Radio resource management for OFDMA systems under practical considerations.

Sivridis, Leonidas

Sivridis, Leonidas (2011) *Radio resource management for OFDMA systems under practical considerations..* thesis, Swansea University.
http://cronfa.swan.ac.uk/Record/cronfa42377

Swansea University
Prifysgol Abertawe

Radio Resource Management for OFDMA Systems Under Practical Considerations

Leonidas Sivridis

College of Engineering

Submitted to Swansea University in fulfillment of the requirements
for the degree of
DOCTOR OF PHILOSOPHY (Ph.D.)

2011

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Declaration

I, Leonidas Sivridis, hereby declare that this thesis is my own work and effort and that it has not been submitted anywhere for any award. Where other sources of information have been used, they have been acknowledged and cited.

Signed ....

Contents

Abstract

Orthogonal frequency division multiple access (OFDMA) is used on the downlink of broadband wireless access (BWA) networks such as Worldwide Interoperability for Microwave Access (WiMAX) and Long Term Evolution (LTE) as it is able to offer substantial advantages such as combating channel impairments and supporting higher data rates. Also, by dynamically allocating subcarriers to users, frequency domain diversity as well as multiuser diversity can be effectively exploited so that performance can be greatly improved.

The main focus of this thesis is on the development of practical resource allocation schemes for the OFDMA downlink. Imperfect Channel State Information (CSI), the limited capacity of the dedicated link used for CSI feedback, and the presence of a Connection Admission Control (CAC) unit are issues that are considered in this thesis to develop practical schemes.

The design of efficient resource allocation schemes heavily depends on the CSI reported from the users to the transmitter. When the CSI is imperfect, a performance degradation is realized. It is therefore necessary to account for the imperfectness of the CSI when assigning radio resources to users. The first part of this thesis considers resource allocation strategies for OFDMA systems, where the transmitter only knows the statistical knowledge of the CSI (SCSI). The approach used shows that resources can be optimally allocated to achieve a performance that is comparable to that achieved when instantaneous CSI (ICSI) is available. The results presented show that the performance difference between the SCSI and ICSI based resource allocation schemes depends on the number of active users present in the cell, the Quality of Service (QoS) constraint, and the signal-to-noise ratio (SNR) per subcarrier.

In practical systems only SCSI or CSI that is correlated to a certain extent with the true channel state can be used to perform resource allocation. An approach to quantifying the performance degradation for both cases is presented for the case where only a discrete number of modulation and coding levels are available for adaptive modulation and coding (AMC). Using the CSI estimates and the channel statistics, the approach can be used to perform resource allocation for both cases. It is shown that when a CAC unit is considered, CSI that is correlated with its present state leads to significantly higher values of the system throughput even under high user mobility. Motivated by the comparison between the correlated and statistical based resource allocation schemes, a strategy is then proposed which leads to a good tradeoff between overhead consumption and fairness as well as throughput when the presence of a CAC unit is considered.

In OFDMA networks, the design of efficient CAC schemes also relies on the user CSI. The presence of a CAC unit needs to be considered when designing practical resource allocation schemes for BWA networks that support multiple service classes as it can guarantee fairness amongst them. In this thesis, a novel mechanism for CAC is developed which is based on the user channel gains and the cost of each service. This scheme divides the available bandwidth in accordance with a complete partitioning structure which allocates each service class an amount of non-overlapping bandwidth resource.

In summary, the research results presented in this thesis contribute to the development of practical radio resource management schemes for BWA networks.

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Professor Jinho Choi for his generous support and supervision. I learned valuable lessons from his vast experience in this field. Also, I would like to thank Dr. Xinheng Wang for his support and encouragement my last year as well as Dr.Yue Li for his help in my second year of my research experience.

With great thanks, I want to acknowledge the financial support of Swansea University without which this PhD thesis would not have been possible. Also, I would like to thank my parents for their constant support and encouragement in facing difficult times and frustration. Finally, I would like to thank my girlfriend Sylvia for her valuable comments and suggestions which helped improve the quality of this dissertation. Her time and devotion to me along with her support are very precious to me and always will be.

List of Figures

List of Tables

List of Abbreviations

AMC	adaptive modulation and coding
BE	best effort
BER	bit error rate
BS	base station
BWA	broadband wireless access
CAC	connection admission control
CCSI	correlated channel state information
CDMA	code division multiple access
CP	complete paritioning
CS	complete sharing
CSI	channel state information
DSA	dynamic subcarrier assignment
ertPS	extended real – time polling service
EPA	Extended pedestrian A
EVA	Extended vehicular A
ETU	Extended terrestrial urban
FDD	frequency division duplex
FDMA	frequency division multiple access
FFT	fast fourier transform
ICI	inter carrier interference
ICSI	instantaneous channel state information
IEEE	institute of electrical and electronic engineers

IFFT	inverse fast fourier transform
ISI	inter symbol interference
ITU	international telecommunications union
LTE	long term evolution
LOS	line of sight
MAC	medium access control
MCS	modulation and coding scheme
MIMO	multiple input multiple output
NLOS	non line of sight
nrtPS	non real – time polling service
OFDM	orthogonal frequency division multiplex
OFDMA	orthogonal frequency division multiple access
PAPR	peak to average power ratio
QoS	quality of service
QPSK	quatenary phase shift keying
RMS	root mean square
RRM	radio resource management
rtPS	real – time polling service
SCSI	statistical channel state information
SNR	signal to noise ratio
TDD	time division duplex
TDMA	time division multiple access
UGS	unsolicited grant service
WiMAX	wireless interoperability for microwave access

Chapter 1

Introduction

1.1 Evolution of Wireless Communication Networks

To this day, four generations of mobile communication networks have been developed. First generation refers to the first-generation of wireless telephone technology standards. These analog telecommunications standards were introduced in the early 1980s. Examples of 1G include NMT (Nordic Mobile Telephone) which was used in the Nordic countries, AMPS (Advanced Mobile Phone System) a system deployed in North America and Australia and the Nippon Telephone and Telegraph (NTT) system used in Japan [1]. First generation networks were primarily analogue as analogue frequency modulation was used to transmit the voice signal. However, digital signalling was also used for the purposes of registering, billing, call setup, and paging [2].

Second generation cellular networks were commercially launched in the early 1990s. Such networks are completely digital, as opposed to the 1G analogue networks. Both voice and data applications are enabled with 2G networks. GSM(Global System for Mobile Communications) is a popular 2G standard. GSM was standardized in 1989 and is the result of the work of group "Groupe Special Mobile" which was founded in 1982. Another 2G standard that has made an important market impact is Interim Standard 95 (IS-95), the first to be based on the spread spectrum scheme. The spread spectrum technique spreads the data over a wider channel than what is strictly necessary allowing for several users to share the same frequencies [3].

Although 2G networks have greater data transmission capabilities than 1G networks, their capacity is limited. Third generation, which is also known as International Mobile Telecommunications (IMT 2000), is a family of standards which allows for enhanced network capacity. Third generation networks are able to deliver higher data speeds than the previous generations. Application services include mobile Internet access, video calls and mobile TV. To meet the standards set by the IMT-2000 working group, a network is required to provide peak data rates of at least 200 kbps. The Universal Mobile Telecommunications System (UMTS) system is a popular 3G standard. It allows operators to support greater numbers of voice and data customers than those that could be supported with 2G. Also, UMTS allows for higher data rates (up to 2 Mbps in indoor environments).

Since the development of UMTS, new proposals and modifications have been introduced in order to further increase the user data rate and/or the system efficiency of the UMTS system. The most significant is High Speed Packet Access (HSPA), which enables peak data rates up to 14 Mbps and 5.8 Mbps in the downlink and uplink respectively [4].

Today, the scope of cellular systems has changed significantly. This is because demands for high-data-rate wireless access are increasing. In order to meet these requirements, 4G networks will be deployed to deliver even higher data speeds. Fourth generation mobile communication systems are expected to become commercially available between 2012 and 2015. The IMT-Advanced group has defined 4G as a future wireless telecommunications technology allowing data transfer rates of up to 1 Gbps at nomadic circumstances and 100 Mbps in a mobile environment [5]. These higher data rates offered by 4G networks will enable more reliable wireless broadband applications. In order to achieve these higher data rate requirements, the common factor between the emerging 4G standards is the use of the OFDM scheme. This is combined with MIMO which is the use of multiple antennas at both the transmitter and the receiver in order to further increase the achievable data-rates.

1.2 Research Motivation

The multitude of bandwidth intensive applications that will be made available by 4G networks means that the available radio spectrum needs to be efficiently utilized. In most cases, the available spectrum will have to accomodate multiple users simultaneously. Resource allocation and CAC are key functions in 4G networks as they not only allow for efficient utilization of resources but also provide QoS guarantees to users. OFDMA, is a multiple access mechanism based on OFDM that will be used on the downlink of future generation networks such as WiMAX and LTE [6]. It results in flexible resource allocation [6] in the sense that instead of allocating whole resources, such as total bandwidth, to only one user at a time, some portion of it can be allocated to each user. Traditionally, resources have been managed statically, and this means including an extra margin in the link budget to maintain acceptable performance for worst case fading scenarios. This, however, results in a waste of spectrum and power. Adaptive resource management for OFDMA systems is important because not only the transmission power and modulation modes can be adapted for every subcarrier, but also the multiple-access scheme can be controlled through dynamic subcarrier allocation. Adapting the modulation mode and coding scheme for each subcarrier becomes more effective when the frame length of the system under consideration is small (the WiMAX frame length is 10msec and the LTE frame length is 1msec). This allows the system to quickly respond to the varying channel conditions of the wireless link. AMC can effectively improve the BER performance of a radio channel which has suffered from multipath and shadowing [10]. The flexibility of OFDMA resource allocation can be used to compensate for wireless channel impairment. However, the resource utilization enhancement that arises from flexible resource allocation depends on the availability of the CSI at the transmitter side. Given the CSI of subcarriers, a transmission can be scheduled over subcarriers that have good status. This process is known as dynamic subcarrier allocation and results in a higher overall system throughput [6]. In practice, the assumption of accurate CSI at the transmitter is impractical.

This imperfect CSI can be correlated with the current state of the channel. The use of correlated CSI at the transmitter requires frequent CSI measurement reports which result in increased system overhead requirements. These overhead requirements further increase when a large number of active users are present in the cell [11]. Instead of feeding back the instantaneous channel coefficients to the BS, it can be more reasonable that users simply send the mean of the subchannel SNR distribution [12] to the transmitter. Clearly, this process will consume considerably fewer resources for overhead purposes. In this thesis, SCSI and CCSI are considered separately.

In 4G network environments, a multitude of services each having different QoS requirements need to be supported. Thus, the CAC unit becomes necessary as the available network resources need to be enough to provide the required QoS to each service. The CAC unit is capable of distributing the network throughput amongst the supported services. In BWA networks employing OFDMA, the CAC functionality is also dependant on the user CSI. All these considerations are the motivation to propose resource allocation and CAC schemes that take into account the aforementioned factors. Efficient resource allocation and CAC which are strongly tied to the commercial growth of BWA networks need to take all these issues into account.

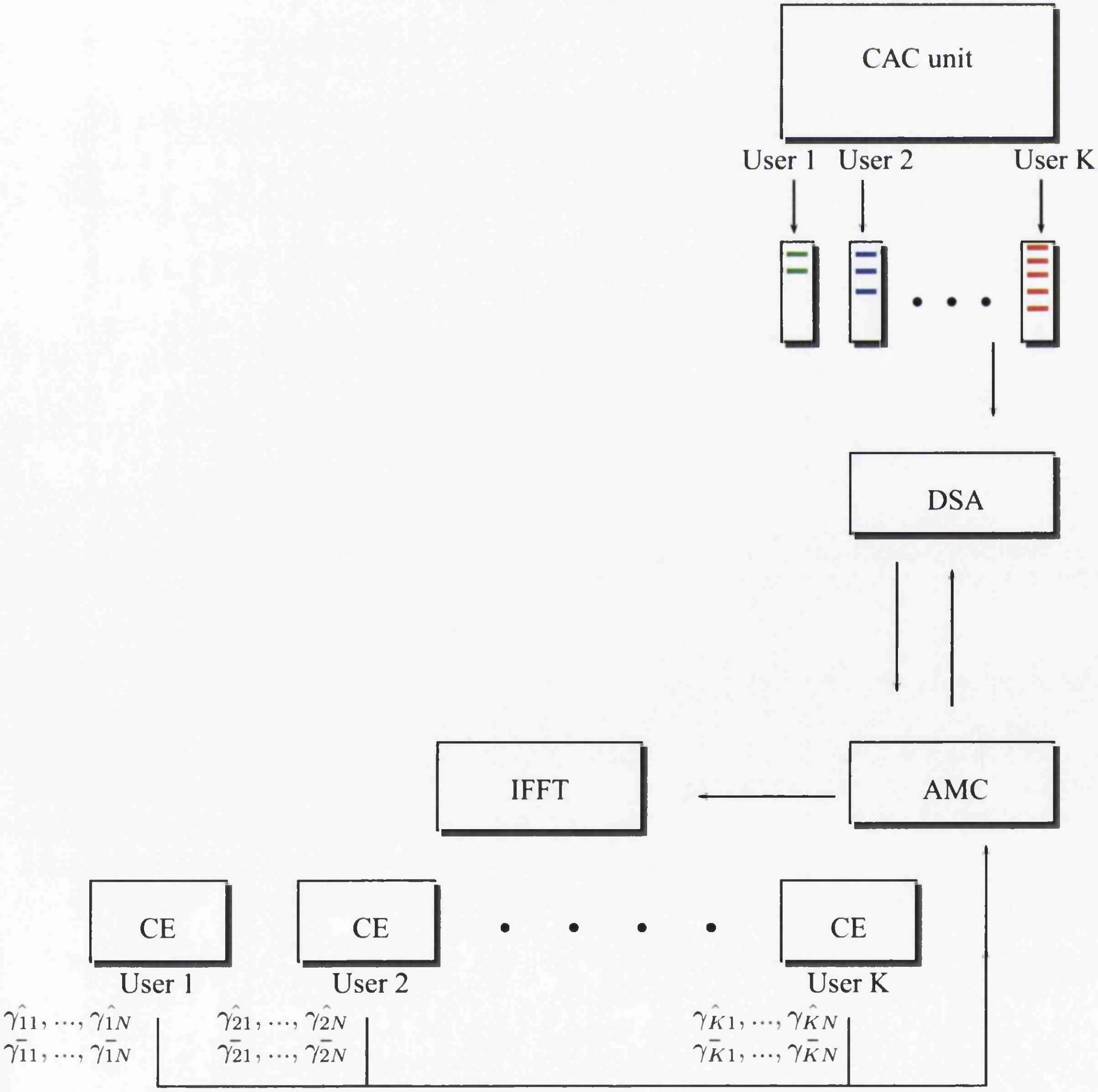
1.3 Problem Description

The wireless radio channel is a challenging medium for reliable high-speed communications. Noise, interference, pathloss and multipath fading impair the wireless channel [7]. One of the most effective techniques to increase spectral efficiency and combat the channel impairments in wireless networks is OFDM. The key concept in OFDM is to split a wideband signal into several orthogonal narrowband signals for transmission. These narrow frequency bands are known as subcarriers. The use of narrowband subcarriers means that OFDM systems are more resilient to frequency selective fading than single carrier systems. The multiple subcarriers used in an OFDM system can overlap without causing interference to each other. This can occur because in OFDM systems the subcarriers are orthogonal, meaning that the peak of one subcarrier coincides with the null of an adjacent subcarrier which will increase the spectral efficiency of the system. Key disadvantages of the OFDM technique as well as proposed solutions to the problems that arise are discussed in[8]- [9]. Given the CSI of subcarriers, adapting the modulation mode and coding scheme for each subcarrier allows the transmitter to send higher transmission rates over the subcarriers with better conditions which improves the throughput and ensures an acceptable bit-error rate at each subcarrier [10]. Despite the use of AMC, deep fading on some subcarriers still leads to low values of the channel capacity. The use of OFDMA further improves spectral efficiency by dynamic subcarrier allocation which allows for frequency and time domain diversity to be effectively exploited. As the probability that all users experience a deep fade on a particular subcarrier is typically quite low, dynamic subcarrier assignment can ensure that subcarriers are assigned to the users who see good channels on them. This allows for the channel to be used in the most efficient way. The larger the number of active users, the strongest channel tends to be stronger and the multiuser diversity gain becomes larger [13]. Another important benefit of dynamic subcarrier allocation is that by scheduling transmission over the subcarriers that have good channel status, less effort for the retransmission of corrupted signals transmitted on the weak subcarriers is required. An optimal frequency usage will be achieved upon optimal subcarrier assignment to users. The intelligent RRM algorithms presented in the literature utilize both dynamic subcarrier allocation and AMC in order to make optimal use of the available wireless resources (i.e maximize system sum-capacity, minimize transmit power)[6].

Optimal allocation depends on the availability of accurate CSI at the transmitter. The CSI, which is also known as CQI, is acquired by the transmitter through feedback following channel estimation by the user terminal. However, a number of different reasons can result in the transmitter working with imperfect CSI. These are user mobility, channel estimation errors and CSI processing delay . Also, in cells with many active users, reporting the CSI of every subcarrier for each user results in a huge communication overhead. An in depth discussion of the factors leading to imperfect CSI can be found in [14]-[17]. Imperfect CSI leads to errors in the performance of AMC and erroneous resource assignments, and, therefore, to a decrease in system performance. Thus the development of practical resource allocation schemes requires accounting for the inaccuracy of the CSI. This thesis is concerned with minimizing the impact of two types of imperfect CSI on AMC: statistical CSI (where users feedback the mean of the subchannel SNR distribution) and correlated CSI (which requires periodic feedback of resources). Not only is the impact of the imperfect channel knowledge on the system performance significant, but the effort required to provide accurate CSI also has to be accounted for. On the one hand, using resources to provide CSI that is correlated with its current state at the transmitter allows for high data rates to be achieved on the downlink. On the other hand, these resources cannot be used to transmit any more data; i.e., signaling overhead is produced which can become prohibitive. Hence, when considering the performance of an adaptive transmission scheme, the signaling overhead also has to be taken into account; i.e., we have a tradeoff between the achievable data rate corresponding to a given CSI accuracy and the signaling overhead required for providing the CSI. This thesis is also concerned with this tradeoff.

Broadband wireless networks support multiple services, which commonly have different data-rate requirements. Therefore, the CAC unit becomes important because it can provide guarantees by distributing the available network resources amongst the supported services. The impact of the CAC unit on resource allocation schemes along with the design of joint CAC/scheduling schemes that can prioritize different services are addressed in this thesis . A diagram depicting the system model of the key areas with which this thesis is concerned with is presented in Figure 1.1. A system with K users and N subcarriers is considered. Packets arrive for each user k that has been admitted into the network by the CAC unit at a rate depending on the application which the user is accessing. The data is stored in the user queues. Each user will estimate their CSI through the use of a channel estimator (denoted by CE in Figure 1.1) and either feedback their average channel state $\bar{\gamma}_{kn}$ or a channel estimate $\hat{\gamma}_{kn}$. At the beginning of each scheduling frame using the CSI and AMC , the BS can compute the achievable data rate of the users for each subcarrier. The scheduler then uses these values along with each user's data-rate requirements of each user to dynamically allocate subcarriers to the users. After modulation and coding the IFFT is applied to create the OFDM signal [7].

Figure 1.1: System Model



1.4 Research Objectives and Contributions

The focus of this thesis is on the downlink transmission channel for OFDMA, since this is typically where the increased performance is needed for mobile broadband wireless access applications [6]. The major objective of this research is to develop practical resource allocation and CAC schemes for OFDMA systems. The developed schemes address practical issues such as inaccuracy of CSI, the support of multiple services with diverse QoS requirements, how to best make use of a limited feedback channel, how to incorporate the user CSI in a CAC policy, and the impact of CAC on resource allocation policies. In this thesis, the developed schemes balance between optimality of resource allocation and computational complexity, while guaranteeing QoS requirements. The research objective is reached over several stages. In the first stage, an approach is developed to solve a rate adaptive resource allocation problem for the case where only the user statistical CSI is known by the transmitter. Based on this developed approach, resources are optimally allocated to achieve a performance that is comparable to what can be achieved when instantaneous CSI is available for the case of Rayleigh fading. The performance difference between the ICSI and SCSI resource allocation schemes depends on the parameters considered and increases with user QoS requirements. This approach is then used to study the performance of OFDMA systems for different vehicular user densities because under significant user mobility it becomes more difficult for users to feedback their CSI. Here, it is also shown that compared to pedestrian users, vehicular users can benefit more from adaptive resource allocation under the assumption that their CSI is known by the transmitter.

In practical systems, only CSI that is correlated to a certain extent with its true state and SCSI can be used to perform resource allocation. In the second stage, resource allocation is performed for the practically relevant case where only a discrete number of modulation and coding levels are available for AMC. First, the CSI inaccuracy effect on the performance of broadband wireless access networks as a function of delay and user velocity was quantified for both correlated and statistical CSI. When resources are allocated in coordination with a CAC unit, the correlated CSI case always yields a significantly higher throughput (even at high user velocities). The throughput gains that are achieved under the correlated CSI based resource allocation scheme come at the cost of increased overhead requirements. The capacity allocated to the dedicated link used for CSI feedback in BWA networks is limited. Therefore, it is useful to limit the amount of overhead bits used for this purpose. A strategy is thus proposed which leads to a good tradeoff between overhead consumption and fairness as well as throughput when the presence of the CAC unit is considered.

CAC is a key RRM function in BWA networks, as it allows for efficient utilization of resources and provides QoS guarantees to users. Typical CAC policies for multiclass wireless networks are the CS and CP schemes [18]. The main concept of CS is that all users share the common resource. If the required resource is satisfied, the connection will be admitted, otherwise it will be rejected. A key drawback of this scheme is that it treats each connection equally without considering their priorities. This leads to users of all traffic classes having equal blocking probabilities. The CP scheme ensures the fairness of different priority calls as the network throughput is divided amongst the services so that users of one service do not overuse the bandwidth.

In OFDMA networks, efficient CAC also depends on the user channel gains. It is therefore useful to develop a CP mechanism that accounts for them. In the third stage, a novel mechanism for CAC is proposed based on the user channel gains and the cost of each service. This scheme divides the available bandwidth in accordance with a CP structure and allocates each accessed service an amount of non-overlapping bandwidth resource.

