	204	-17.6	249	-12.2	294	-6	339	-0.8
25	-0.9	70	-6	115	-13.1	160	-18.3	205	-17.6	250	-12.1	295	-5.8	340	-0.7
26	-1	71	-6.2	116	-13.2	161	-18.5	206	-17.5	251	-12	296	-5.7	341	-0.6
27	-1	72	-6.3	117	-13.3	162	-18.6	207	-17.4	252	-11.8	297	-5.6	342	-0.6
28	-1.1	73	-6.5	118	-13.4	163	-18.7	208	-17.3	253	-11.7	298	-5.4	343	-0.5
29	-1.2	74	-6.6	119	-13.4	164	-18.7	209	-17.2	254	-11.6	299	-5.3	344	-0.5
30	-1.3	75	-6.8	120	-13.5	165	-18.7	210	-17	255	-11.4	300	-5.1	345	-0.4
31	-1.4	76	-6.9	121	-13.5	166	-18.6	211	-16.9	256	-11.3	301	-5	346	-0.4
32	-1.5	77	-7.1	122	-13.6	167	-18.5	212	-16.7	257	-11.2	302	-4.9	347	-0.3
33	-1.5	78	-7.2	123	-13.6	168	-18.4	213	-16.6	258	-11	303	-4.7	348	-0.3
34	-1.6	79	-7.4	124	-13.7	169	-18.2	214	-16.4	259	-10.9	304	-4.5	349	-0.3
35	-1.7	80	-7.6	125	-13.7	170	-18.1	215	-16.3	260	-10.7	305	-4.4	350	-0.2
36	-1.8	81	-7.7	126	-13.8	171	-18	216	-16.1	261	-10.6	306	-4.3	351	-0.2
37	-1.9	82	-7.8	127	-13.8	172	-17.9	217	-15.9	262	-10.4	307	-4.1	352	-0.1
38	-2	83	-8	128	-13.9	173	-17.7	218	-15.8	263	-10.3	308	-4	353	-0.1
39	-2.1	84	-8.1	129	-13.9	174	-17.6	219	-15.7	264	-10.1	309	-3.9	354	-0.1
40	-2.3	85	-8.3	130	-13.9	175	-17.5	220	-15.5	265	-10	310	-3.8	355	-0.1
41	-2.4	86	-8.4	131	-14	176	-17.4	221	-15.4	266	-9.8	311	-3.6	356	0
42	-2.5	87	-8.6	132	-14.1	177	-17.3	222	-15.3	267	-9.7	312	-3.5	357	0
43	-2.6	88	-8.8	133	-14.1	178	-17.2	223	-15.2	268	-9.6	313	-3.4	358	0
44	-2.7	89	-8.9	134	-14.2	179	-17.1	224	-15	269	-9.4	314	-3.3	359	0

Figure A.1: Horizontal Antenna Gains for Kathrein model 741 989.