1.5 Structure of Thesis

In Chapter 2, the relevant literature review is discussed. The problem formulation for resource allocation in OFDMA networks for the SCSI based resource allocation scheme, along with the heuristic algorithm used to solve it, are presented in Chapter 3. First, the approach to optimal loading of the subcarriers when only a statistical description of the CSI is available is introduced. Performance evaluations are presented over two parts: in the first part, the results illustrate the gain achieved when subcarriers are optimally loaded using this approach. In the second part, a comparison is performed with ICSI based resource allocation in order to gain an insight into the factors that affect the performance difference between the ICSI and SCSI based resource allocation schemes. The approach used to optimally load the subcarriers is then employed to evaluate the performance of the OFDMA downlink for different vehicular user densities. In practical systems, it is impossible for the transmitter to constantly work with perfectly known CSI. In fact, the transmitter can only work with either CSI that is correlated to a certain extent with its true state which is obtained through periodic feedback via the user terminal or statistical CSI which requires less frequent feedback. When such CSI is treated as perfect, the system performance is significantly deteriorated due to the non-zero probability of outage events. Chapter 4 first quantifies the performance degradation created by outages for both cases for the practically relevant case where only a discrete number of MCS schemes are available for AMC. Since in a fast-fading environment, periodic CSI measurements require large amounts of feedback, a strategy is proposed that considers the presence of the CAC unit and provides a good tradeoff between overhead consumption and fairness as well as system throughput. CAC is an important RRM function in 4G networks. Chapter 5 explains that CAC in OFDMA networks is also dependant on the user channel gains. A novel mechanism for CAC is proposed based on the user channel gains and the cost of each service. This scheme divides the available bandwidth in accordance with a Complete Partitioning (CP) structure and allocates each accessed service an amount of non-overlapping bandwidth resource. The exact amount of bandwidth allocated to each service depends on its cost. Numerical results demonstrate that the cost model has a strong impact on the blocking probability associated with each traffic class. In Chapter 6 critical evaluations of the new schemes/approaches this thesis has proposed against the related work are provided. The contribution of this thesis is summarized in Chapter 7. In addition, possible future research works relevant to this thesis are also discussed here.

Chapter 2

Literature Review

The idea of using channel information at the transmitter to improve the performance of communication systems has been around since at least 1968 [19]. The main concept is to use knowledge about the channel in order to adjust transmission parameters accordingly to maximize communications performance. This is known as adaptive modulation and coding.

Adaptive modulation and coding in single-user wireless communication systems have been studied extensively [20],[21]. The extension of the adaptive modulation concept to resource allocation in multiuser wireless networks has also been very well studied since the introduction of the concepts of multiuser diversity [22].

Most works on OFDMA resource allocation have considered the case where instantaneous CSI is available at the transmitter and various algorithms based on instantaneous CSI have been developed. The research in this area can be broadly divided into two categories, namely margin-adaptive and rate-adaptive. Margin adaptation is the minimization of the transmit power subject to minimum Quality of Service (QoS) requirements for each user [6]. Examples of such work are [23] and [24]. Rate adaptation is the maximization of the data rates subject to QoS constraints [6]. These QoS constraints could be a combination of data rate, bit error rates, or delays. An example of rate adaptation is presented in [25]. The solution to these problems depends on the availability of ICSI at the transmitter.

When the ICSI is available, the transmitter can optimally assign rates to users; and outages will not occur.

The effect of imperfect CSI has been well considered for single user OFDM systems [26]- [28]. In [26], it is shown that the effect of outdated channel information on the performance of an adaptive OFDM system can be overcome by taking advantage of channel prediction. The effect of OFDM channel estimation errors, as well as that of outdated CSI, were studied in [27], whereas optimal power loading algorithms for data-rate maximization were developed in [28]. In [27], it was shown that the effect of outdated CSI is detrimental; whereas the effect of CSI errors on adaptive OFDM is minimal when a reasonably good channel estimator is used.

For the case of multiuser OFDM, in [29], the authors considered the problem of ergodic weighted sum-rate maximization for resource-allocation, and studied the impact of channel estimation error, where channel estimation error resulted from pilot-aided MMSE channel estimation. However, the effect of outdated CSI has a more detrimental effect and is not considered. Moreover, no assurances are provided for the users' QoS requirements.

In [30], Stefanatos and Demetriou considered outage capacities for resource allocation purposes under the partial CSI assumption where the channel estimation error may arise because of the channel estimation process accuracy and feedback/processing delay. A (close to) optimal resource allocation algorithm was derived that allocates subcarriers, rates, and power in a manner that will maximize the sum-goodput of the system, by exploiting the statistical description of the partial CSI. In order to make the algorithm more practical, closed form expressions for the data-rate that maximizes the goodput are obtained. This avoids the need for a numerical search. However, the closed form expression is only applicable when reliable channel estimates are considered and therefore does not apply to the case where only the SCSI is known by the transmitter. In [31], the same authors used this approach to investigate the problem of partial CSI based resource allocation in downlink OFDMA with minimum user data-rate requests.

Works on resource allocation under the outdated CSI assumption were performed in [32]. In [32], the authors consider the probability distribution function of the current CSI conditioned on outdated CSI in the resource allocation framework and propose an algorithm that minimizes the total transmit power of the system subject to constraints on users' conditional expected capacities. In [33], the authors use a channel prediction model to determine the pdf of the CSI conditioned on the outdated values. They then use this pdf in the evaluation of the Block Error Rate of the users' transmission and exploit the result in the development of a novel scheduling rule. However, as the user CSI (i.e $h|\hat{h}, f_{dT})$ follows a conditional distribution, it is preferable to use outage capacities rather than ergodic capacities which apply to the case where each codeword is long enough to experience a sufficiently long number of different channel states.

Resource allocation based on channel distribution information was investigated in [12]. The authors addressed the problem of subchannel assignment and power allocation, to maximize the ergodic weighted-sum rate under long-term fairness, minimum data rate requirements, and power budget constraints for the mobile WiMAX environment. It was shown that in fading environments, even when only the channel distribution information is known by the transmitter, adaptive resource allocation strategies can provide performance enhancements for OFDMA systems. However, these authors considered maximization of the ergodic weighted sum-rate of the system. In [34], the authors considered rate-adaptive resource allocation problems under the SCSI assumption. Their paper then proposes an adaptive algorithm based on stochastic approximation methods that do not require knowledge of the CDI. However, ergodic rates were maximized in this case too. Ergodic capacities are calculated based on the assumption that the channel fading transitions through all possible fading states; and therefore, this definition might not be very useful in practice for source transmission with fixed delay constraints.[35].

Wireless scheduling algorithms can exploit multiuser diversity by prioritizing the users with the best channel conditions. In order to fully extract multiuser diversity gain, the BS needs to know the CSI of every active user in the network. Therefore, a key disadvantage of this scheme is the amount of overhead required to carry the instantaneous channel states from the users to the BS. In particular, the feedback information in the network linearly increases as the number of active users present in the cell grows. A number of feedback information reduction schemes have been proposed. These can be broadly split into two groups. One is feedback rate reduction related to quantization whilst the other is a threshold-based technique where the threshold is set in a way that allows only users with a large probability of being scheduled to feedback their CSI. The aim of these feedback reduction schemes is to reduce the feedback load whilst maintaining the performance of the scheduling algorithm which can be significantly deteriorated because of insufficient feedback information.

In [11], the threshold based scheme is first introduced, and then the authors showed that the feedback load associated with carrying the instantaneous channel states from all active users to the BS is largely unjustified. The threshold based scheme presented in [11] is referred to as selective multi-user diversity. With this scheme, each user k compares its absolute SNR to a certain threshold value $\gamma_k(t)$ and only those who fall above it are allowed to feedback their downlink transmission rates to the BS. Other users remain silent. This scheme, however, suffers in a system with asymmetric user fading statistics as fairness becomes a significant problem. In order to solve this problem, these authors also consider the use of normalized values. In this case, each user compares the value of its normalized SNR $\frac{\gamma_k(t)}{\bar{\gamma}}$ to a threshold value A, and feeds back its CSI only when its instantaneous channel state exceeds this value. In [36]- [38], the authors use the same threshold value to reduce the feedback information.

A key issue associated with the threshold based technique is that the number of users which simultaneously feedback their instantaneous CSI to the BS may vary. This results in variable-bit rate feedback channels. This is in contrast with the situation in practice where the bandwidth allocated for the feedback of instantaneous channel states is limited

[39]. Consequently, such schemes are difficult to realize in practice. In systems such as 3GPP/LTE and WiMAX/802.16, the BS controls the feedback process by sending signaling messages each time it requires a CQI report from a user [39]. With these proposed schemes, users locally decide whether they should attempt to access the feedback channel or not. In [40], the threshold value is adaptively adjusted in order to limit the number of users that feed back their CSI. However, with this scheme, the number of users that feed back their CSI can still vary between different time slots.

The relationship between fairness and feedback reduction was further discussed in [40]-[42]. In [40], only the users that have a high probability of being selected by a proportionally fair scheduler feed back their CSI. In [41], a modification of the best M CQI feedback scheme [43] is presented. Here, users feed back their best M weighted CQI values rather than their M best absolute CSI values. Similar ideas were proposed in [42]. In these papers, the key idea is to give more feedback opportunity to the users who are scheduled the most. However, the users who are more often to be scheduled rapidly vary between scheduling intervals when fairness oriented algorithms are used. Frequently revoking and reallocating the resources reserved for feedback purposes results in a lot of bandwidth being consumed for signaling [43].

In [44], the authors attempted to further reduce the amount of CQI feedback by using quantised CQI to represent user channel states. Instead of sending real-valued variables, users send a discrete value to represent their true channel state. These quantized values indicate the appropriate modulation levels that are associated with the CSI.

In [45], Sanayei and Nosratinia reduced the feedback rate to one bit. In particular, the BS sets a threshold value α which is the same for all users. Each user compares the absolute value of their channel gain to this threshold. Whenever the channel gain exceeds the threshold, a 1 will be transmitted to the base station; otherwise a 0 will be transmitted. This leads to a fixed rate feedback channel. However, this scheme cannot be used in a practical scenario employing AMC. This is because with this scheme, the instantaneous channel states cannot be fed to the receiver. This will lead to either high outage probabilities or to transmission on MCS levels that are too conservative with respect to the real channel state.

In order to guarantee the QoS in multiservice BWA networks, it is critical to implement the CAC mechanism [46], [47]. CAC is an important step for the provision of QoS guaranteed services because it can prevent the system from being overloaded. Also, CAC can assist a multiservice network to provide different classes of traffic loads with different priorities by manipulating their blocking probabilities [47].

In [47], a framework of 2-D CAC was proposed to accommodate various features of fixed WiMAX networks (802.16). Specifically, the 2-D uplink and downlink WiMAX CAC problem was decomposed into two independent 1-D CAC problems. The authors then formulate the 1-D CAC optimization, in which the demands of service providers and subscribers were jointly accounted for. To solve the optimization problem, a utility and fairness constrained optimal revenue policy as well as its corresponding approximation algorithm were developed. In this work, the presence of the fairness requirement guarantees that the blocking probabilities of all traffic classes will be kept relatively uniform so that no traffic class is unfairly treated. However, in a multiservice BWA network the QoS should be guaranteed for certain services (i.e voice) even when the network is extremely overloaded. Also, as in [47] a fixed WiMAX network is considered the approach cannot be applied to cases where the users are mobile and subject to fading. This is because it does not account for the inherently dynamic nature of mobile wireless networks. Other works on bandwidth allocation and CAC policies for fixed WiMAX can be found in [48]-[50].

In WCDMA and OFDMA based wireless networks, there is no evident relationship between the number of users that can be admitted into the network and the available capacity of the system [51]. A number of CAC schemes have been proposed for WCDMA networks [52]-[54]. In WCDMA networks, radio resources are based on the level of received interference in the uplink and the total transmit power in the downlink [55]. Therefore, WCDMA admission control techniques are not directly applicable to OFDMA based systems like LTE and mobile WiMAX where the available bandwidth is divided into subchannels [55]. For example, a well-known strategy in WCDMA, when congestion is detected, is to lower the bit rates of connected bearers until the load returns to an acceptable level [52]-[53]. However, in pure OFDMA based systems such as LTE and mobile WiMAX, the data-rate is maintained at the MAC scheduler [55]. Thus, if congestion is detected in a cell, the system must remove a subset of the connected bearers until the load is reduced to an acceptable level. Therefore, congestion control is particularly important in OFDMA based systems.

Connection admission control schemes for OFDMA based cellular networks have been investigated in [55]-[58]. In [56], the authors propose a CAC scheme for IEEE 802.16e/OFDMA systems supporting mobile multimedia Internet services. The admission decision is based on the capacity of the cell as well as on the expected traffic load that each connection will bring to the network. The proposed scheme achieves to fulfill the QoS demands of the connections; but, in overloaded situations, only non-real time connections can be admitted.

In [57], a novel admission control scheme is proposed for handling multiclass services in LTE systems. An objective function of maximizing the number of admitted users is proposed to evaluate the system capacity. However, in the presented numerical results, there is no plot that distinguishes the blocking probabilities for the different traffic types; and therefore, no information about how the multiclass services are actually handled.

In [55], a novel predictive admission control scheme is presented. The authors propose a new cell load measurement method and mechanisms for predicting the load increase due to the acceptance of new connections. However, the threshold value, which decides whether a new connection of each traffic class should be admitted, is fixed. Moreover, there is no discussion about how the value of this threshold should be computed. It is useful to introduce an adaptive threshold algorithm which reduces the operators effort in parameter tuning [55].

In [58], the authors proposed a pre-emptive congestion control scheme for LTE. However, with the use of this method, high priority requesting bearers can displace lower priority connected bearers.

Chapter 3

Impact of CSI on RRM for the OFDMA downlink

3.1 Introduction

OFDMA is based on OFDM; and thus, inherits its key benefits while allowing for multiuser diversity to be exploited [59]. This leads to more efficient RRM as spectrum can be allocated to the users with the better channel conditions. For these reasons, RRM solutions for OFDMA systems have attracted significant interest. Most of the works on OFDMA resource allocation have considered the case where ICSI is available at the transmitter, and various algorithms based on ICSI have been developed. In [24], the rate adaptive problem was investigated. The objective was to maximize the total sum continuous rate over all users subject to power and BER constraints. It was shown that in order to maximize the total capacity, each subcarrier should be allocated to the user with the best gain on it; and the power should be allocated using the water-filling algorithm across the subcarriers. In [60] it was shown that equal power allocation accross all of the subcarriers can dramatically reduce the complexity of the resource allocation algorithm at the cost of a slight decrease in system throughput. This is due to the nature of OFDMA systems, where subcarriers are commonly assigned to the users with the best channel gains. In [24], no fairness amongst the users was considered. Thus, the users that have the best channel conditions will be assigned all the resources, which leads to bandwidth starvation of some users. This problem was addressed in [25].

There are several reasons which lead to unavailable user ICSI at the transmitter. Firstly, feedback of CSI will increase the overhead occupying more wireless resources, such as transmit power and bandwidth. Secondly, under significant user mobility, the small coherence time makes channel estimation procedures less accurate. Other reasons that contribute towards unavailable ICSI are prediction errors as well as feedback/processing delays. Therefore, in some cases, it is more reasonable for the users to send back statistical information about the channel. This knowledge is referred to as SCSI. Under SCSI based resource allocation, users only need to feed back the mean of the channel SNR distribution. This leads to fewer wireless resources such as transmit power and bandwidth being consumed for feedback purposes.

When the BS assumes that the SCSI values reflect the true channel state, a significant performance deterioration will be realized. This is due to the non-zero probability of outage events [61], [62]. The BS will assign a rate to a user based on the SCSI which may not be able to be supported by the true channel state. This will result in an unsuccessful transmission which leads to a waste of system resources. A simple solution to this problem would be to transmit at a lower rate than what is dictated by the user SCSI. However, transmitting at rates that are too low will result in unnecessary underutilization of the channel. The use of SCSI to perform resource allocation was investigated in [12]. The authors maximized the expected value of the sum-rate capacity (ergodic rates) subject to user minimum rate and power constraints.

In this chapter, a rate adaptation problem for users whose SCSI is only available at the transmitter is solved. It is shown that for Rayleigh fading channels, loading the subcarriers with a rate equal to the Lambert W function of the average SNR leads to optimal channel utilization and important system throughput gains . To further enhance the performance of the system, a well-known heuristic algorithm presented in [25] is extended. In order to investigate the impact of ICSI knowledge on RRM for the OFDMA downlink, the ICSI based resource allocation scheme is compared with the SCSI based resource allocation scheme.

Under significant user mobility it becomes more difficult for the transmitter to acquire accurate user CSI and SCSI based resource allocation is more convenient. In the final section of this chapter, simulation results show that vehicular users can benefit more from adaptive resource allocation when their CSI is perfectly known by the transmitter.

The rest of this chapter is organized as follows. The system and channel model are described in section 3.2. The methodology used to optimally load subcarriers using SCSI is discussed in section 3.2. In section 3.3, the problem formulation for the SCSI and ICSI cases is presented. A description of the algorithm used to solve the problem is also given here. In section 3.4, results that show the gains realized through optimal loading of the subcarriers are given. In section 3.6, results that depict the key factors affecting the performance difference between the ICSI and SCSI resource allocation schemes are presented. In section 3.7, the efficiency of the OFDMA downlink for different vehicular user densities is discussed followed by a summary in section 3.8.

3.2 System and Channel Model

3.2.1 System Model

A downlink OFDM system with K users and N subcarriers is considered. Each subcarrier n has a total bandwidth equal to B. The k_{th} user's minimum bit rate is denoted by R_k. Resource allocation is performed for each subcarrier (i.e subcarriers cannot be shared between users). An assignment indicator c_{kn} is defined for the k_{th} user and the n_{th} subcarrier. Therefore, $c_{kn} = 1$ when carrier n is allocated to user k and 0 otherwise.

3.2.2 Channel Model

The complex channel gain of subcarrier $n \in 1, 2, ..., N$ for user $k \in 1, 2, ..., K$ is assumed quasi-static over the duration of one frame and is denoted by h_{kn}. The channel gains are assumed to follow a complex Gaussian distribution with zero mean and unit variance, $h_{kn} \sim NC(0, 1)$.

3.3 Optimal Capacity of subcarriers under SCSI and ICSI assumptions

3.3.1 Capacity of subcarrier when ICSI is available

When the ICSI is known by both the transmitter and the receiver, the BS can adapt its transmission strategy without errors. Therefore, in this case, there is no notion of capacity versus outage where the transmitter sends bits that cannot be decoded. When the ICSI is known, the maximum capacity that can be reliably transmitted on a subcarrier n by user k is the Shannon capacity [61]:

$$C_{kn} = B \log_2(1 + \frac{p_{kn}|h_{kn}|^2}{N_o}), \tag{3.1}$$

where p_{kn} is the power allocated for transmission on this sub-carrier, and N_o is the noise spectral density.