Appendix B

Vertical antenna gains

Angle	Gain	Angle	Gain	Angle	Gain	Angle	Gain	Angle	Gain	Angle	Gain	Angle	Gain	Angle	Gain
0	-1.7	45	-13	90	-13.9	135	-13.7	180	-0.1	225	-14	270	-11.7	315	-13.7
1	-0.7	46	-13.3	91	-14.2	136	-13.3	181	-0.7	226	-13.8	271	-11.8	316	-13.8
2	-0.1	47	-13.7	92	-13.9	137	-13	182	-1.7	227	-13.7	272	-11.7	317	-14
3	0	48	-13.8	93	-14	138	-12.8	183	-3.2	228	-13.5	273	-11.7	318	-14.4
4	-0.4	49	-13.6	94	-14.1	139	-12.6	184	-4.9	229	-13.2	274	-11.7	319	-13.8
5	-1.3	50	-13.3	95	-14.2	140	-12.6	185	-6.8	230	-12.9	275	-11.7	320	-13
6	-2.7	51	-13.2	96	-14.1	141	-12.7	186	-8.8	231	-12.6	276	-11.8	321	-12.4
7	-4.5	52	-13.2	97	-14	142	-13.1	187	-10.7	232	-12.4	277	-11.9	322	-11.8
8	-6.4	53	-13.4	98	-13.9	143	-14	188	-12.1	233	-12.3	278	-12	323	-11.4
9	-8.4	54	-13.5	99	-13.9	144	-13.6	189	-12.7	234	-12.2	279	-12	324	-11
10	-10.3	55	-13.5	100	-13.9	145	-12.5	190	-13.2	235	-12.1	280	-12	325	-10.9
11	-11.3	56	-13.3	101	-14	146	-11.7	191	-14.5	236	-12.1	281	-12.1	326	-10.9
12	-11.5	57	-13.1	102	-13.9	147	-11.1	192	-13	237	-12.1	282	-12.1	327	-11
13	-11.8	58	-13	103	-13.6	148	-10.7	193	-12	238	-12.2	283	-12.2	328	-11.3
14	-12.8	59	-13	104	-13.3	149	-10.6	194	-11.4	239	-12.5	284	-12.3	329	-11.8
15	-13.6	60	-12.9	105	-13.2	150	-10.7	195	-11	240	-12.9	285	-12.3	330	-12.5
16	-12.2	61	-12.9	106	-13.1	151	-10.9	196	-11.1	241	-13.4	286	-12.4	331	-13.3
17	-11.3	62	-13	107	-13.2	152	-11.3	197	-11.4	242	-13.8	287	-12.4	332	-14
18	-10.9	63	-13.2	108	-13.4	153	-11.9	198	-12.1	243	-13.6	288	-12.5	333	-13.5
19	-10.9	64	-13.4	109	-13.6	154	-12.7	199	-12.9	244	-13.3	289	-12.6	334	-13.1
20	-11.3	65	-13.7	110	-13.8	155	-13.8	200	-13.8	245	-13.3	290	-12.7	335	-13
21	-12.1	66	-13.8	111	-13.8	156	-13.8	201	-13.8	246	-13.4	291	-13	336	-13.3
22	-13	67	-13.7	112	-13.6	157	-13.3	202	-13.6	247	-13.8	292	-13.3	337	-13.8
23	-13.5	68	-13.6	113	-13.6	158	-13.3	203	-13.8	248	-14	293	-13.8	338	-14.2
24	-13.3	69	-13.6	114	-13.6	159	-13.5	204	-14.2	249	-13.8	294	-14	339	-13.8
25	-13.3	70	-13.6	115	-13.7	160	-13	205	-13.8	250	-13.3	295	-13.8	340	-13.6
26	-13.8	71	-13.8	116	-13.8	161	-12.1	206	-13.3	251	-13	296	-13.4	341	-13.8
27	-13.8	72	-13.8	117	-13.7	162	-11.3	207	-13	252	-12.7	297	-13.3	342	-13.8
28	-12.7	73	-13.6	118	-13.4	163	-10.9	208	-13.1	253	-12.6	298	-13.3	343	-12.9
29	-11.9	74	-13.4	119	-13.2	164	-10.9	209	-13.5	254	-12.5	299	-13.6	344	-12.1
30	-11.3	75	-13.2	120	-13	165	-11.3	210	-14	255	-12.4	300	-13.8	345	-11.4
31	-10.9	76	-13.1	121	-12.9	166	-12.2	211	-13.3	256	-12.4	301	-13.4	346	-11.1
32	-10.7	77	-13.2	122	-12.9	167	-13.6	212	-12.5	257	-12.3	302	-12.9	347	-11
33	-10.6	78	-13.3	123	-13	168	-12.8	213	-11.8	258	-12.3	303	-12.5	348	-11.4
34	-10.7	79	-13.6	124	-13	169	-11.8	214	-11.3	259	-12.2	304	-12.2	349	-12
35	-11.1	80	-13.9	125	-13.1	170	-11.5	215	-11	260	-12.1	305	-12.1	350	-13
36	-11.7	81	-14	126	-13.3	171	-11.3	216	-10.9	261	-12.1	306	-12.1	351	-14.5
37	-12.5	82	-13.9	127	-13.5	172	-10.3	217	-10.9	262	-12	307	-12.1	352	-13.2
38	-13.6	83	-13.9	128	-13.5	173	-8.4	218	-11	263	-12	308	-12.2	353	-12.7
39	-14	84	-13.9	129	-13.4	174	-6.4	219	-11.4	264	-12	309	-12.3	354	-12.1
40	-13.1	85	-14	130	-13.2	175	-4.5	220	-11.8	265	-11.9	310	-12.4	355	-10.7
41	-12.7	86	-14.1	131	-13.2	176	-2.7	221	-12.4	266	-11.8	311	-12.6	356	-8.8
42	-12.6	87	-14.2	132	-13.3	177	-1.3	222	-13	267	-11.7	312	-12.9	357	-6.8
43	-12.6	88	-14.1	133	-13.6	178	-0.4	223	-13.8	268	-11.7	313	-13.2	358	-4.9
44	-12.8	89	-14	134	-13.8	179	0	224	-14.4	269	-11.7	314	-13.5	359	-3.2

Figure B.1: Vertical Antenna Gains for Kathrein model 741 989. Assumed is an angle direction of 3 degrees downtilt.