3.3.2 Optimal Capacity of subcarrier when only SCSI is available

When the instantaneous channel conditions are unknown by the transmitter but known by the receiver, the capacity of each subcarrier is viewed as a random variable and is given as [62]:

$$C(\nu) = B \log_2(1 + \nu\bar{\gamma}), \tag{3.2}$$

where ν is exponentially distributed as Rayleigh fading is considered, and $\bar{\gamma}$ is the average value of the SNR. Under these conditions, there is a non-zero probability that the actual channel conditions cannot support an assigned rate ρ. This value is given as [62]:

$$\begin{aligned} P_{out} &= \Pr(C(\nu) < \rho) \\ &= 1 - \exp[-(\bar{\gamma}^{-1})(2^{\rho} - 1)]. \end{aligned} \tag{3.3}$$

In this case, it is observed that only the rate $\rho = 0$ is compatible with $P_{out} = 0$. The goodput is defined as the average successfully transmitted rate [30]. For user k and subcarrier n, this value is expressed as:

$$G_{kn} = \rho_{kn}(1 - P_{out}). \tag{3.4}$$

In Figure 3.1, curves of goodput vs data-rate are plotted for different values of the average user SNR $\bar{\gamma_{kn}}$. Each subcarrier can be optimally loaded by selecting the value of $\rho_{k,n}$ that maximizes $G_{k,n}$. These values of ρ for each k, n pair are given as:

$$\rho_{kn} = \frac{W(\bar{\gamma_{kn}})}{\ln(2)}, \tag{3.5}$$

where W denotes the Lambert-W function, the solution to the transcendental equation $W(x)e^{W(x)} = x$(A proof of Eq. 3.5 is presented in Appendix A of this thesis). Therefore, the maximum goodput user k can achieve on subcarrier n is:

$$G_{k,n} = \frac{W(\bar{\gamma_{kn}})}{\ln(2)}(\exp[-(\bar{\gamma}_{kn}^{-1})(2^{\frac{W(\bar{\gamma}_{kn})}{\ln(2)}} - 1]). \tag{3.6}$$

In Figure 3.2 a curve of the optimum goodput that can be achieved on a subcarrier vs the average user SNR is presented:

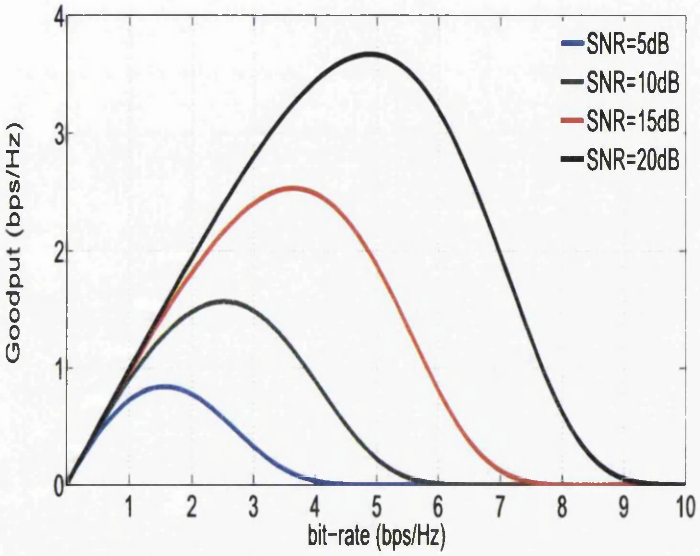


Figure 3.1: Goodput vs data-rate for different values of the average SNR

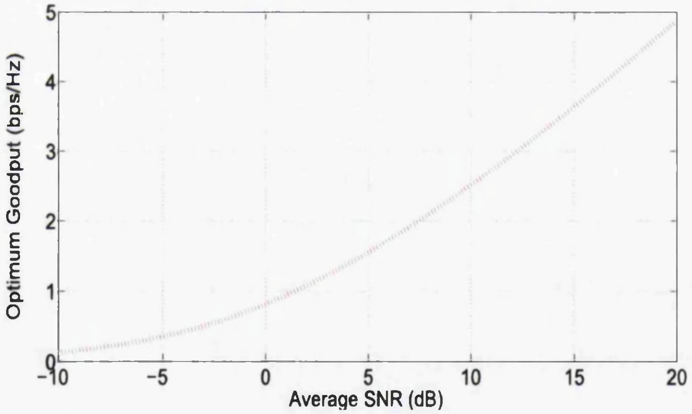


Figure 3.2: Optimum Goodput vs average user SNR

3.4 Problem Formulation

The resource allocation problem for the case where SCSI is only known at the transmitter side is formulated in this section. The objective is to maximize the sum-goodput of the OFDMA downlink under minimum user data-rate requirement constraints denoted by

R_k. Equal power allocation across all subcarriers is assumed as this reduces the complexity of the problems and minimally decreases the data throughput of a multiuser OFDM system [60]. For the SCSI based scheme, the problem can be mathematically formulated as follows:

$$\begin{aligned} P1 : &\quad \max_{c_{kn}} \sum_{k=1}^{K}\sum_{n=1}^{N} c_{kn} G_{k,n} B \\ Subject\ to : & \\ &\quad C1 : \sum_{n=1}^{N} c_{kn} G_{k,n} B \geq R_k, \forall k \\ &\quad C2 : \ If\ c_{k'n} = 1,\ then\ \ c_{kn} = 0 \,\forall\, k \neq k'. \end{aligned} \tag{3.7}$$

It should be noted that the first constraint, C1, ensures that the data-rate requirement is met for all users k, while the second constraint ensures that a single carrier is not shared between different users. By replacing the goodput with the Shannon capacity in P1, the formulation for the case of ICSI can be obtained:

$$\begin{aligned} P1 : &\quad \max_{c_{kn}} \sum_{k=1}^{K}\sum_{n=1}^{N} c_{kn} \log_2(1 + \frac{\nu P_t}{N_o}) B \\ Subject\ to : & \\ &\quad C1 : \sum_{n=1}^{N} c_{kn} \log_2(1 + \frac{\nu P_t}{N_o}) B \geq r_k, \forall k \\ &\quad C2 : \ If\ c_{k'n} = 1,\ then\ \ c_{kn} = 0 \,\forall\, k \neq k'. \end{aligned} \tag{3.8}$$

3.4.1 Complexity of the Problem

In $P1$, both the goodput and the Shannon capacity can be precalculated for all users and channels before allocation. Therefore, these values can be treated as constants. Hence, $P1$ is an integer linear programming problem which is one of the earliest members of the NP-hard class [63]. There exist KN integer variables and $K + N$ constraints, where the number of subcarriers is high in practice (i.e N= 1024/2048 used in our simulations). As the complexity of the problem grows exponentially with the values of KN and $K+N$ [64], it cannot be solved by using standard integer linear programming methods such as Branch and Bound. The associated complexity will clearly be too high for real time applications. Thus, a heuristic algorithm which will allocate subcarriers to users needs to be developed.

3.4.2 Heuristic Subcarrier-Bit Allocation Algorithm

In this work, an extended version of the algorithm presented in [25] is used to solve $P1$. In [25], the algorithm used initially allocates subcarriers to the users who can transmit the highest amount of data on them without taking any user QoS constraints into consideration. This will maximize the sum-capacity of the system [24]. Without considering the QoS constraints, the optimization problem becomes:

$$\max_{c_{kn}} \sum_{k=1}^{K} \sum_{n=1}^{N} c_{kn} G_{k,n} B \tag{3.9}$$
$$C1 : \; If \; c_{k'n} = 1, \; then \;\; c_{kn} = 0 \; \forall \; k \neq k',$$

where c_{kn} are variables, and G_{kn} are precalculated constants. The maximum can be easily achieved by setting:

$$k_n^* = \arg\max_{k} G_{kn} \tag{3.10}$$

and then letting $c_{k_n^*,n} = 1$ and $c_{k,n} = 0 \; \forall k \neq k_n^*$.

However, this process will lead to fairness not being guaranteed as many users will not have their rate constraints satisfied. This suggests, that in order to meet the rate constraints, some subcarriers need to be reallocated from the users that have the highest channel gains on them to the users whose rate constraints have not been met. This will lead to a reduction of the system sum capacity. A subcarrier should not be reallocated from a user k^* if this process leads to a violation of his data-rate requirements. In [25], the reallocation process is based on a cost function given by the following equation:

$$\delta_{kn} = \frac{G_{k^*n} - G_{kn}}{G_{k^*n}} \forall n. \tag{3.11}$$

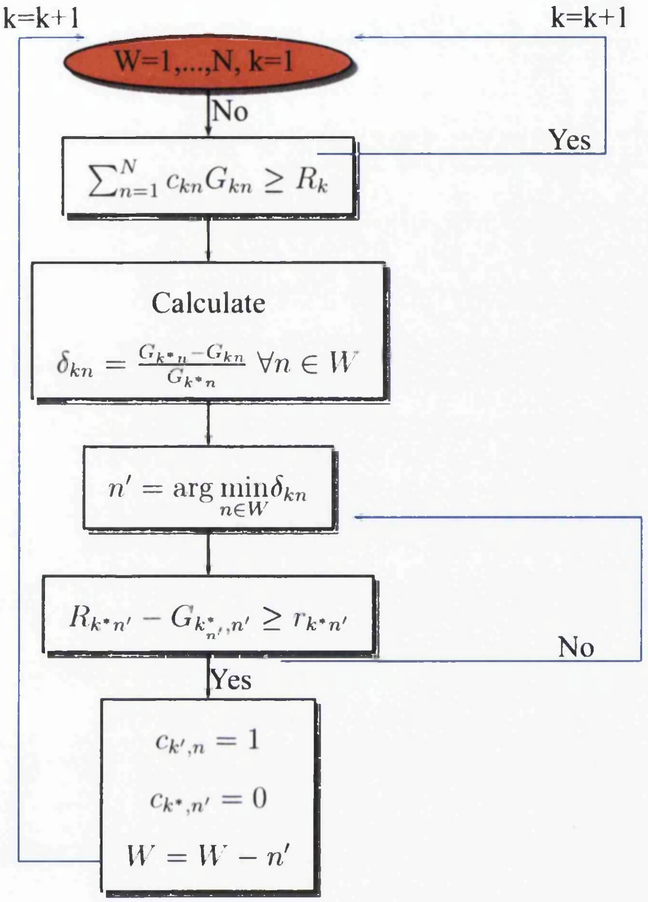


Figure 3.3: Subcarrier Reallocation Process

The cost function ensures that any reallocations cause a small reduction in the overall sum-capacity of the system, and that the running time of the algorithm is minimized. The flowchart of the reallocation process is illustrated in Figure 3.3.

3.4.3 Extension to Algorithm

Here, the aforementioned algorithm is extended as when carriers are allocated to users who do not meet $C1$, their requests will be rejected, and any carriers allocated to them will be wasted. Consider $X \subseteq \{1, 2, ..., K\}$ to be the set of users who have not had their data-rate constraints satisfied following the execution of the algorithm proposed in [25]. The cardinality of this set equals l. It is expected that a number of these users will still have been allocated some subcarriers. The data-rates allocated to these l users by [25] are given in vector **y**=[y_1,...,y_l] whereas their associated data-rate constraints are **z**=[t_1,....,t_l]. Furthermore, it is assumed that $W \subseteq \{1, 2, ..., N\}$ is the set of subcarriers that have been allocated to these l users through [25]. In order to improve the overall performance, the following extension is proposed:

```
$Initialize:\ W \subseteq \{1,2,....,N\},\ X \subseteq \{1,2,....,K\}, y, z$
$\forall$ users $u \in$ X
$Calculate\ r(u) = z(u) - y(u)$
$u^* = \arg\min_{u} r(u)$ //find the user $u^*$ closest to meeting his QoS requirement
while $y_{u^*} < t_{u^*}$ & $|W| \neq 0$ do
    $i^* = \arg\min_{i \in W} \frac{G_{k^*i} - G_{u^*i}}{G_{u^*i}}$ // $k^*$: user who was orginally allocated carrier $i^*$
    $y_{u^*} = y_{u^*} + \rho_{u^*i^*} B$ // give subcarrier $i^* \in$ W to user $u^*$
    $y_{k^*} = y_{k^*} - \rho_{k^*i^*} B$ // remove subcarrier $i^* \in$ W from user $k^*$
    $W = W - i^*$ // remove subcarrier $i^*$ from W
end while
$X = X - \{u^*\}$ // remove $u^*$ from X
```

In the algorithm given above, $\rho_{u^*i^*}$ is the optimum goodput user u^* can achieve on carrier i^*. The use of this extension enables any unused subcarriers to be assigned to the users closest to meeting their data rate requirements. Therefore, the number of utilized subcarriers increases, and a higher overall system performance can be achieved. Moreover, the number of satisfied users grows.

3.4.4 Suboptimal Properties of Heuristic Algorithm

The suboptimal properties of the heuristic algorithm used to perform resource allocation are demonstrated in Figure 3.4. The results of the algorithm are compared with the optimal results obtained through a brute force search. The simulations are run 10,000 times. The values of the channel gains used for this simulation are according to the ITU Vehicular A power delay profile.

Due to the long computational time, only eight subcarriers and three users are considered. As only eight subcarriers are considered, the data-rate constraint for each user is set to 4000 bps. All parameters used are given in Table 3.1. The algorithm efficiency is defined as the ratio of the goodput achieved through the use of the suboptimal algorithm to the goodput that can be achieved through a brute force search. Figure 3.4 shows that the algorithm exhibits excellent suboptimal properties (within 99.99% of the result achieved through the brute force search method) when SCSI is used to perform resource allocation.

Table 3.1: Simulation Parameters

Parameter	Value
Number of subcarriers	8
Number of users	3
Channel Model	ITU Vehicular A
Bandwidth per subcarrier	9.7656 kHz
Average received SNR per subcarrier	0 dB
Data-rate constraint per user	4000 bps

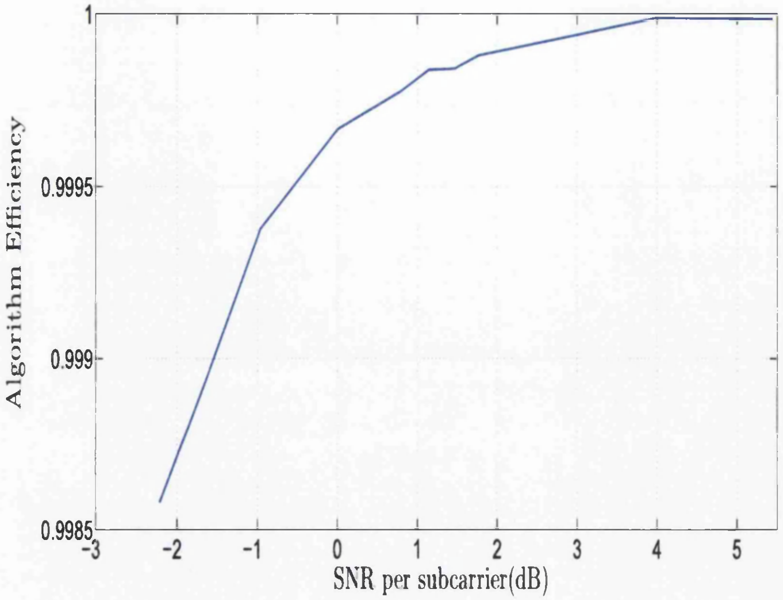


Figure 3.4: Performance comparison between brute force search and suboptimal algorithm used to perform resource allocation

3.4.5 Impact of Extension to Original Algorithm

When an optimal solution for $P1$ exists (i.e the user data rate requirements in $C1$ are not particularly high) the original and extended versions of the algorithm exhibit the same algorithm efficiency (suboptimal properties). However, when the network traffic load is high (i.e appropriate admission control is not performed hence a high number of active users with high data-rate requirements are admitted into the network) and only the users' SCSI is available, an optimal solution to $P1$ cannot be found. In this case, using the extended version of the algorithm to perform resource allocation becomes beneficial.

In order to present the importance of the final part of the algorithm, a throughput gain factor is used which is defined as the ratio of the goodput achieved with the original version to the goodput that can be achieved using the extended version. The average received power per subcarrier is set equal to 0 dB. Here, the number of users present in the cell are increased to 4 whereas there are still 8 subcarriers that can be assigned to them.

Figure 3.5 shows that when the data-rate constraint is high and SCSI is used for resource allocation, the use of the extended version leads to an important increase in throughput. As the user data-rate requirements grow, it becomes increasingly difficult for the SCSI based resource allocation scheme to meet these demands. Therefore, a higher number of unsatisfied users will be realized. With the original algorithm, any subcarriers allocated to them would have been discarded. Hence, the impact of this algorithm becomes more visible when the user demands are high.

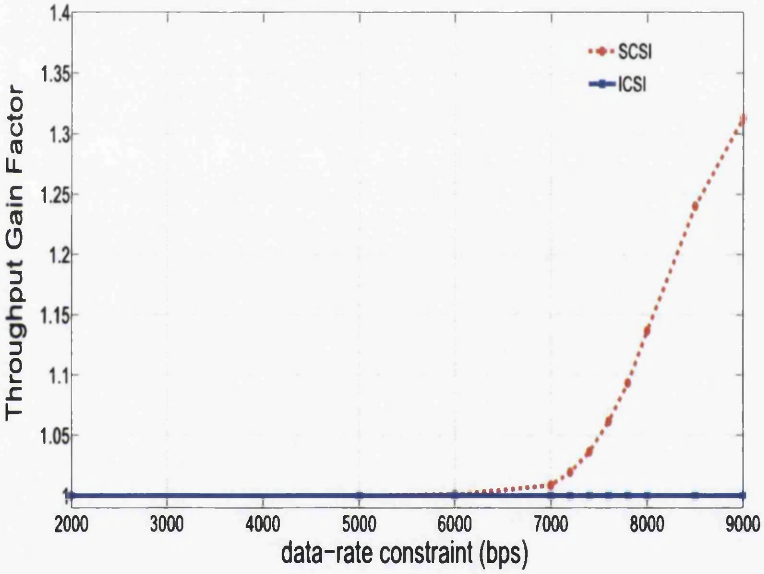


Figure 3.5: Impact of the proposed extension to the original algorithm

3.5 Importance of optimally loading subcarriers when SCSI is used

In this section, a comparison is made between the system performance when subcarriers are optimally loaded in accordance with Eq. 3.5 to the case where carriers are loaded without considering outages. The performance is evaluated in terms of satisfied user probability and goodput. Twenty users each having a data-rate constraint equal to 500 kbps were considered. The users were uniformly distributed within a cell whose radius varies between 0.5 km and 1.1 km so that the value of the average SNR of all users could be altered. The transmit power and noise power spectral density equal 46 dBm and -174 dBm/Hz. The performance is evaluated in multipath channel environments modeled as a tapped delay line with six taps as specified in the ITU Vehicular A channel model [65]. The propagation environment is described by the COST-Hata model [66].

In Figure 3.6, the impact of optimally loading the subcarriers on the sum-goodput of the system is presented. These gains are important and increase with the value of the average user SNR. Therefore, maximizing over the expectation of the distribution as in [12],[34] will lead to deteriorated performance. In Appendix B of this thesis, it is shown that as the user SNR tends towards infinity the ratio of the optimal goodput that can be achieved on a subcarrier to the goodput that can be achieved on the subcarrier without the use of optimal loading is equal to the value of e. The importance of optimal loading is also presented in Figure 3.7. A close observation of this figure shows that significantly more user QoS requirements can be satisfied when the subcarriers are optimally loaded.

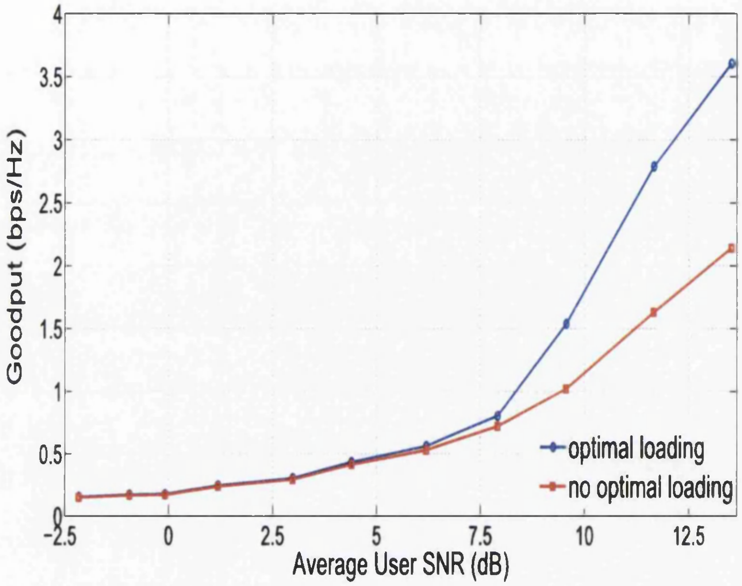


Figure 3.6: Impact of optimal subcarrier loading on sum-goodput of the system

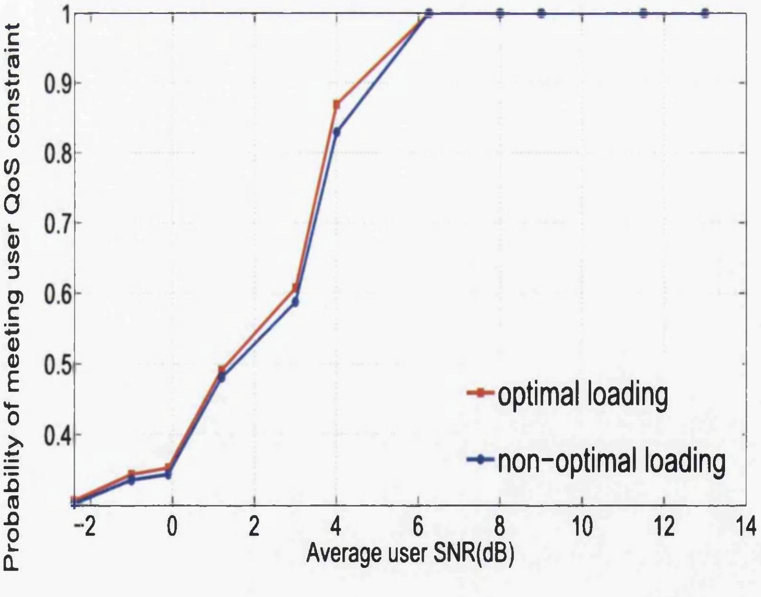


Figure 3.7: Impact of optimal subcarrier loading on satisfied user probability

3.6 Comparison of the ICSI and SCSI based resource allocation schemes

The impact of ICSI on resource allocation for the OFDMA downlink can be evaluated by comparing it to the SCSI resource allocation scheme. The SCSI based scheme considers subcarriers that are loaded with a rate given by Eq. 3.5. The key simulation parameters used to compare the two schemes are listed in Table 3.2. In order to measure the impact of ICSI knowledge, a capacity gain factor is defined. This is the ratio of the optimum sum-capacity of a system where the user ICSI is known to the optimum sum-goodput of that same system when only the user SCSI is available. The performance is evaluated in multipath channel environments modelled as a tapped delay line with six taps as specified in the ITU Vehicular A channel model [65]. The performances are evaluated using simulations over 10,000 instances of independent channel realizations. When averaging over a large number of channel realizations, accurate information regarding the comparison of the two cases can be obtained. Results were obtained using Matlab R2007B.

Table 3.2: Simulation Parameters used to compare SCSI and ICSI based resource allocation schemes

Parameter	Value
Number of sub-carriers	1024
Tx Power	$8 \rightarrow 35\ mW$ per sub-carrier
Noise power density	$10^{-10}\ W/Hz$
Channel Model	ITU Vehicular A
Bandwidth	10 MHz
data-rate constraint	1 Mbps, $0 \rightarrow 1.6\ Mbps$

3.6.1 Impact of a varying transmit power on the difference between ICSI and SCSI

In this subsection, the effect of a varying transmit power on the performance difference between the SCSI and ICSI based resource allocation schemes is presented. It is worthwhile to investigate the significance ICSI knowledge has on the probability of satisfying the user data-rate constraints. In Figure 3.8, it is noticed that an important number of data-rate constraints can be met by using SCSI. However, when the instantaneous channel realizations are known, nearly all of the user data-rate requirements can be satisfied.

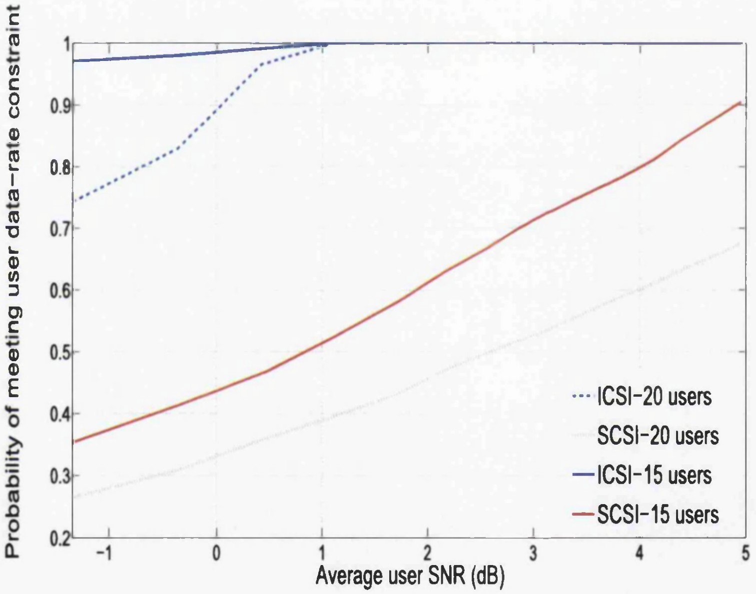


Figure 3.8: Probability of user data-rate requirements being met versus average user SNR for the SCSI and ICSI based resource allocation schemes (data-rate constraint=1Mbps)

A further comparison between the SCSI and ICSI based resource allocation schemes is made in Figure 3.9 where 4 users each have a data-rate requirement equal to 1Mbps. Here, it is observed that the SCSI scheme requires approximately 6 dB more power per subcarrier for the performance of the two schemes to be equal. Figure 3.10 presents the variation of the capacity gain factor with SNR per subcarrier for 4 users without data-rate constraints. For comparison purposes, a curve corresponding to the same amount of users each requiring 1 Mbps is also given. It can be observed that the values of the capacity gain factor are lower when there are no data-rate constraints. This is because more subcarrier reallocation operations are needed when SCSI is used for the data-rate demands to be met. As the SNR per subcarrier increases, the performance difference between the two cases is reduced because more data-rate constraints can be met using SCSI. Also, the actual value of the capacity gain factor decreases as the power grows. In Appendix A of this thesis, it is shown that as the power approaches infinity, the value of the capacity gain factor will be equal to one. In general the figures show that the performance difference between the ICSI and SCSI based resource allocation schemes is small. This is attributed to the Lambert-W approach which optimally loads the subcarriers in a Rayleigh environment. Optimal loading of subcarriers will lead to an improved performance when the channel gains are distributed according to a different p.d.f (i.e Ricean). However, closed form expressions that relate the optimum goodput with the data rate need to be developed.

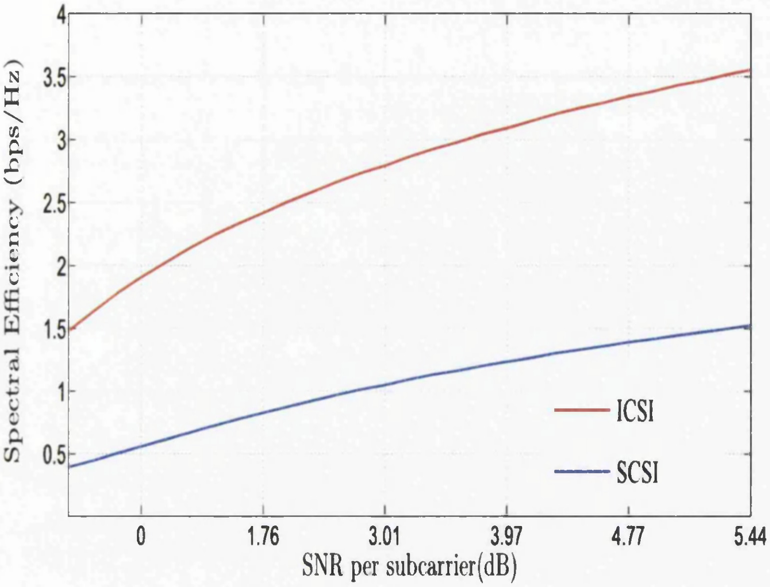


Figure 3.9: Maximized sum-capacity versus average subcarrier SNR

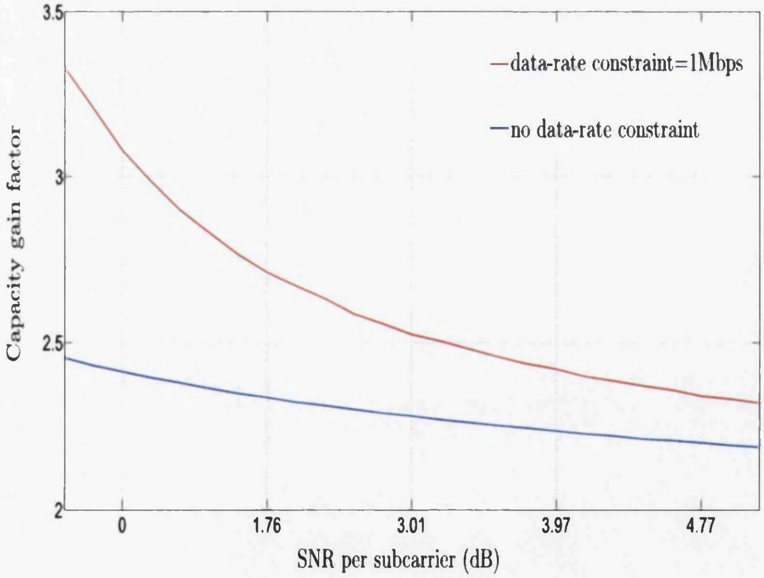


Figure 3.10: Impact of SNR per subcarrier on the capacity gain factor

3.6.2 Impact of data-rate constraint on difference between SCSI and ICSI resource allocation schemes

The user data-rate constraint in P1 plays an important role on the performance difference between the SCSI and ICSI based resource allocation schemes. Figure 3.11 shows the variation of the capacity gain factor with the user data-rate requirement. The transmit power equals 10 mW (SNR per subcarrier is equal to 0 dB). A close observation of this figure indicates that the value of this factor grows with the data-rate constraint. As the user demands increase, the number of subcarrier reallocation operations required to satisfy users whose instantaneous channel realizations are unavailable grows. This process has a negative impact on the optimum sum-goodput of the SCSI based scheme. Moreover, multiuser diversity has a strong effect on the results of Figure 3.11. When the users' ICSI is known by the transmitter, it is easier for multiuser diversity to be exploited. However, when SCSI is used to perform resource allocation, the benefits of multiuser diversity outweigh the drawbacks of subcarrier reallocation only when the data-rate constraint is low (i.e < 0.1 Mbps in our simulations).

It is important to further investigate the significance of multiuser diversity on these results. In Figure 3.12, it is noticeable that when the user data-rate requirement is equal to 0.1 Mbps and the SCSI based resource allocation scheme is used, the overall system throughput increases with the number of users. However, it becomes increasingly difficult to exploit multiuser diversity as the data-rate requirements grow. In this case, the effects of subcarrier reallocation counteract multiuser diversity even when a relatively low number of active users are present in the cell. On the other hand, using ICSI based RRM enables multiuser diversity to be utilized much more efficiently when the data-rate demands are high.

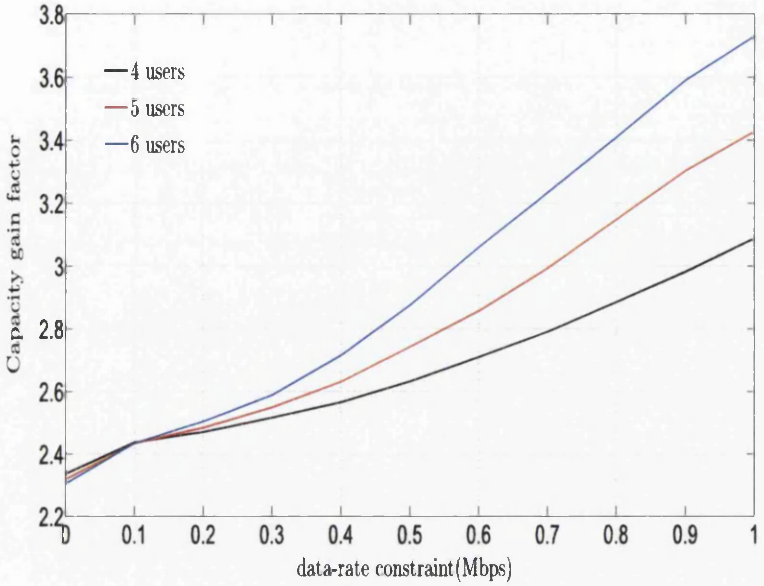


Figure 3.11: Capacity gain factor dependence on data-rate constraints

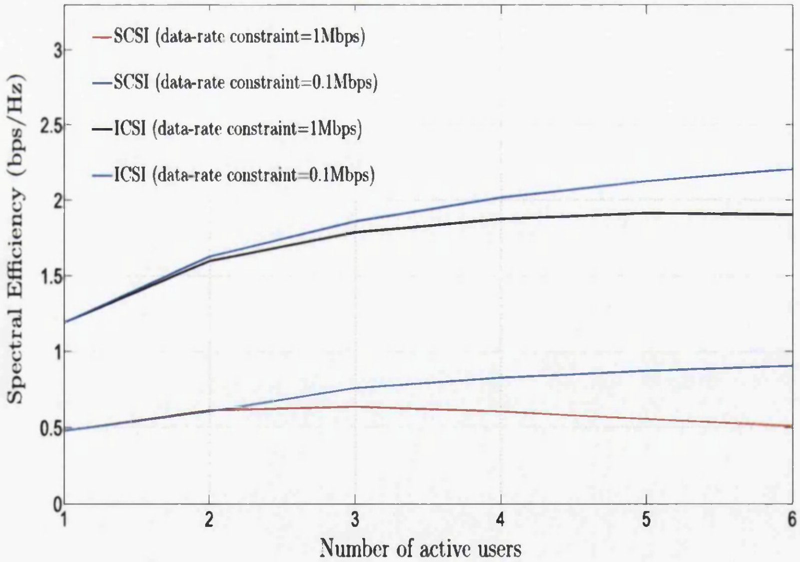


Figure 3.12: Effect of data-rate constraint on multiuser diversity

3.7 Impact of channel model mixtures on RRM for the OFDMA downlink

Operators commonly specify a certain percentage mix of the ITU representative models to capture the propagation characteristic specifics of a given environment [67]. This is because some of the active users that are present in the cell will be pedestrian whereas others will be moving at vehicular speeds. This implies widely varying coherence time and bandwidth. In this section, the efficiency of the OFDMA downlink for different vehicular user densities is studied. The evaluation of LTE techniques demands channel models with increased bandwidth compared to UMTS models, to reflect the fact that the characteristics of the radio channel frequency response are connected to the delay resolution of the receiver [43]. In 3GPP the 20 MHz LTE channel models were based on a synthesis of existing models such as the ITU and 3GPP models [43]. In a 3GPP LTE environment, the following channel models are used: EPA (This model covers pedestrian users with speeds upto 3 km/hr), EVA (this model covers mobile users with speeds up to 50 km/hr) and ETU (This model covers moving vehicles with speeds upto 90 km/h) [43]. It is assumed that the exact proportion of users adhering to each channel model (vehicular or pedestrian) depends on the geographic location of the BS. In particular, we assume that the users which exhibit high mobility (i.e those with vehicles) are most likely to be found in a rural area. Table 3.3 shows the vehicular user density for each considered geographic area.

Table 3.3: Vehicular user density in each considered geographic area

Location	ITU Ext Veh A(%)	ITU Ext Ped A(%)
Rural Areas	40	60
Urban Areas	10	90
Sub-Urban Areas	30	70

In order to evaluate the spectral efficiency of the OFDMA downlink in the considered geographic locations, simulations over 10,000 instances of independent channel realizations are performed. Fourty users each having a 500 kbps data rate requirement are assumed to be uniformly distributed within a cell of varying radius. The cell is varied between 0.5 km and 1.1 km. The total bandwidth is set to equal 20MHz which is divided into 2048 subcarriers.

As under significant user mobility, the CSI cannot be considered as perfect two sets of results for Figure 3.13 are plotted. For the first set of results, simulations are performed under the realistic assumption that ICSI can only be used to allocate resources to the pedestrian users. For the second set of results we perform simulations under the assumption that ICSI can be used to allocate subcarriers to both vehicular and pedestrian users. From this figure, it can be observed that when ICSI is only used to allocate resources to pedestrian users the highest overall system throughput can be achieved when the transmitter is in an urban area where a large number of pedestrian users can be found. However, if we assume that ICSI based resource allocation can be used for both vehicular and pedestrian users then the highest overall system goodput can be found in rural areas. To see why this is so, the spectral response of the two PDP's accross 1024 subcarriers is plotted in Figure 3.14. It is noticeable from this figure that more multipath is associated with the extended vehicular power delay profile. This is attributed to the much larger delay spread of the channels. Multiuser diversity can thus be better exploited when the user channels are characterized by the EVA PDP. Therefore, when the user CSI is known, adaptive subcarrier allocation leads to higher gains in the vehicular environment than in an area with little detected multipath propagation. In order to verify these results, all users are located at an equal distance from the transmitter (i.e only multipath fading is considered). Results for this simulation are presented in Figure 3.15. Again, it is noticed that when the ICSI of all users is perfectly known, higher system throughputs are realized in the rural area.

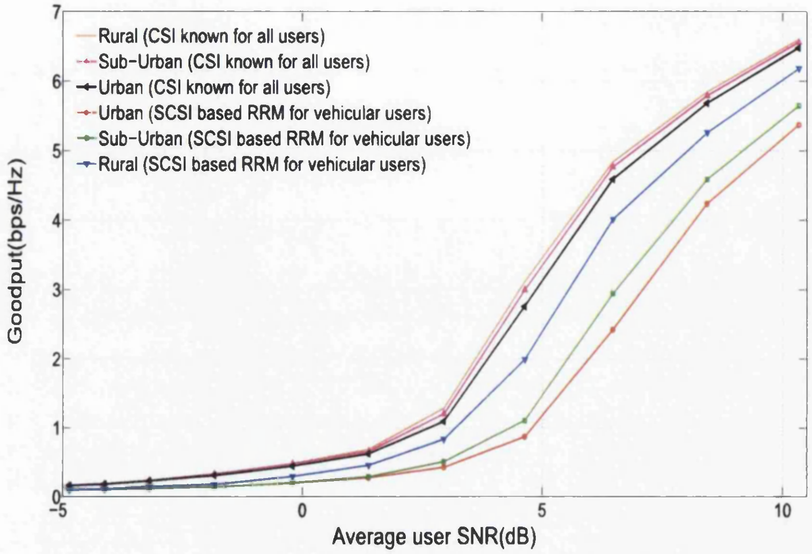


Figure 3.13: Goodput vs average user SNR for different geographic locations of the transmitter

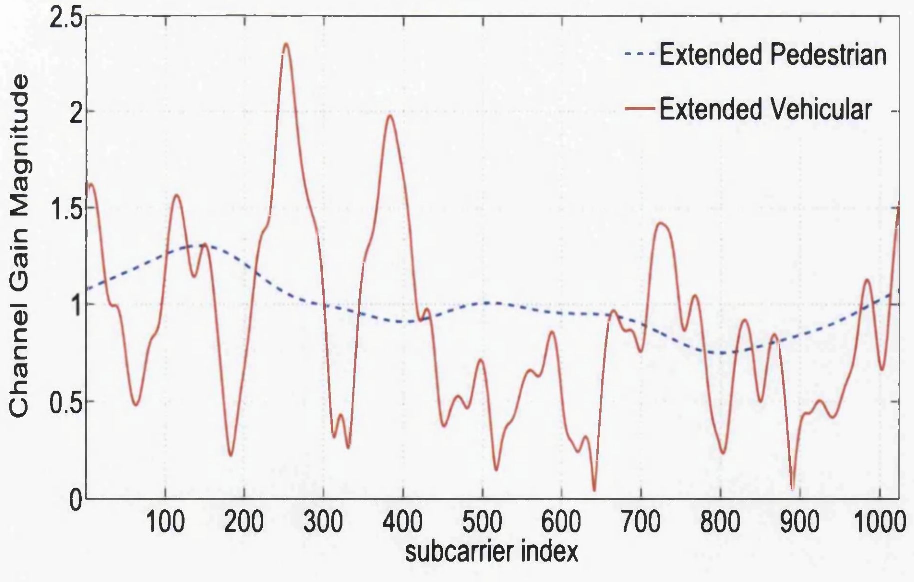


Figure 3.14: Spectral response of the EVA and EPA power delay profiles

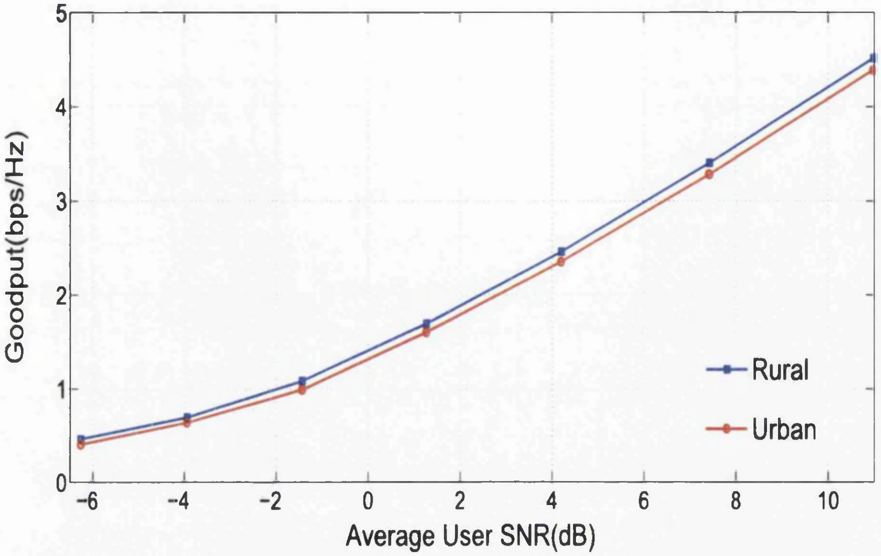


Figure 3.15: Goodput vs average user SNR when users are located at an equal distance from the transmitter and the ICSI of all users is perfectly known

3.8 Summary

In this chapter, a rate adaptive resource allocation problem was solved for users whose SCSI is only known by the transmitter. Under these conditions, it was shown that optimal channel utilization for Rayleigh fading channels, can be realized by transmiting at a rate that depends on the Lambert W function of the average SNR. To further enhance the overall spectral efficiency, a well-known sub-optimal algorithm was extended. Using the proposed approach, numerical results showed that a significant number of user data-rate requirements could be met using SCSI. However, an important loss in performance was observed when a performance comparison between the SCSI and ICSI based resource allocation scheme was made. This degradation was dependant on the number of active users present in the cell, the data-rate constraint, and the transmit power.

The efficiency of the OFDMA downlink in different geographic locations was also investigated. It was shown that in geographical areas with a high proportion of vehicular users, a significant throughput loss is observed because of the lack of accurate CSI. However, under the assumption that the CSI of the vehicular users is known, adaptive resource allocation leads to higher throughput gains in the areas where many vehicular users can be found.

Chapter 4

RRM for fast fading environments

In the previous chapter, the significance of accurate CSI on RRM techniques for the OFDMA downlink was investigated. This chapter investigates the significance of periodic CSI feedback on RRM techniques for the OFDMA downlink in a fast fading environment.

4.1 Introduction

Efficient RRM over an OFDMA system depends on the availability of accurate CSI at the transmitter [6]. Its knowledge allows for the system bandwidth to be fully exploited through adaptation of the transmission parameters used on each OFDMA carrier. Although the system throughput can be improved with accurate CSI [6], this assumption is unrealistic.

RRM in OFDMA systems involves the application of AMC. AMC provides the flexibility to match the applied MCS scheme to each user's channel conditions. It was concluded in [27] that AMC is not significantly affected by noisy channel estimation if a reasonably good estimator is used. However, when the mobility of the users is high, the CSI becomes more outdated, which will lead to errors in the resource assignments; and thus, to a decrease in overall system performance. As user mobility is a principal driving force for mobile-OFDMA based wireless systems such as LTE, it becomes more important to consider the time-varying nature of channels for resource allocation problems in order to further enhance the system throughput. In LTE, resources are allocated to users every 1msec [43] which allows for a quick response to the varying channel conditions of the link. However, under significant user mobility the small coherence time means that the CSI measurement reports need to be frequent; otherwise, the CSI used by the BS to perform AMC will not be correlated with the true channel states. Here, correlated CSI (CCSI) is defined as the CSI that is correlated to some degree with the true channel state. Frequent CSI measurement reports lead to increased system overhead requirements, especially when a large number of active users are present in the cell [11]. Instead of feeding back the instantaneous channel coefficients to the BS, it is more reasonable that users simply send the mean of the channel's SNR distribution [12]. Using SCSI to perform resource allocation requires significantly less resources to be occupied for feedback purposes.

In practical systems, only imperfect/outdated CSI and SCSI can be used to perform resource allocation. Therefore, practical resource allocation schemes that account for CSI inaccuracy are required. The impact of imperfect and outdated CSI on the performance of adaptive OFDM has already been studied in the literature [26],[27]. Using SCSI to perform resource allocation for OFDMA systems has been analyzed in [12]. In this chapter, CCSI and SCSI based resource allocation strategies are addressed.

Most works [68] – [69], [30] on resource allocation for OFDMA systems under the partial CSI assumption do not differentiate between the SCSI and CCSI concepts. In [12], a comparison that focused on the continuous rate case (i.e Shannon capacity based formulation) was performed between the two cases. In the first part of this paper the two cases are compared for the more practically relevant case where only a discrete number of modulation and coding levels are available. In order to do so, the performance degradation created by CSI errors is first quantified. These results are then applied to allocate resources to users with the objective of maximizing the overall system throughput while ensuring that the target bit error rates are met and that each user's queue length is kept within stable bounds. The impact of this maximization on the system throughput for both the SCSI and CCSI based resource allocation schemes is investigated.

The majority of research on OFDMA resource management does not consider the presence of a CAC unit [12], [30], [46] [68] – [69]. Works are focussed on providing solutions to rate-adaptive or margin-adaptive resource allocation problems [6]. The CAC unit limits the number of admitted flows in order to maintain the user QoS. Also, it distributes the network throughput between the supported services [70]. Motivated by the comparison between the CCSI and SCSI based resource allocation schemes in the second part of this paper a strategy is proposed that leads to a good tradeoff between overhead consumption and fairness as well as throughput when the presence of the CAC unit is considered.

The rest of this chapter is organized as follows: the system model is presented in section 4.2. The methodology used to optimally load subcarriers when a constraint is imposed on the BER is discussed in section 4.3. The scheduling algorithm used in this simulations and the strategy proposed to make good use of the available resources for feedback is presented in section 4.4. The network simulation is described in section 4.5. Simulation results are presented in section 4.6 followed by a summary in section 4.7.

4.2 System Model

A downlink OFDM system with K users and N subcarriers is considered. The time axis is divided into TTIs of length 1msec as specified in the LTE standard [43]. During each TTI, packets of fixed length arrive for each user at a given rate. Each user requests access to the same service whereas a CAC unit is assumed to limit the number of incoming flows so that the network can offer each flow the QoS that it requires. If a user request is accepted, the arriving packets are buffered in separate queues. At the beginning of each TTI the BS schedules bandwidth transmission and allocates resources to each user according to their queue state and estimated/average SNR.

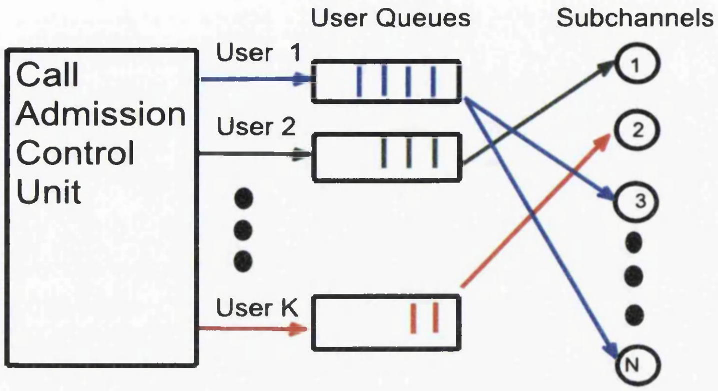


Figure 4.1: System Structure

A set of MCS levels are employed for AMC. Each MCS level is indexed by $m \in \{1, ..., M\}$, and its selection is determined by the value of the imperfect SNR which is denoted by $\hat{\gamma}$. The range of SNR values used for transmission is divided into M intervals by M SNR thresholds which are denoted by Γ_m. The values of Γ_m are chosen in such a way that the information rate is supportable subject to a target BER constraint. Therefore, MCS m will be applied when $\Gamma_m \leq \hat{\gamma} < \Gamma_{m+1}$. Also, it is assumed that $R_1 < R_2 < ..... < R_M$, where R_m is the spectral efficiency of MCS m. Each user k's complex channel gain on subchannel n is denoted by h_{kn}. The channel gains are assumed to follow a complex Gaussian distribution with zero mean and unit variance, $h_{kn} \sim NC(0, 1)$.

4.3 Maximization of Throughput using CCSI and SCSI

Consider mobile user k whose outdated channel gain for subchannel n is denoted by $\hat{h}_{kn}$ and the current channel gain which is h_{kn}. According to [27] and MMSE estimation theory [71], when both $\hat{h}_{kn}$ and h_{kn} are complex Gaussian variables with a zero mean and unit variance, the conditional pdf $P(h_{kn}|\hat{h}_{kn})$ follows a complex Gaussian distribution with a mean equal to $\mu_k = \rho_k \hat{h}_{kn}$ and a variance of $\sigma_k = 1 - \rho_k^2$. In these expressions, ρ_k denotes the correlation coefficient between h_{kn} and $\hat{h}_{kn}$ and its value is $\rho_k = J_0(2\pi f_{d,k}\tau)$. Here, J_0 is the zeroth order Bessel function of the first kind, $f_{d,k}$ is the maximum Doppler frequency for user k, and τ is the channel feedback delay. In practical systems, if the maximum Doppler frequency can be estimated, this correlation coefficient can be obtained [72]. As $P(h_{kn}|\hat{h}_{kn})$ is Gaussian distributed, $P(|h_{kn}| \,\big|\, \hat{h}_{kn})$ will follow a Ricean distribution. When the CSI is outdated, there will always be a nonzero probability of selecting an MCS level which is not optimized for the region where the current SNR lies. P_{knm} is defined as the probability of successfully selecting an MCS level m on sub-channel n for user k. It is given as

$$P_{knm} = P(\gamma_{kn} > \Gamma_m) = Q_1\left(\frac{\rho_k\sqrt{\frac{\hat{\gamma}_{kn}}{\bar{\gamma}_{kn}}}}{\sigma_k}, \frac{\sqrt{\frac{\Gamma_m}{\bar{\gamma}_{kn}}}}{\sigma_k}\right), \tag{4.1}$$

where

$$Q_1(\alpha, \beta) = \int_{\beta}^{\infty} x I_o(\alpha x) \exp^{(-\frac{\alpha^2+\beta^2}{2})}, \tag{4.2}$$

is the Marcum Q function [73] and $\bar{\gamma}_{kn} = \frac{P_r}{N_o B}$ is defined as user $k's$ mean SNR on subcarrier n. The outdated and current SNR are $\hat{\gamma}_{kn} = \frac{P_r|\hat{h}_{kn}|^2}{N_o}$ and $\gamma_{kn} = \frac{P_r|h_{kn}|^2}{N_o}$, respectively, where P_r is the received power and N_o is the noise spectral density. In Figure 4.2, the impact of $f_d\tau$ on the probability of successfully exceeding a given SNR threshold is presented. When SCSI is used to select MCS levels, P_{knm} for a Rayleigh fading channel becomes [2]:

$$P_{knm} = P(\gamma_{kn} > \Gamma_m) = \exp\Big(-\frac{\Gamma_m}{\bar{\gamma}_{kn}}\Big). \tag{4.3}$$

As the use of outdated CSI or SCSI leads to errors when AMC is performed, it is obviously useful to minimize them. In order to achieve optimal utilization of the available spectral resources, the optimum MCS level m^* for each user k on subchannel n is chosen according to the following rule:

$$m^* = \arg\max_{m\in M} P_{knm}R_m \tag{4.4}$$

where P_{knm} is given by either Eq. 4.1 or Eq. 4.3 depending on whether outdated CSI or SCSI is used.

In this setting, it is denoted that the maximum bandwidth normalized information rate that can be reliably transmitted by user k on subchannel n by G_{kn} is given as:

$$G_{kn} = P_{knm^*}R_{m^*}. \tag{4.5}$$

The quantity G_{kn} is referred to as throughput throughout this chapter.

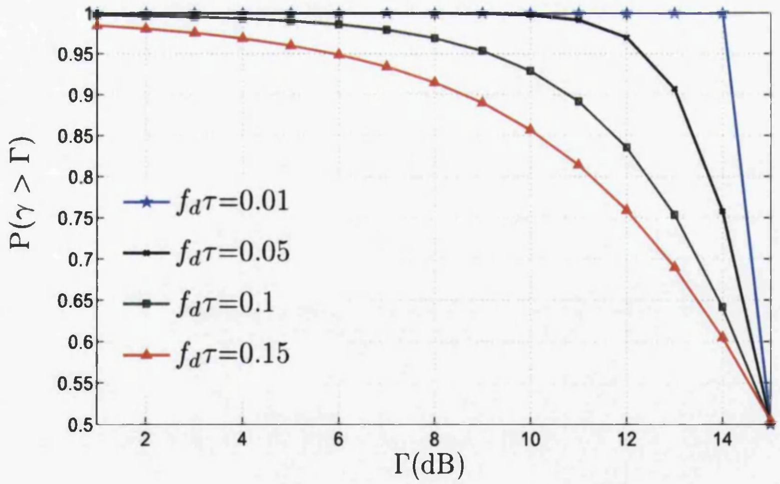


Figure 4.2: Probability of Exceeding an SNR threshold (Γ) vs Γ for different values of the Doppler-Delay product

4.4 Optimal Scheduling under practical assumptions

4.4.1 Scheduling Algorithm

A number of practical scheduling algorithms which assign resources to a set of selected users can be integrated into an OFDMA system [74]. Queue and Channel aware algorithms are able to meet user QoS requirements, as one may need to sometimes schedule users whose delays/queues are becoming large even though their current channel state is not the most favorable. It was shown in [75], that Queue and Channel aware scheduling algorithms are throughput optimal; and that they lead to significant performance improvements for the LTE system [76]. In this work, equal power allocation across all subcarriers is assumed. It is noted in [20] that the throughput degradation arising from such an assumption is negligible when AMC is applied as with the case of LTE.

In order to maximize the throughput of the system whilst meeting the target BER of each user, a simple Queue and Channel aware algorithm will allocate a subcarrier n to the user k^* for which the following holds:

$$k^* = \arg\max_k P_{knm^*} R_{m^*} W_{k^*}(t) \tag{4.6}$$

where $W_k(t)$ is the head-of-line packet delay or queue length for user k during TTI t; and $P_{knm^*} R_{m^*}$ is the channel capacity that meets the target BER/BLER requirement assuming that the optimum MCS level m^* is selected according to Eq. 4.4. The throughput user k^* can achieve on subchannel n is:

$$G_{k^*n} = P_{knm^*} R_{m^*} B \tag{4.7}$$

where B is the bandwidth of a subchannel.

For a subcarrier n to be assigned in accordance with Eq. 4.6, the scheduler needs to search KM values of $P_{knm}R_m$ as the queue length can be considered a constant. Therefore, the computational complexity associated with allocating N subcarriers to K users is $O(KMN)$.

4.4.2 Strategy to reduce overhead load

As the user velocity increases, the CSI has to be updated more frequently. Thus, large amounts of spectrum resources need to be reserved for overhead purposes if AMC ans scheduling is to be performed using CSI values that are correlated with the current value of the CSI. This leads to increased overhead requirements. The overhead load increases when a large number of active users are simultaneously present in the cell [11]. In particular, NKM bits per timeslot are required where M is the number of bits required to quantize a real number with negligible quantization error [41]. This is clearly impractical for future mobile OFDMA systems such as LTE, as the capacity allocated for CQI feedback purposes is limited [77]. Therefore, it is useful to limit the amount of users that feedback their ICSI.

In this work, ICSI feedback is confined to a set of users. Others only feedback their SCSI which requires much fewer overhead resources. When users feedback their SCSI only a single value needs to be reported for the whole system bandwidth [75]. Also, this value does not need to be updated as frequently as ICSI. In this work, the feedback of ICSI is confined to a set of users by making use of the following considerations:

In networks that support heterogeneous applications with diverse throughput requirements, the CAC unit becomes crucial. Firstly, it limits the number of incoming flows so that the required QoS can be provided to each flow. Secondly, it provides QoS guarantees by distributing the network throughput between the supported services [70]. This is particularly important as a particular group of subscribers may be more demanding than the rest, which results in an allocation of the network resources to the former and leaves the latter unsupported. Limiting the aggregate rate that the group of demanding subscribers receives imposes fairness and guarantees QoS for each service.

When the CAC unit functionality is accurate, all users admitted into the network need to be satisfied in terms of QoS. A user situated at the edge of the cell requesting access to a specific service will be allocated more channels than a user with a high average SNR accessing the same service. Therefore, by limiting ICSI feedback to the set of users with the lowest geometry (i.e., users SNR induced by the path-loss/shadowing model), more of the channels assigned by the scheduler will be loaded with CCSI. This scheme is also expected to increase the system fairness. This is because the scheduling rule is a function of both the maximum achievable data-rate and the queue length. Therefore, when users with poor channel conditions feedback their ICSI and the BS only knows the SCSI of the users with good channel conditions, they are more likely to meet the criterion of Eq 4.6. Otherwise, when the BS only knows the SCSI of the users with a low average SNR, they will be scheduled only when their queue lengths become very large. When only the users with a high average SNR feedback their CSI the fairness will decrease even further.

4.5 Network Simulation

In order to evaluate the importance of periodic CSI feedback as well as the strategy used to reduce the amount of feedback bits, system level simulations are performed. Simulation parameters are based on [78], and these are typical values used for LTE simulation studies [75]. A system with $10MHz$ of bandwidth divided into 666 subcarriers is considered, of which 624 are used for data. The remaining 32 subcarriers are used as guard subcarriers which also need to be accounted for. The width of each carrier equals 15KHz. Resource allocation cannot be performed on a per subcarrier basis due to the resulting overhead and is based on subchannels. In LTE, [75] each subchannel consists of 12 subcarriers. Thus, N= 52 subchannels can be assigned to the users. The wireless environment is typical Urban non Line of Sight, and the carrier frequency equals 2GHz. The cell diameter is 1km; and the distance, d_k, between the k_{th} user and the BS is a 2-D uniformly distributed random variable. The most suitable path loss model in this case is the COST 231 Walfisch-

Ikegami (WI) [79] as it allows estimation of the pathloss from 20m [79]. The system level simulation parameters are summarized in Table 4.1. Fig 4.3 shows the empirical CDF of the users geometry, i.e., users SNR induced by the path-loss/shadowing model (shadowing is constant for each user in each simulation run).

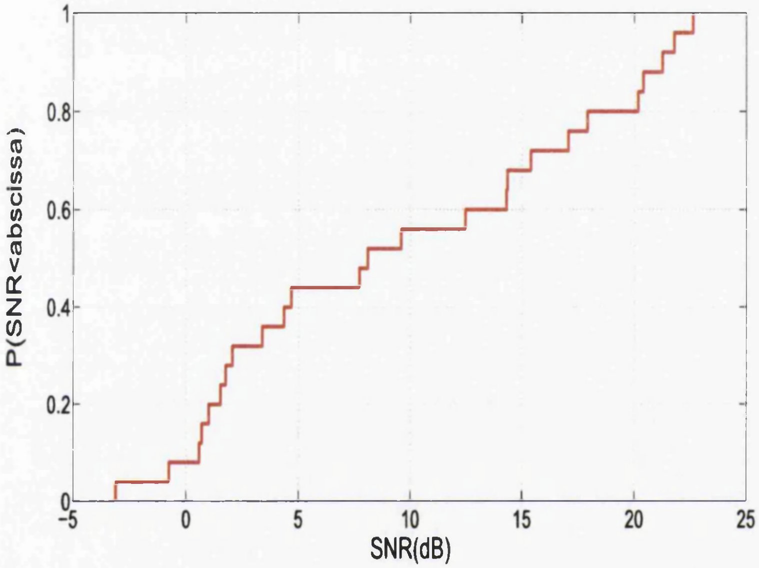


Figure 4.3: Empirical CDF of users' geometry

A schematic diagram of the simulation flow is given in Figure 4.4. When the simulation begins, each of the (K=25) users moves in a given random direction. The simulator updates the user location every 100 TTIs. During each TTI, packets arrive for each user k's queue at a rate equal to the packet arrival rate. The packet size is selected such that the system capacity is roughly equal to 1 packet/user/TTI. In order to assign subchannels to these users so that packet transmission can occur, the optimum MCS level m^* is required $\forall k, n$. When SCSI is used, the pathloss model leads to the the average user SNR $\bar{\gamma}$ through which m^* can be obtained using Eq. 4.4. However, when CCSI is used, the values of $\hat{h}_{kn}$ are also needed. These depend on the power delay profile and the distribution of h_{kn} ($h_{kn}|\hat{h_{kn}} \sim NC(\rho\hat{h}_{kn}, \sqrt{1-\rho^2})$. Moreover, the values of ρ are also required. These are a function of τ (delay time between the channel estimation and the actual transmission) and the user velocity. The value of τ differs in each simulation run and is added to the CSI processing delay to obtain the total delay time. Using these values, Eq. 4.4 allows the optimum MCS level m^* to be found $\forall k, n$ for the case of CCSI. Finally, the queue and channel aware scheduler allocates subchannels to the users in accordance with Eq. 4.6.

For each point in the figures presented in the Results section, the simulator is run for 1000 TTI's which for the LTE system is equivalent to a 1 second real-time period. In order to reduce the computational load, link level simulation results are prepared in advance in the form of look-up tables for the throughput calculation; and these give the required SNR values needed to meet a specific target bit error rate. The defined MCS levels use coding rates between $1/8$ to $2/3$ combined with QPSK, 16QAM, and 64QAM modulation schemes. The MCS levels used in our simulations are shown in Table 4.2.

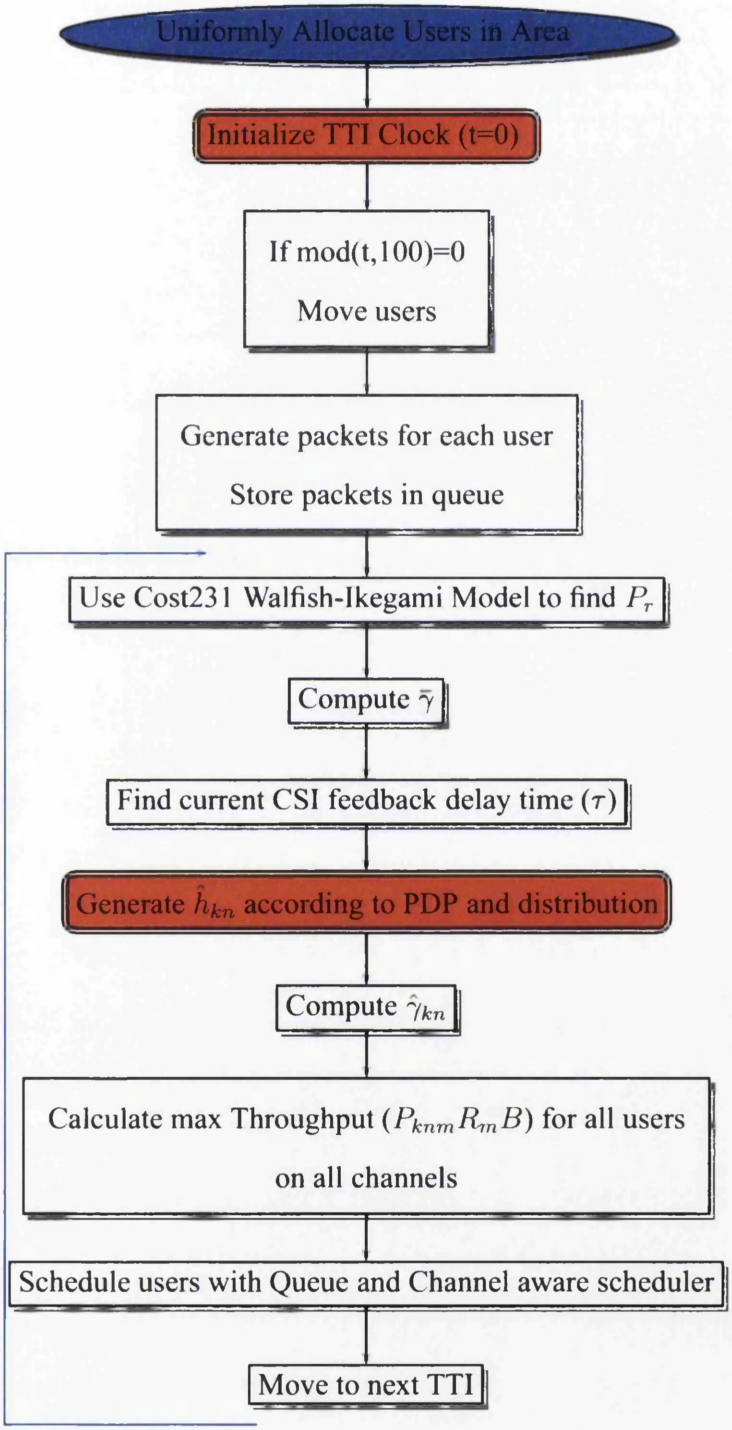


Figure 4.4: Block Diagram of Simulation Flow

Table 4.1: Key Simulation Parameters used to evaluate CCSI and SCSI based resource allocation schemes

Parameter	Value
Carrier Frequency	2 GHz
Cell Configuration	single cell
Cell Radius	1 km
Channel Bandwidth	10 MHz
subcarrier spacing	15 KHz
BS Tx Power	46 dBm
BS Antenna Height	50 m
MS Antenna Height	2 m
Mean Buliding Heights	12 m
Mean Width of Streets	50 m
Mean Building Separation	100 m
Incident angle relative to street	90^o
Path-Loss Model	COST 231 Walfisch-Ikegami
Propagation Model	ITU Vehicular A
Shadowing Log-Normal Deviation	8 dB
Thermal Noise Density	-174 dBm/Hz
Number of active Users	25
Packet Size	500 bits
Packet Arrival Rate (P.A.R)	1,3 packets/user/TTI
CSI Measurement Error	Ideal
CSI Processing Delay	1 TTI
CSI Reporting Period	2 TTI
TTI length	1 msec

Table 4.2: Selection of MCS based on received SNR and the corresponding throughput

SNR(dB)	Modulation	Coding Rate	Throughput (bps/Hz)
~ -5	No use		
-5 ~ -1.9	QPSK	1/8	0.25
-1.9 ~ 1.8	QPSK	1/4	0.5
1.8 ~ 3.8	QPSK	1/2	1
3.8 ~ 7.1	QPSK	2/3	1.33
7.1 ~ 9.3	16QAM	1/2	2
9.3 ~ 11.3	16QAM	2/3	2.67
11.3 ~ 14.5	64QAM	1/2	3
14.5 ~ 17.2	64QAM	2/3	4
17.2 ~ 19.5	64QAM	0.81	4.86
19.5 ~	64QAM	2/3	5.25

4.6 Results

In this section, the impact of optimally selecting MCS levels on the system throughput when either CCSI or SCSI are used to perform AMC is presented. System throughput refers to the maximum spectral efficiency that can be reliably transmitted. A comparison between the SCSI and CCSI resource allocation schemes in a fast fading environment is also presented. Then, it is shown that the limited resources reserved for feedback purposes can be optimally used when the users with the lowest average SNR's send their CSI to the BS.

4.6.1 Impact of optimally selecting MCS levels

As AMC is always performed using imperfect CSI, there will be a nonzero probability that the selected MCS level will not be optimized for the region where the current SNR lies. Figure 4.5 compares the performance of a system when MCS levels are selected using the outdated/average SNR with the performance of the same system when Eq. 4.4 is used to perform AMC.

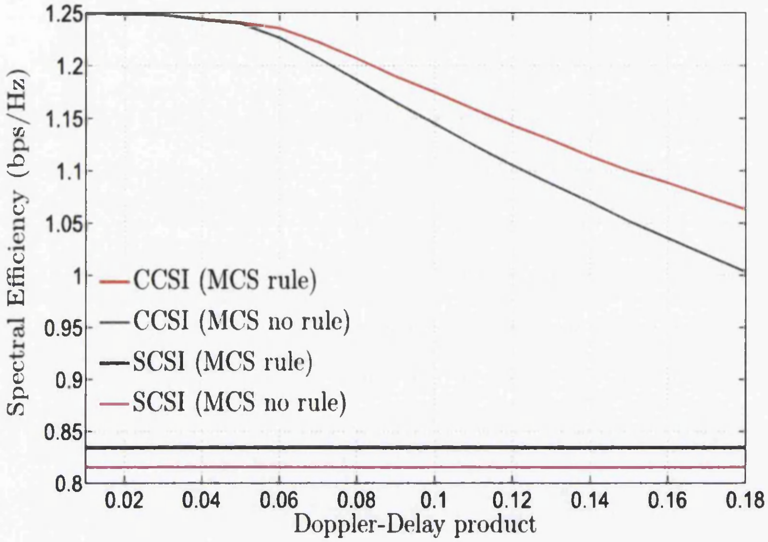


Figure 4.5: Impact of optimally selecting MCS levels

For the periodic CSI feedback scheme, the performance gains are lower when the value of the Doppler-Delay product $f_d\tau$ is small. In this case, the feedback delay is low, and few MCS level selection errors occur. As the value of $f_d\tau$ grows, the CSI becomes more outdated, and more MCS level selection errors occur. Therefore, the impact of the MCS level selection scheme grows with the value of $f_d\tau$. A close observation of this figure shows that using the rule of Eq. 4.4 to select MCS levels also leads to important performance gains when channels are adaptively modulated using SCSI. These results show that for the case where only a discrete number of modulation and coding levels are available for AMC maximizing the expected sum-rate leads to important throughput gains. These results are in agreement with those of [12] which focus on the continuous rate case (Shannon-capacity based formulation) and therefore have more theoretical rather than practical significance.

It is also important to present the impact of imperfect CSI on the bit error rate of the system as the value of $f_d\tau$ varies when a constraint on the target BER is not imposed on the system (MCS levels are not selected according to Eq. 4.4). As shown in [20], the instantaneous BER for M-QAM modulation schemes (as well as for BPSK) can be approximated for each user as:

$$BER(\gamma) \approx 0.2\exp^{-1.6\frac{\gamma}{2^{\lfloor r(\gamma)\rfloor}-1}}, \tag{4.8}$$

where $r(\gamma)$ denotes the number of bits per symbol corresponding to the applied modulation scheme and $\lfloor x \rfloor$ is the floor operation which provides the largest integer not greater than x. Figure 4.6 presents results showing the variation of the bit error rate with $f_d\tau$.

In this case, results show that the use of CCSI leads to improved BER performance. The value of the BER of the two cases approaches each other as the Doppler-Delay product grows.

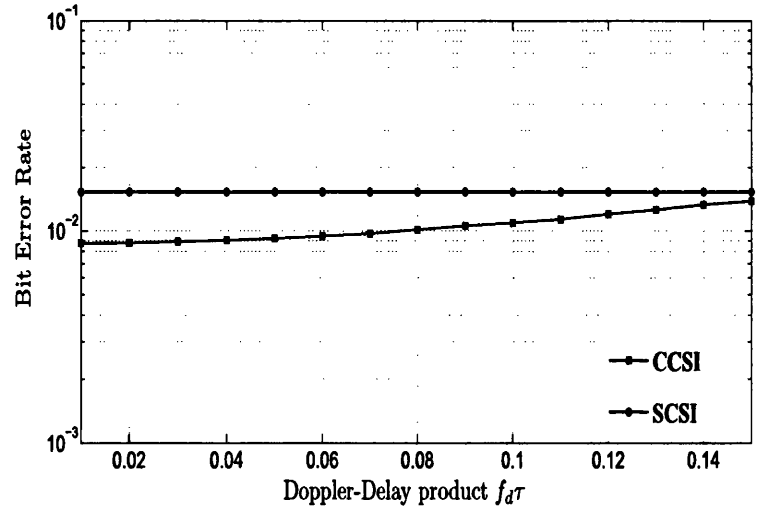


Figure 4.6: BER vs $f_d\tau$ for CCSI and SCSI based resource allocation schemes when a constraint is not imposed on the target BER.

Figure 4.7 shows throughput vs velocity curves for a CSI measurement reporting period equal to 2 TTI's. The packet arrival rate for each user equals 1 packet/user/TTI. A processing delay time of one TTI is assumed with velocities ranging from 2 km/h through 40 km/h. Beyond these speeds, there is little correlation between the outdated and current CSI for the parameters considered. When compared to SCSI, this figure shows that periodic CSI measurements lead to significant throughput gains for all of the velocities considered. In particular, a 43% throughput gain is observed when the user velocity equals 40 km/h.

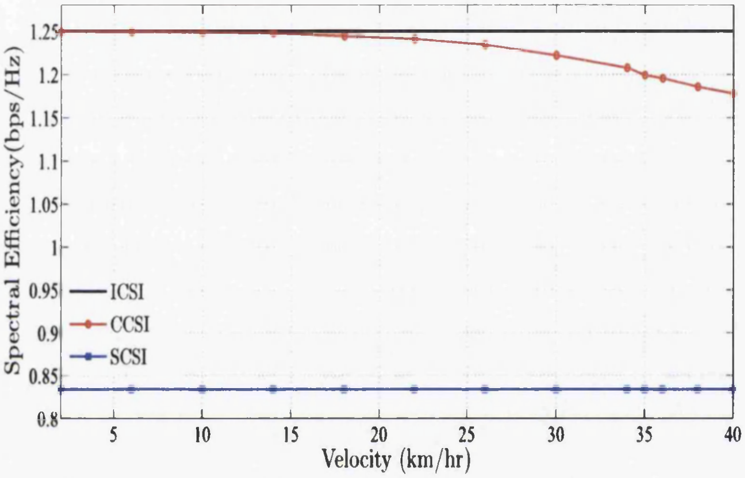


Figure 4.7: Spectral Efficiency vs velocity when ICSI, SCSI and CCSI is used to perform resource allocation (packet arrival rate=1 packet/user/TTI)

4.6.2 Providing a tradeoff between overhead and throughput/fairness

It was established in the previous subsection that when the scheduler works with CSI that is correlated with its current state, significantly higher overall system throughput can be achieved. These gains are important even at high user velocities. However, they come at the cost of increased overhead requirements. This overhead load increases with user velocity and the number of active users present in the cell. Therefore, in a fast fading environment, it is impractical to assume that all users are able to feedback their ICSI to the BS. In this section, the strategy presented in Section 3 is applied to the network under consideration.

The $K = 25$ active users present in the cell are divided into two groups according to the value of their average SNR. The first group consists of the 13 users with the highest average SNR which are named Group A users. The remaining 12 users are called Group B. For the duration of this simulation, users of Group B cannot move closer to the BS than any of the Group A users so that accurate results can be obtained. The following cases are considered for evaluation:

Case 1: Periodic CSI is received by the BS for all of the active users in the cell.

Case 2: Periodic CSI is received by the BS for only Group B users.

Case 3: 12 users (half from Group A and the other half from Group B) periodically send their CSI to the BS.

Case 4: Only Group A users send their CSI to the BS.

Case 5: No users send their CSI to the BS (only the average SNR of each active user is known by the BS).

Cases 1 and 5 are benchmark cases and present results for the CCSI and SCSI cases respectively. Case 2 is the proposed scheme which allows only the users with low average SNR's to periodically feedback their instantaneous SNR. Case 4 is similar to the scheme presented in [11] where only users whose absolute SNR exceeds a certain threshold feedback their instantaneous channel states to the BS. Case 3 is an inbetween scenario where it is assumed that the BS randomly selects an equal amount of users from each group to feedback their CSI. A similar scheme was proposed in [80] where users are grouped according to their SNR values and each group can access a given number of feedback channels via random access. However, as the BS controls the CQI feedback process here it is assumed that for the whole duration of each simulation (1000 TTIs) the same users feedback their CSI. This saves signaling bandwidth as frequently revoking and reallocating the bandwidth allocated for CQI feedback by means of radio resource control is not feasible on a fast time scale [43].

In order to quantify the degree of fairness in our simulations, Jain's fairness index is used which is defined as [81]:

$$J = \frac{(\sum_{k=1}^{K} T_k)^2}{N \sum_{k=1}^{K} T_k^2} \tag{4.9}$$

where T_k is user $k's$ average throughput. This factor measures the spread in the users' average throughputs T_k, and its value will always be within the range $1/N$ to 1 [81]. It can easily be verified that J = 1 indicates absolute fairness, whereas $J = 1/N$ indicates no fairness at all (all resources are allocated to a single user).

Figures 4.8 and 4.9 present the variation of the system's throughput and Jain's fairness factor with velocity for each of the cases considered. These figures show that when the users of Group B periodically send their CSI to the BS, higher throughput and fairness compared to the other cases (Case 3,4,5) is achieved. The employed CAC scheme has admitted 25 users into the network and the packet arrival rate equals 1 packet/user/TTI. When the BS only knows the SCSI of the Group A users , their average queue lengths is not significantly affected as can be seen from the difference between Case 2 and Case 4 in Figure 4.10. This implies that allowing Group A users to feedback their ICSI does not lead to an efficient usage of the resources allocated for overhead purposes. However, as users move further away from the BS, they require more channels to meet their data-rate requirements. By enabling the users furthest from the BS to feedback their ICSI, a higher overall system throughput can be realized. This is because a higher number of allocated channels will have been bit-loaded with CCSI. Therefore, unlike [11], where only the users whose subchannel SNR exceeds a certain threshold feedback their channel quality, these results show that when a CAC unit is employed in conjunction with a scheduler, the queue states need to also be accounted for when deciding which users should feedback their CSI to the BS. Figure 4.11 shows results for the case where only 5 users are able to feedback their CSI. Here, it can be observed that the performance of Case 2 and Case 5 (random users feedback their ICSI) is comparable in terms of throughput with Case 5 surpassing Case 2 at the higher velocities. The reason this occurs is because when fewer users feedback their CSI and the velocities are high, less network throughput is available. Therefore, if the CAC scheme accepts requests from the same users it is no longer accurate which leads to the creation of significant queue lengths for the Group A users. This is also noted in Figure 4.12. Hence, selecting which group of users will feedback their ICSI clearly depends on whether a reliable CAC scheme is employed or not. When the CAC scheme is reliable, the users with the highest SNR's do not need many subchannels to meet their QoS requirements; whereas the low SNR users will be allocated more, so that their queues will be kept within stable bounds. However, as the accuracy of the CAC scheme decreases, randomly selecting the users who feedback their CSI will lead to more channels being loaded with CCSI. If there is no CAC mechanism present and a high number of users requesting a bandwidth intensive service were admitted into the network, then Case 4 (users with high average SNR feedback their CSI) would yield the highest throughput. The absence of CAC is depicted in Figure 4.13 which shows the very high queue lengths associated with Group B users when the packet arrival rate for each of the 25 users equals 3 packets/TTI. In terms of fairness, the Jain factor remains higher for Case 2 regardless of the CAC scheme. This can be observed in Figures 4.9, 4.14 and 4.16.

All these results show that as at high velocities, a high amount of overhead is needed for the BS to work with CSI that is correlated with the current CSI; and the system-wide spectral resources available for feedback are limited, allowing only the users with the lowest average SNR to periodically send their ICSI to the BS leads to a better tradeoff between the bandwidth occupied for feedback and throughput/fairness when an accurate CAC scheme is employed. This strategy can also be applied to feedback schemes where users can only send their M highest channel gains to the BS.

It has been established, that when only a fixed number of users are allowed to feedback their ICSI over a given time period, it is useful to select the ones with the lowest average SNR. In order to further evaluate the impact of the introduced scheme, it is also compared with the idealized case presented in [40]-[42] where only the users whose scheduling metric (i.e $W_k(t)R_k$ where $W_k(t)$ is user k's queue length and R_k is the maximum throughput user k can achieve) exceeds a threshold value feedback their CQI information. The simulation parameters considered are the same as those used for Figures 4.8-4.10. It can be observed from Figure 4.17 that approximately a 10% loss in sum-capacity is realized compared to the ideal case. The realized capacity loss is less at lower values of the user velocity as more throughput can be gained through ICSI feedback (i.e lower values of the channel estimation error). In this case, the users most likely to be scheduled are the ones with the lowest average SNR as the users with a higher SNR can meet their data-rate requirements by occupying a small number of channels. However, as the velocity grows, users with high average SNR's need more channels to satisfy their data-rate requirements.

With the idealized scheme, there are important incurred penalties. To frequently revoke and reallocate the bandwidth reserved for feedback is expensive from a radio resource perspective as these are high priority signaling messages [43].

It is useful to examine how often the resources reserved for ICSI feedback will be reallocated to different users when the proposed scheme is used. Periodic feedback schemes require only one signaling message for the allocation of a CQI channel and one for its release [39]. The allocation message indicates the location and periodicity of the CQI channel slots. Once a CQI channel is allocated, the user transmits CQI feedback messages on the slots of this channel until it receives a deallocation message [39]. Therefore, using the proposed feedback reduction scheme, signaling bandwidth for the allocation/deallocation of channels reserved for ICSI feedback will be consumed when the users of Group A become users of Group B/users of Group B become the users of Group A. In order to investigate how often signaling bandwidth is consumed for the allocation/release of CQI channels , simulations were performed using the parameters of Table I. The 12 users with the lowest average SNR were only allowed to feedback their instantaneous channel states. The user velocities were uniformly distributed within a specified range: (0-10)km/hr, (0-20)km/hr, (0-40)km/hr and (0-80)km/hr. Results were averaged over 1000 different uniform user locations and random directions. For each different user location the simulations were run for the equivalent of a 1 second real-time period.

Figure 4.18 presents results for the average number of times the resources reserved for CQI feedback need to be reallocated to different users when the proposed scheme is used. It is observed, that when the user velocities are low (0-10km/hr) there is a very low probability that the resources reserved for CQI feedback will be reallocated to different users. Also, it can be seen, that even when user velocities upto 80 km/hr are considered there is still a low probability that bandwidth needs to be consumed for the allocation/release of CQI channels. Table 4.3 presents the number of times that the resources reserved for CQI feedback need to be released and reallocated to different users when the idealized case presented in [40]-[42] is considered and the user velocities are uniformly distributed within the range $0-40km/hr$. It can be seen that the resources reserved for ICSI feedback need to be allocated/reallocated approximately 3×10^4 times more often compared to the proposed scheme.

Table 4.3: Number of times resources reserved for ICSI feedback need to be allocated/deallocated for idealized and proposed schemes

Considered Scheme	No of times resources need to allocated/deallocated
Proposed scheme	0.13 (average)
Idealized scheme [40]-[42]	3120

4.7 Summary

This work shows that optimally selecting MCS levels leads to a performance enhancement when either SCSI or CCSI is used to perform resource allocation. A comparison between the SCSI and CCSI schemes shows that the use of CCSI leads to important throughput gains even under significant user mobility. Since in a fast fading environment, excessive overhead is required for the scheduler to continuously work with CCSI, a strategy which leads to optimal usage of the resources reserved for feedback purposes was presented. Simulation results showed that this strategy leads to a higher overall fairness and system throughput when ICSI feedback is confined to a specific set of users and CAC functionality is considered.

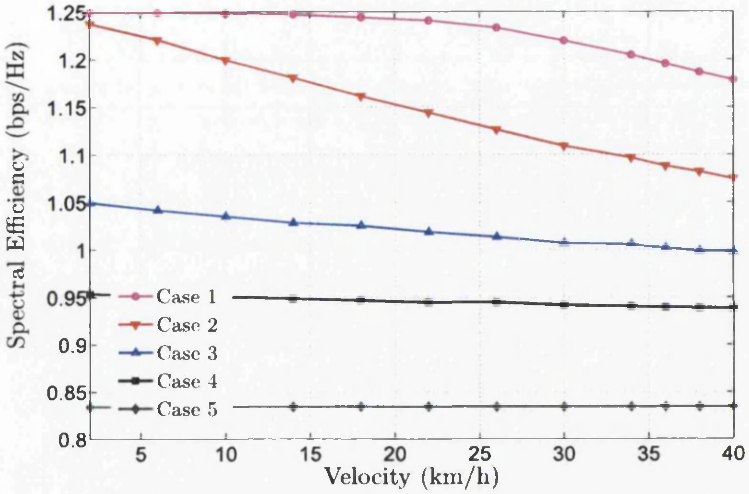


Figure 4.8: Spectral Efficiency vs user velocity for the 5 different cases considered (packet arrival rate =1 packet/user/TTI)

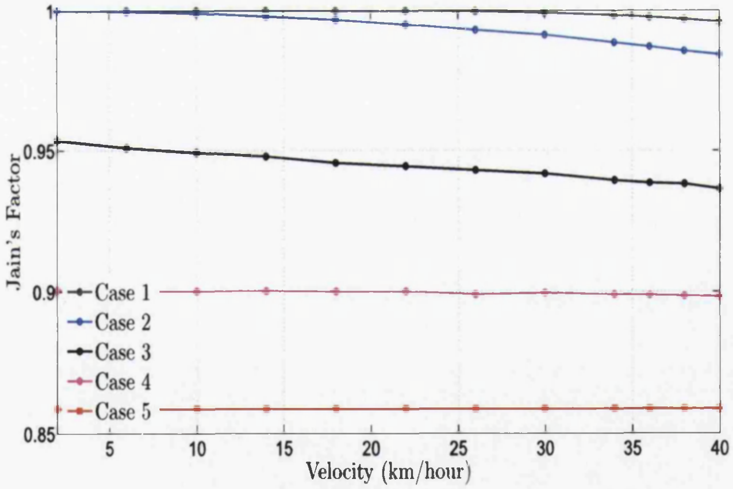


Figure 4.9: Jain Factor vs user velocity for the 5 different cases considered (packet arrival rate =1 packet/user/TTI)

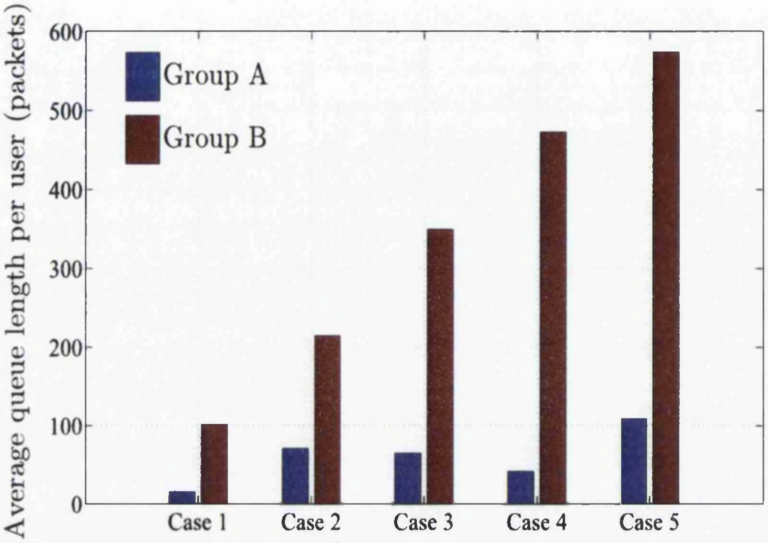


Figure 4.10: Average Queue length vs velocity for the 5 different cases considered (user velocity=40km/h)

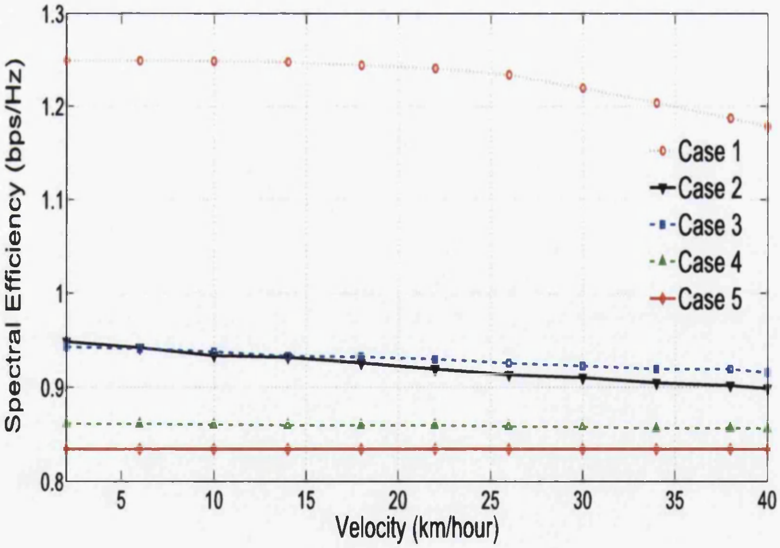


Figure 4.11: Spectral Efficiency vs velocity for each different case when 5 users feedback CSI

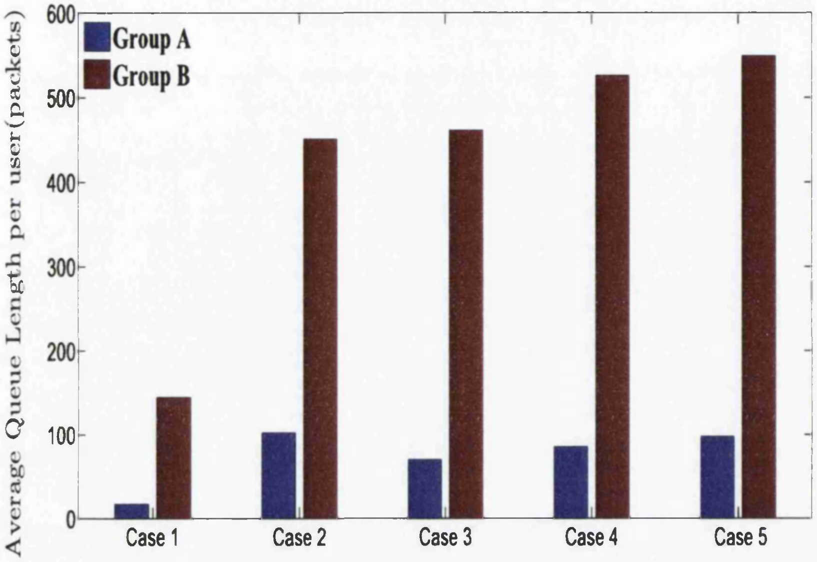


Figure 4.12: Average queue length for each different case when 5 users feedback CSI (user velocity=40km/h)

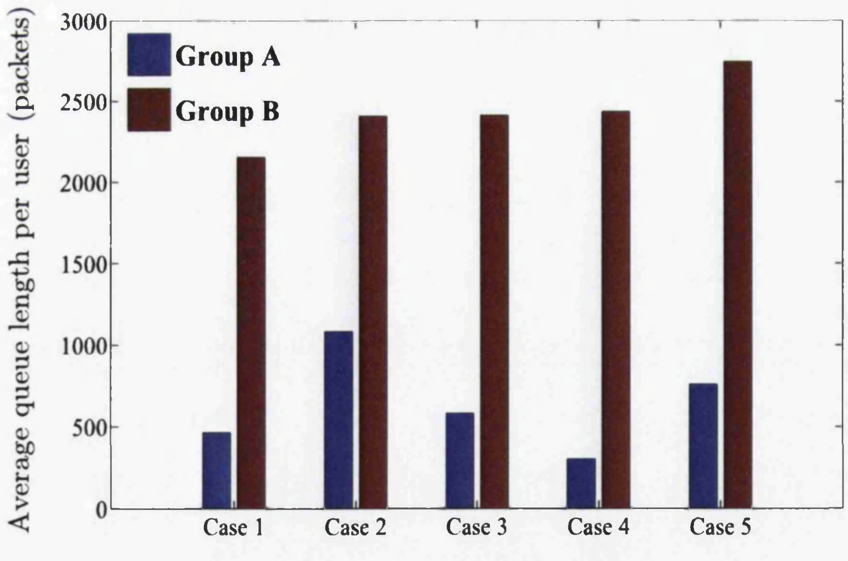


Figure 4.13: Average queue length when CAC unit not employed (packet arrival rate =3 packet/user/TTI, user velocity= 40km/hour)

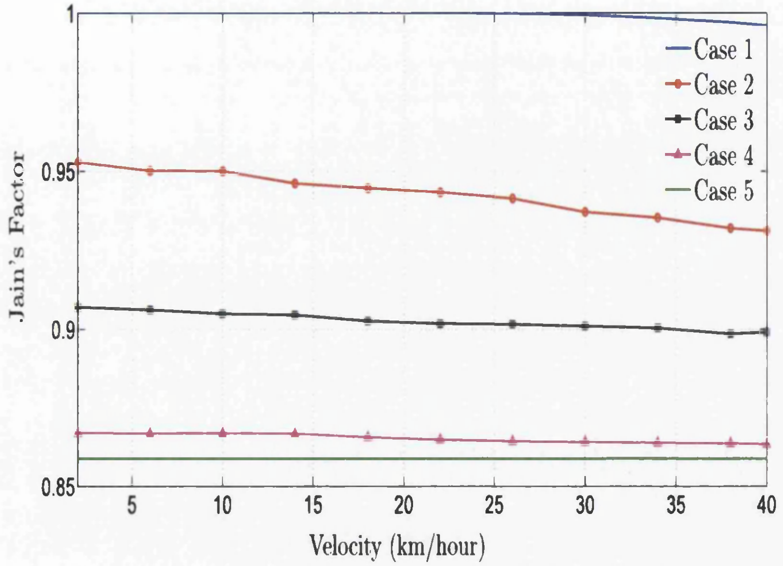


Figure 4.14: Jain factor vs velocity for each different case when 5 users feedback CSI

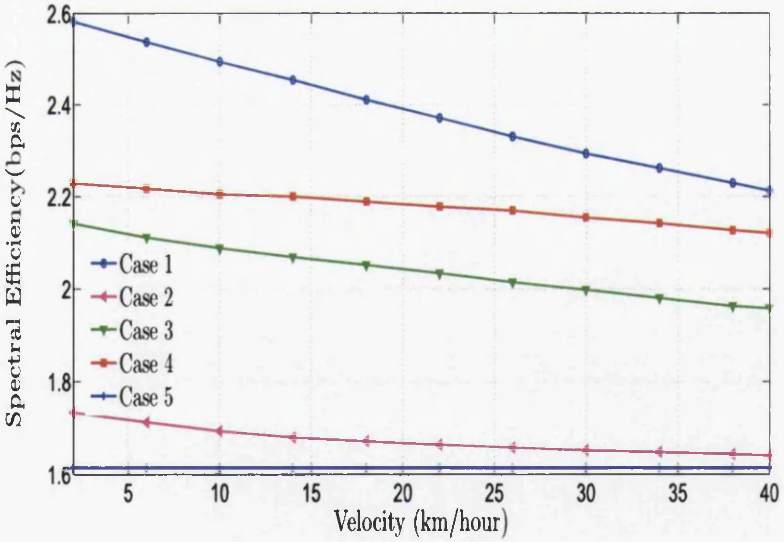


Figure 4.15: Spectral Efficiency vs velocity when CAC unit not employed (packet arrival rate =3 packets/user/TTI)

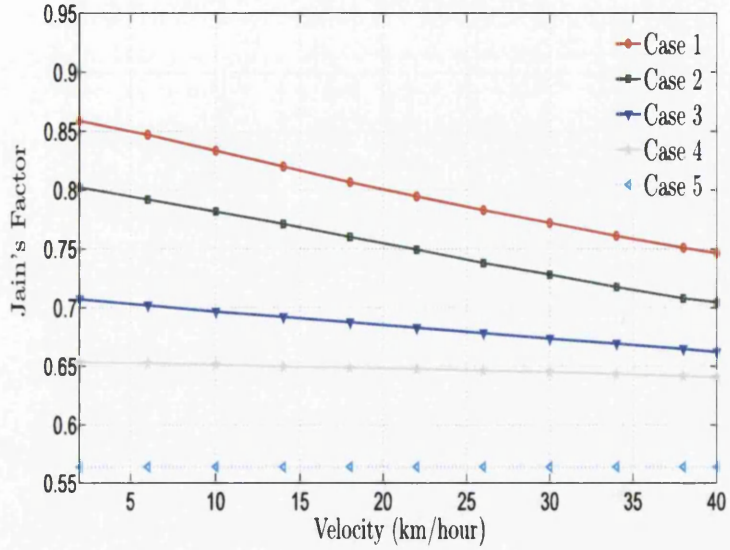


Figure 4.16: Jain Factor vs velocity when CAC unit not employed (packet arrival rate =3 packets/user/TTI)

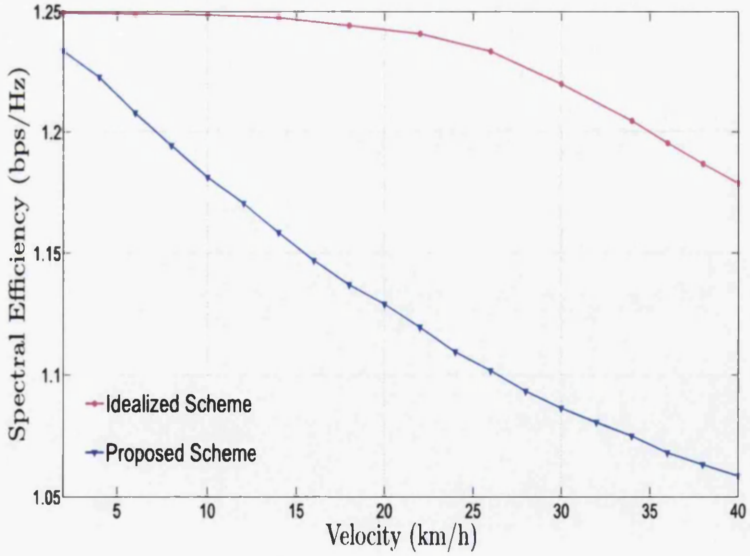


Figure 4.17: Spectral Efficiency vs velocity for idealized and proposed scheme

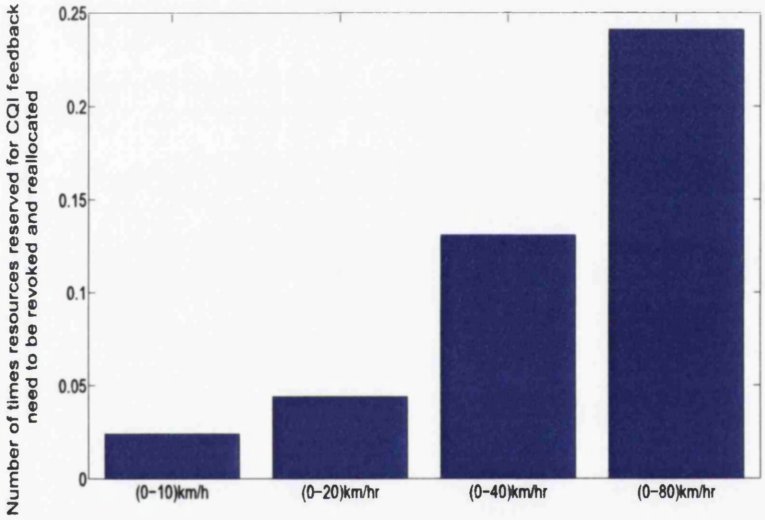


Figure 4.18: Average number of times resources reserved for feedback are reallocated for different user velocities

Chapter 5

A novel RRM policy for mobile WiMAX networks

In Chapter 4, the impact of a CAC unit on practical resource allocation strategies for the OFDMA downlink was discussed. In this chapter, a mechanism for resource allocation and CAC is proposed using the cost of each service and the user channel gains to partition the available bandwidth in accordance with a complete partitioning structure. The focus is on mobile WiMAX networks.

5.1 Introduction

In BWA networks, a multitide of different services need to be supported, and it becomes very important to effectively manage the available bandwidth so that QoS can be provided to each service. In a mobile WiMAX network environment, a variety of applications can be supported which can be generally split into five groups: 1) unsolicited grant service (UGS); 2) real-time polling service (rtPS);3) extended real-time polling service (ertPS); 4) non-real time polling service (nrtPS); and 5) best effort (BE) services [13].

UGS is designed to support real-time services that generate fixed sized data packets on a periodic basis (e.g. VoIP). The rtPS service is fitted to support real-time services that generate variable sized data packets on a periodic basis (e.g. video conference, MPEG, etc.). The nrtPS service is similar to rtPS but is delay tolerant; hence, only a minimum data rate is specified. ertPS is a scheduling mechanism that builds on the efficiency of both UGS and rtPS. The ertPS is designed for realtime traffic with a variable data rate (such as voIP service with silence suppression). Best effort is designed to support a data transmission when no minimum service level is required [13].

The goal of any CAC scheme is to avoid the network congestion and maintain the delivered QoS to different connections by means of restricting the number of ongoing connections in the system or rejecting the new connection request. Well known admission control schemes are: 1)the complete sharing scheme and 2)the complete partitioning scheme [82]. The main concept of the complete sharing scheme is that many users with different QoS requirements share a common resource [82]. If the required resource is satisfied, the call will be admitted into the system; otherwise, it will be rejected [82]. In [83]-[85], the authors proposed a CS resource allocation scheme in making admission control decisions. The drawback of this scheme is that it treats each call equally and does not consider the priorities of different calls. This will lead to users of all traffic classes having equal blocking probabilities. Consequently, the QoS of each service cannot be easily controlled. The CP scheme [84]-[85] ensures the fairness of different priority calls. When a CP scheme is employed, new connection requests are rejected when the resource allocated to the corresponding service type is used up [82].

The support of multiple services in OFDMA based wireless networks implies that the aggregate rate allocated to users of each traffic class should not exceed a partition of the throughput (CP scheme) [86]. Limiting the aggregate rate that a group of demanding subscribers receives imposes fairness and guarantees QoS [70],[86] . In addition, the amount of resource required to service the arriving traffic load depends on the channel gains (user SNR). Higher user channel gains will lead to fewer subchannels/subcarriers needed to service the arriving traffic load. Likewise, lower channel gains require more subcarriers/subchannels to service the same amount of traffic load. Therefore, it is important that the CAC mechanism takes this into account so that the admitted traffic load is neither 1) too high which would lead to unstable user queues, nor 2) too low: which leads to inability of the scheduler to utilize the inherent frequency diversity.

In this Chapter, a mobile WiMAX environment is considered where the arriving traffic load is higher than the bandwidth capacity. A method to calculate the blocking probability of each user under the assumption of a CS policy is first developed. These values are then varied according to the price of each traffic class. These values are then used in a CP policy that aims to optimize the operator revenue under fairness constraints. Numerical results will demonstrate that the pricing model can play a key role in ensuring that traffic classes requiring guaranteed QoS can obtain enough channels to satisfy their user requirements even under conditions of extreme network overload.

The rest of this chapter is organized as follows: the system model is presented in Section 5.2. The methodology used to calculate the value of the blocking probability when all users of each traffic class are equally blocked is discussed in Section 5.3. The method used to prioritize users of different traffic classes is also described here. Existing CAC optimization strategies are discussed in Section 5.4. The CP structured policy and the heuristic algorithm used to partition the bandwidth are presented in Section 5.5. Simulation results are presented in Section 5.6 followed by the conclusions in Section 5.7.

5.2 System Model

A downlink OFDM system with N available subcarriers is considered here. It is assumed that M classes of traffic share the available bandwidth. Connections for each traffic class arrive at a rate equal to λ_i and each connection has a holding time equal to $\frac{1}{\mu_i}$. The data rate requirement of each traffic class i is equal to R_i , and each class has a BER requirement which is denoted by BER_m. Adaptive M-QAM modulation with modulation levels 0,2,4,6,... is employed at the transmitter, whereas the channels cannot be shared between the users. Equal power allocation accross all of the subcarriers is assumed.

The channels are described by Rayleigh fading which follow the ITU Pedestrian A model. The WiMAX Forum has recommended that both the ITU Pedestrian B and Vehicular A models are used to model the channels in the 802.16e environment [87].

5.3 Methodology

In this chapter, a cost-based CP policy is developed for a mobile WiMAX environment. It is assumed that the network is overloaded. The bandwidth is partitioned through the use of a blocking probability threshold value which differs according to the traffic class. In order to calculate this value, two phases are required. Initially, absolute fairness is assumed (CS scheme). In this case, users of each traffic class are equally likely to be blocked. The methodology used to calculate the blocking probability of each traffic class is calculated in subsection 5.3.1. Using this value, thresholds for the blocking probability of each traffic class are obtained in subsection 5.3.2.

5.3.1 Absolute Fairness (Complete Sharing Policy)

When the incoming traffic is greater than what can be serviced, a number of users need to be denied access due to a lack of spectral resources. According to [47], in order to achieve absolute fairness (AF) at the CAC level, each traffic class is given the same value of the blocking probability. This means that all users, regardless of their traffic class, have an equal chance to access the bandwidth. In the scenario of a stressfull network where the arriving traffic load is higher than the capacity of the network, the blocking probability of each traffic class is denoted by Bl^{AF} and is given by the following equation:

$$Bl^{AF} = 1 - \frac{B_{av}}{B_{req}}, \tag{5.1}$$

where B_{av} is the available bandwidth, and B_{req} is the requested bandwidth. In a cellular OFDMA environment, B_{av} is the number of available subcarriers; whereas B_{req} is the number of subcarriers needed to fulfil all of the user requirements. When the value of the requested bandwidth (B_{req}) exceeds the available bandwidth (B_{av}), some of the users have to be blocked by the CAC unit.

In a cellular environment utilizing AMC, the value of the blocking probability is a function of the user SNR. For example, when users are in generalcx located close to the BS, then the value of Bl^{AF} will be lower (as higher order MCS levels can be used) than when they are far; in which case, more subcarriers are needed to satisfy the users. Therefore, in a cellular environment it is convenient to express this value in terms of requested subcarriers.

5.3.2 Calculation of Bl^{AF}

Initially, the bandwidth required to satisfy the ongoing connections needs to be calculated. In order to calculate this value the following assumptions are made:

1) Each user terminal reports one CQI value for the whole system bandwidth (wideband CQI feedback [43]).

2) Each subcarrier/subchannel can only be occupied by one user.

3) All users are subject to the same type of fading.

According to [88], if adaptive M-QAM is employed at the transmitter, the number of subcarriers c_k needed by a user k to satisfy his data-rate requirements is given as follows:

$$c_k = \left\lceil \frac{R_k}{\frac{B}{N} * f(\log(1 + \frac{-1.5}{\ln(BER_m)}\bar{\gamma}_k)} \right\rceil , \tag{5.2}$$

where B is the total available bandwidth (MHz), N is the total number of available subcarriers/subchannels, BER_m is the BER requirement of each service of traffic class m,

R_k is user k's average data-rate requirement and $\bar{\gamma}_k$ is the average SNR of each user k. In Eq. 5.2 $f(x)$ is:

$$f(x) = 2\left\lfloor \frac{1}{2}x \right\rfloor. \tag{5.3}$$

If there are K_1 ongoing connections accross all traffic classes then the total bandwidth needed to satisfy all ongoing connections is denoted by B_{old} and is given as:

$$B_{old} = \sum_{k=1}^{K_1} c_k. \tag{5.4}$$

In order to calculate the value of the bandwidth needed to satisfy the requested number of new connections B_{new} the same assumptions are made. When there are K_2 new users requesting access to M different traffic classes the value of B_{new} is given as:

$$B_{new} = \sum_{k=1}^{K_2} c_k. \tag{5.5}$$

Therefore using the value of $B_{req} = B_{old} + B_{new}$, Bl^{AF} can now be calculated.

Figure 5.2 shows the variation of the average value of Bl^{AF} with the connection arrival rate. All users considered have an average data-rate requirement equal to 512 kbps, whereas users request access to services of different priority even though their data-rate requirements are equal. The average user SNR equals 15 dB. Two values of the BER requirement considered are 10^{-4} and 10^{-6}. From this figure it can be observed that the value of the Bl^{AF} increases significantly with the connection arrival rate. The absolute fairness policy is equivalent to the CS CAC policy where all calls are treated equally and share a common resource. However, this policy is in essence unfair since it treats each traffic class equally and does not prioritize services.

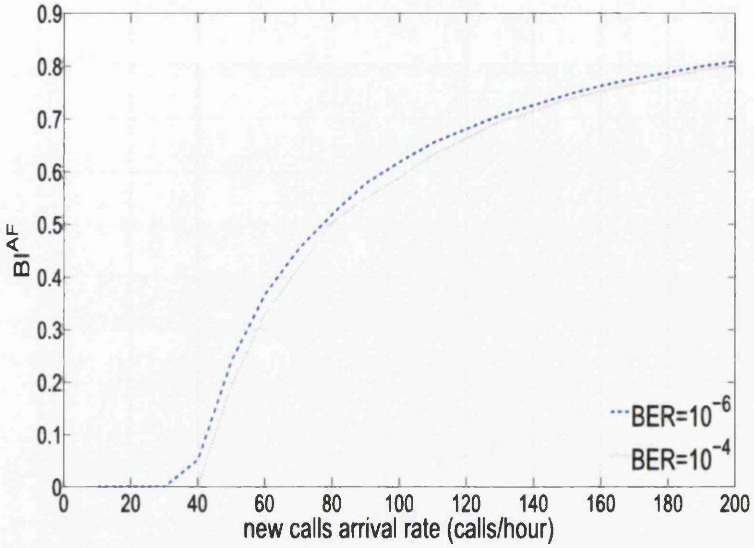


Figure 5.1: Absolute Fairness Blocking Probability versus new call arrival rate

5.3.3 Cost based Connection Admission Control

In this section, the value of the blocking probability calculated in the previous section is modified. With the CS scheme, all new users have equal blocking probabilities regardless of their traffic class. However, in a multiservice OFDMA cellular network the QoS needs to be guaranteed for voice services even when the network is extremely stressed, whereas the more expensive rtPS and ertPS applications should be allocated more bandwidth. Therefore, CS is not a convenient policy. Here, the blocking probability of each traffic class is a function of its cost. It is assumed that the price per bit of prioritized traffic classes such as UGS is more expensive than the price per bit of traffic classes such as Best Effort. Firstly, the value of the blocking probability calculated in the previous section needs to be modified so that it can account for the cost of each service. his allows prioritization of certain traffic classes. The threshold value for each traffic class m must satisfy the following condition:

$$Bl^{AF} = \frac{\sum_{m=1}^{M} Bl_{th}(m)}{M} \tag{5.6}$$

Under this assumption, a threshold for the blocking probability of each traffic type can be set as:

$$Bl_{th}(m) = Bl_{th}^{AF} * (1 - \frac{S_m - \bar{S_m}}{\bar{S_m}}), \tag{5.7}$$

where S_m is the cost per bit for traffic class m, M is the total number of traffic classes, and $\bar{S_m}$ is the average bit price. This equation ensures that the overall value of the blocking probability equals Bl^{AF} when all traffic classes have an equal price per bit. The value of $Bl_{th}(m)$ will increase if the cost per bit for traffic class m is less than the average bit price. This increase, is proportional to the difference between the average bit price and S_m. Likewise, $Bl_{th}(m)$ will decrease if the cost per bit for traffic class m is greater than the average bit price of the traffic classes. Therefore, the application of this formula will allocate more bandwidth to users requesting more expensive services and less to those that require the cheaper ones. When the following holds:

$$S_m \leq 2 * \bar{S}_m. \tag{5.8}$$

the value of $Bl_{th}(m)$ will always equal zero regardless of the value of Bl^{AF}. This means that when the network is stressed, the blocking probability of this traffic class will be equal to one. Also, when

$$S_m \geq 2 * \bar{S}_m \tag{5.9}$$

the blocking probability of traffic class m will equal zero.

5.4 Existing CAC optimization strategies

Consider M different classes of traffic sharing N subcarriers. The bandwidth requirement vector can be defined as $\vec{b} = (b_1, b_2, ..., b_M)$, and the system state vector as $\vec{n} = (n_1, n_2, ..., n_M)$; where n_i is the number of class i connections in the system. In the context of a wireless system, b_i is the average minimum number of subcarriers required to satisfy a connection of traffic class i. Based on these parameters, Ω_{CS} is defined as the set of all possible system states that can be expressed as $\Omega_{CS} = \vec{n}|\vec{n}\vec{b} \leq B$ [47]. The subscript CS stands for complete sharing. A CAC policy can be defined as any arbitrary subset of Ω_{CS}. Given Ω, a connection request will be accepted when the system state vector remains in Ω after the connection is accepted.

5.4.1 Optimal Revenue Strategies

The long term average revenue of a particular CAC policy can be calculated as [47]:

$$R(\Omega) = \sum_{\vec{n} \in \Omega} \vec{n}\vec{r}P_{\Omega}(\vec{n}), \tag{5.10}$$

where $P_{\Omega}(\vec{n})$ is the steady state probability that the system is in state $\vec{n}$ and $\vec{r} = r_1, r_2, ..., r_M$; where $r_i = S_i b_i$ is the corresponding average revenue generated by accepting a class i connection. When the arrival and service processes are both memoryless, then $P_{\Omega}(\vec{n})$ can be calculated as [83]:

$$P_{\Omega}(\vec{n}) = \frac{1}{G(\Omega)} \prod_{i=1}^{M} \frac{\rho_i^{n_i}}{n_i!}, \quad \vec{n} \in \Omega, \tag{5.11}$$

where $G(\Omega)$=$\sum_{\vec{n} \in \Omega} \prod_{i=1}^{M} \frac{\rho_i^{n_i}}{n_i!}$ and $\rho_i = \frac{\lambda_i}{\mu i}$. The blocking probability of class i traffic is [47]:

$$Pb_i(\Omega) = \frac{G(\Omega_i^b)}{G(\Omega)}, \tag{5.12}$$

where $\Omega_i^b = \vec{n}|\vec{n} \in \Omega$ and $\vec{n} + \vec{e_i} \notin \Omega$. Here, $\vec{e_i}$ is a M-dimensional vector of all zeros, except for its i_{th} element.

5.5 CP structured policy and Algorithm used to partition the bandwidth

The objective of CAC optimization can be chosen to satisfy either service providers or subscribers. Service providers wish to optimize their revenues whereas subscribers prefer fairness and throughput. Here, a fairness constrained optimal revenue resource allocation policy is studied. In particular, the aim is to maximize the revenue of the service provider during network congestion, while ensuring that the $Bl_{th}(m)$ threshold values for each traffic class are met. This satisfies both users and service providers since a) the revenue is maximized, and b)users who pay more for a particular service are prioritized. The blocking probability threshold for each traffic class should be greater than the values of $Bl_{th}(m)$ derived using the costs of each traffic class in Section 5.3 which is denoted by $Bl_{lb}(m)$. If this is not the case, no CAC policy can be found to meet the fairness constraints [47]. Therefore, $Bl_{lb}(m)$ is subject to the follwing relationship: $Bl_{lb}(m)$ $<Bl_{th}(m) <1$. Correspondingly, the normalized blocking probability threshold is pb^{th}, and its value ranges from zero to one. The relationship between $Bl_{th}(m)$ and pb^{th} is given as: $Bl_{th}(m) = (1 - Bl_{th_{lb}}(m))pb^{th} + Bl_{th_{lb}}(m)$.

5.5.1 CP structured Admission Control Policy

In order to specify this policy, the brute force searching method is a straighforward technique. However, it requires calculating the long term average revenue and the blocking probabilities using Eq. 5.11 and Eq. 5.12. As the brute force searching method has unbearable complexity and is very difficult to implement, the CP structure has become particularly popular amongst many researchers [47]. The CP structured admission control policy allocates each traffic class a certain amount of non-overlapped bandwidth resource [47]. Therefore, the blocking rate of one traffic class will not influence the blocking rate of others [47]. Hence, as there are M traffic classes, a CP policy can be split into M independent sub-policies [47]. In this work, the CP policy separates the total number of subcarriers into M non-overlapped parts which are denoted by N_{CP}^{m}, where N_{CP}^{m} belongs to traffic class m. For a given CP policy, the m_{th} sub-policy can be modelled by a $M/M/N/N$ queuing system [47] in which the total number of servers s_m equals N_m/n_m where n_m is the minimum number of subcarriers needed to satisfy a user of traffic class m and N_m is the total number of subcarriers allocated to traffic class m. The value of n_m is given by:

$$n_m = \left\lceil \frac{R_m}{B_n * l_m} \right\rceil, \tag{5.13}$$

where R_m is the data-rate requirement of traffic type m , B_n is the width of a subcarrier, and l_m is the average spectral efficiency the users of traffic class m can achieve. If there are k_m users requesting access to this class, then the value of l_m is given by the following formula:

$$l_m = \frac{\sum_{i=1}^{k_m} f(\log(1 + \frac{-1.5}{\ln(BER_m)}\bar{\gamma}_i))}{k_m}. \tag{5.14}$$

According to Eq 5.11 and Eq. 5.12, the overall long-term average revenue of the CP policy is:

$$R_m = \sum_{m=1}^{M} \sum_{j=0}^{s_m} S(m)j \frac{\frac{\rho_i^j}{j!}}{\sum_{k=0}^{s_m} \frac{\rho_i^k}{k!}}, \tag{5.15}$$

where $\rho_i = \frac{\lambda_i}{\mu_i}$ is the offered Erlang load of traffic class i and $S(m)$ is the revenue per server. Furthermore, the blocking probability of class m traffic is given by the Erlang B formula and is [89]:

$$Pb_m(CP) = B(s_m, \rho_m) = \frac{\frac{\rho_m^{s_m}}{s_m}}{\sum_{k=0}^{s_m} \frac{\rho_m^k}{k!}} \tag{5.16}$$

It is noted that the Erlang B formula can be computed using the following recursion:

$$B(s_m, \rho_m) = \frac{\rho_m B(s_m, \rho_m)}{s_m + 1 + \rho_m B(s_m, \rho_m)} \tag{5.17}$$

where $B(0, \rho_{(m)})$=1. These discussions have shown that the optimal CP problem is to find the best bandwidth partitioning scheme.

5.5.2 Fairness Constrained Greedy Revenue Algorithm

The revenue of the accepted traffic can be calculated by subtracting the revenue of arriving traffic from the revenue of rejected traffic. It is assumed that class m traffic is assigned $j = N_m/n_m$ servers. According to the theory of marginal economic analysis [90], the corresponding revenue can be given as:

$$R_i(CP) = r_i\rho_i - r_i\rho_i B(j, \rho_i)[B(j-1, \rho_i) - B)j, \rho_i] \tag{5.18}$$

If the server number is reduced to $j - 1$, the following holds:

$$R_i(CP) = r_i\rho_i - r_i\rho_i B(j-1, \rho_i). \tag{5.19}$$

Therefore, the revenue brought by the j_{th} server is written as:

$$R_i(j, \rho_i) = R_i(CP)|_{B_{CP}=jb_i} - R_i(CP)|_{B_{CP}=(j-1)b_i} = r_i\rho_i[B(j-1, \rho_i) - B(j, \rho_i)]. \tag{5.20}$$

5.5.3 Heuristic Algorithm used

The above discussions have shown that the optimal CP problem is to find the best bandwidth partitioning scheme. Here, the heuristic used for the fairness-constrained optimal revenue CP policy is described. In order to partition the bandwidth, a variation of the greedy heuristic algorithm proposed in [47] to solve their fairness constrained greedy revenue problem is applied to wireless networks. As this algorithm does not guarantee that the QoS of the ongoing connections will be satisfied a second stage is then proposed. The algorithm proposed in [47] proceeds in two phases. In the first phase, it allocates each traffic class enough subcarriers until the value of the Erlang blocking probability drops below the value of the blocking probability threshold for each traffic class. This process ensures that the required grade of service will be provided at all times. This condition is fulfilled step by step, beginning with the most expensive class. Then, any remaining bandwidth is allocated in a manner that will optimize the overall revenue.

Algorithm 1 Algorithm used to optimally partition the bandwidth so that blocking probability thresholds can be met

```
$Input : N(i) = 0; Pb(m) = 1; s(i) = 0$
$Input : Bl_{th}(m); N_{free} = N_{total}$
Phase 1 : Sort traffic classesin ascending order using
the cost and allocate servers until $Pb(m)$ drops
below $Bl_{th}(m)$
for i=1:M do
  while $Pb(i) > Bl_{th}(m)$ do
    $Nfree = Nfree - n(m)$
    $N(m) = N(m) + n(m)$
    $Pb(i) = B(si, \rho(i));$
    si=si+1;
  end while
end for
Phase 2 : Allocate remaining free bandwidth
according to the optimal revenue strategy
for i=1:M do
  while $N(m) < n(m) * \rho(m), Nfree - n(m) >= 0$ do
    $N(m) = N(m) + n(m)$
    $Nfree = Nfree - n(m)$
  end while
end for
```

This algorithm outputs the values of $N(m)$ which are the total number of subcarriers allocated to traffic class m.

5.5.4 Process used to ensure ongoing connections are not dropped

The values of $N(m)$ which the previous algorithm outputs do not ensure that all ongoing connections can be satisfied. Therefore, a further step is proposed which ensures that no existing connections are dropped . In order for this step to be executed, the values of the bandwidth required to satisfy the ongoing connections of each traffic class are required. As the number of ongoing connections of each traffic class are known, these values can be calculated using Eq 5.2 and Eq 5.4. For convenience they are denoted by $B(1), B(2), ..., B(m)$ where $B(m)$ is the number of subcarriers that can satsify all ongoing connections of traffic class m. Furthermore, as with the case of the previous algorithm, the M traffic classes are ordered in order of priority (i.e traffic class 1 is the most expensive class and traffic class M is the least).

Algorithm 2 ensures that the ongoing connections can be satisfied. Initially, the algorithm checks to see whether the bandwidth required to satisfy all ongoing connections exceeds the system's available bandwidth N_{total} (i.e users who were close to the BS when admitted into the network have now moved to the cell edge). In this case, it is assumed that the MAC scheduler can cope with a short period of overloading with a graceful performance degradation that will not be too noticeable to users [58]. When the bandwidth required to satisfy all ongoing connections does not exceed the system's available bandwidth, the algorithm checks to see whether the number of channels allocated to a traffic class $N(m)$ exceeds the number of channels needed to satisfy its ongoing connections $B(m)$. If it does, then this traffic class does not need to be allocated more bandwidth to satisfy the requirements of the ongoing connections. However if the allocated bandwidth is less than what is required by the ongoing connections the algorithm removes bandwidth/carriers from other traffic classes as long as the ongoing connections of those traffic classes can be satisfied. Carriers are first removed from the traffic class that offers the least revenue. If there is still not enough bandwidth to satisfy the ongoing connections, the algorithm will then remove bandwidth from the traffic class that offers the next least amount of revenue. This process is repeated up to the most expensive class until all existing connections of the traffic class under consideration have been satisfied.

Algorithm 2 Ensure ongoing connections can be satisfied

```
Require: $N(m) \geq B(m) \forall m$
  $jj = M$
  for ii = 1 to M do
    $Needed = B(ii) - N(ii)$
    { Calculates how much bandwidth is needed to satisfy the ongoing connections}
    while Needed > 0 & jj > 0 do
      $Transfer = min(N(jj) - B(jj), Needed)$
      { Transfers bandwidth as long as data-rate requirements of class $jj$ are not violated}
      if Transfer <= 0 then
        $jj = jj - 1$
        {continue}
      end if
      $N(ii) = N(ii) + Transfer$
      $N(jj) = N(jj) - Transfer$
      $Needed = Needed - Transfer$
      $jj = jj - 1$
    end while
    if jj == 0 then
      { break}
    end if
  end for
```

5.5.5 Complexity of heuristic algorithm

The first phase of this algorithm has a complexity of $O(N)$. Also, there are $O(N)$ iterations in the second phase of the algorithm. For each iteration, the algorithm searches through M system states to locate the traffic class that leads to the maximum possible revenue. Therefore the complexity of the second phase equals NM. By combining phases 1 and 2, it can be concluded that this algorithm has a complexity of $O(NM)$. The algorithm that ensures the ongoing connections are satisfied also has a linear complexity ($O(M)$).

5.6 Simulation Results and Discussion

In this section, the impact of the cost on the developed CP CAC scheme is investigated. In the first subsection, the significance of increasing the cost of a UGS bit is presented. Afterwards, the performance of the bandwidth partitioning scheme under different user SNR conditions is discussed. In order to evaluate the performance of the scheme, a bandwidth sufficiency factor is introduced. It is defined as the ratio of the number of subcarriers allocated to the connection requests of traffic class m to the number of subcarriers that are needed to service all users of that traffic class. Five different traffic classes including two different rtPS classes are considered. The number of traffic classes and their minimum data rate requirements needed so that the connections can be satisfied are listed in Table 5.1 as are the number of users requesting access to services in the cell. The number of users requesting access to traffic class i is assumed to equal $\left\lceil \frac{\lambda_i}{\mu_i} \right\rceil$ where λ_i is the arrival rate of connections for traffic class i during the period of network overload and $\frac{1}{\mu_i}$ is the average connection holding time. A bandwidth of 10 MHz and 1024 subcarriers were assumed for all simulations. The average user SNR for the results presented in Section 5.6.1 was set to equal $12dB$. The value of pb^{th} is set to 0.05.

Table 5.1: Simulation Parameters Used to evaluate bandwidth paritioning scheme

Traffic Class	User no	data-rate requirement	Target BER
UGS	11	64 kbps	10^{-4}
rtPS	2	5 Mbps	10^{-6}
rtPS1	1	3 Mbps	10^{-6}
nrtPS	50	30 kbps	10^{-6}
BE	50	10 kbps	10^{-6}

Due to the characteristic of delay-tolerant services, when making CAC decision, only minimal traffic rates need to be satisfied in order for BE and nrtPS services to be accepted.

5.6.1 Impact of the cost

This section shows that during the period of network overload a large amount of UGS users can be satisfied in the network if the price per bit of the UGS traffic class grows. It is useful to investigate the significance of varying the price per bit of each traffic class on the efficiency of the bandwidth partitioning scheme. For Figures 5.2 and 5.3, the price per bit for each considered traffic class is given in Table 5.2. In Figure 5.3, the impact of varying the price of a UGS bit for under a constant Bl^{AF} equal to 0.92 is investigated.

Table 5.2: Cost of Each Service

Traffic Class	Cost Per Bit
UGS	variable
rtPS	0.3
rtPS1	0.28
nrtPS	0.25
BE	0.2

It is evident that as the cost of this traffic class increases, more subcarriers are allocated to it. In particular, when the price of a UGS bit exceeds 0.4, all the subscribers of this class can be satisfied. Therefore, even though under an absolute fairness or CS CAC scheme, the blocking probability of the UGS traffic class would have equalled 92%; increasing the cost of this traffic class ensures that all of the UGS offered traffic load can be serviced. When the price of a UGS bit is closer to the cost of the rtPS1 class, then 18% less users can be satisfied. Figure 5.4 shows curves of the bandwidth sufficiency factor vs the blocking probability experienced by the UGS users if the absolute fairness (CS) scheme is employed. In order to vary Bl^{AF}, integer multiples of the number of users given in Table 5.1 are used. A close observation of this figure indicates that even when the traffic intensity of the connections is very high, a significant proportion of the average offered UGS traffic load can be serviced. As the price of the UGS traffic class grows, this value clearly increases. In the case of WiMAX, the number of slots allocated to the UGS service is fixed [13] due to the stringent delay requirements associated with the class. Varying the cost of UGS ensures that a specific portion of the frame is always allocated to it. The higher the cost, the greater the portion of that frame. This will always lead to guaranteed QoS for the UGS users, regardless of the amount of network overload. However, under a CS CAC scheme, no services can be prioritized.

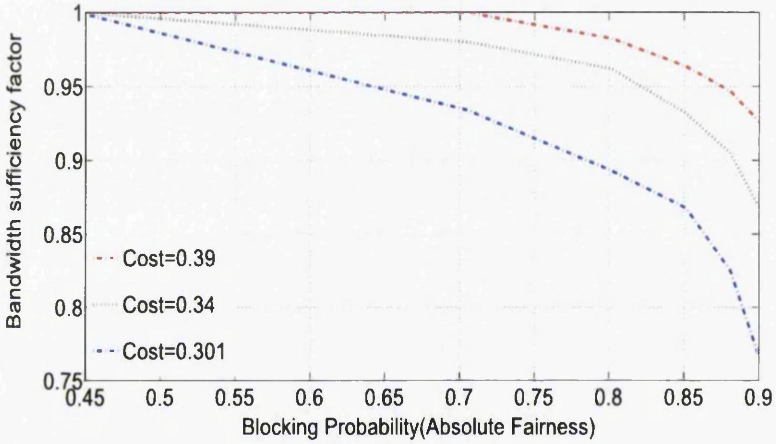


Figure 5.2: Impact of UGS price on bandwidth sufficiency factor

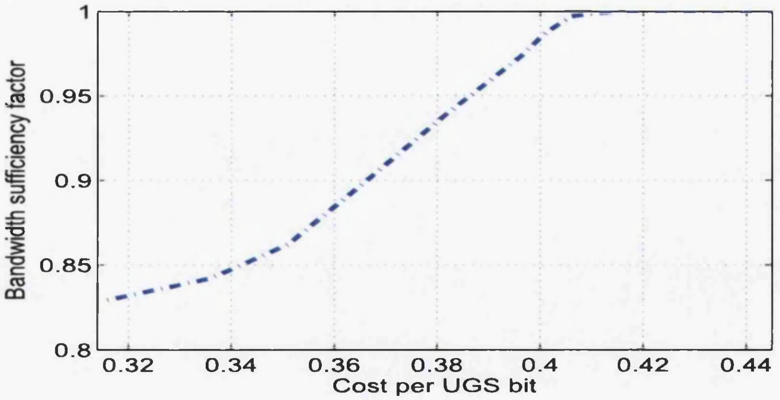
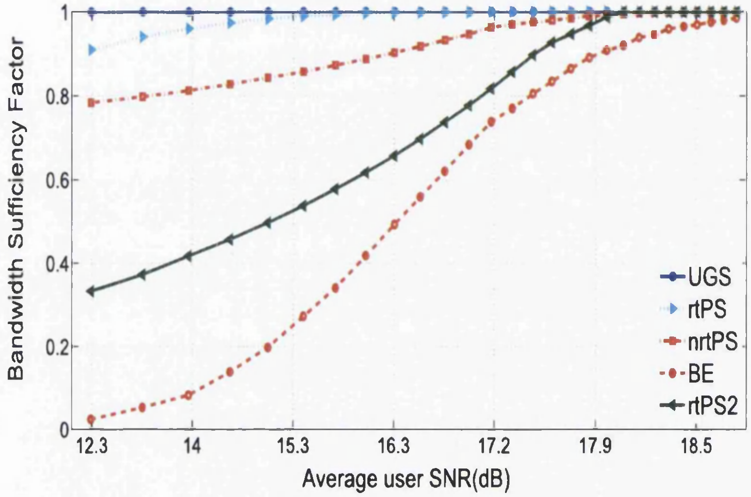


Figure 5.3: Impact of UGS price on bandwidth sufficiency factor

5.6.2 Impact of user SNR

The impact of the user SNR's on the efficiency of the bandwidth partitioning scheme is investigated here. Figure 5.4 shows the variation of the blocking probability of each traffic class with the average user SNR. The parameters of Table 5.1 are considered. Both Figures 5.2 and 5.3 depict that when the cost of the UGS traffic class exceeds 0.4, all users have access to bandwidth; therefore, the price of a UGS bit is set to equal 0.45 in order to ensure that these connections have guaranteed QoS. The price of all other traffic classes follows Table 5.2. A close observation of Figure 5.4 indicates that a higher number of users from all traffic classes can be satisfied as the average user SNR grows. It can also be seen that the number of subcarriers allocated to UGS is sufficient, regardless of the value of the user SNR. This is due to the price of a UGS bit which equals 0.45. Figure 5.5 also shows that the rtPS traffic class is allocated enough bandwidth to satisfy the user demands when the average user SNR equals 17.2 dB. Furthermore, the nrtPS curve is located above the rtPS class 1 curve at lower values of the SNR. This is due to the nature of the algorithm employed. As the QoS requirements of rtPS1 equal 3 Mbps, one server is equivalent to a high number of subcarriers. At the lower values of the user SNR, there are commonly not enough carriers to meet these demands. These carriers need to be allocated to a traffic class so that they are not wasted. Therefore, the bandwidth is passed on to the next most expensive traffic class (nrtPS) which has a much lower QoS requirement (30 kbps). Figure 5.5 presents plots of the revenue versus the average user SNR. These figures show that higher revenues can be earned when a cost based CP CAC policy is employed regardless of the value of the user SNR. Therefore, CP is a more convenient policy than CS in terms of revenue too. However, as the average user SNR grows, it can be observed that the difference between the two cases decreases. This is because in such a scenario, the overall available system throughput increases. Thus, a higher proportion of the offered traffic load from the more expensive traffic classes will be serviced. Both Figures 5.5 and 5.6 depict that the difference between the CS scheme and the CP scheme decreases as the average user SNR grows. Hence, although the CS scheme offers better resource utilization, a proper pricing scheme will lead to higher revenues for the operator and improved QoS for prioritized (more expensive) traffic classes.



,

Figure 5.4: Blocking Probability vs average user SNR

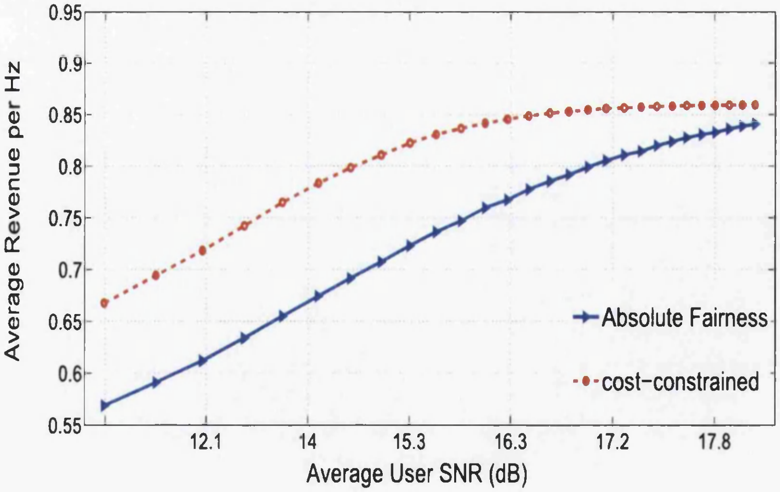


Figure 5.5: Revenue per Hz versus average user SNR

5.7 Summary

In this chapter, the channel gains and the cost of each traffic class were used to develop a model that partitions the available bandwidth in accordance with a CP-structure. The model ensures that the ongoing connections can be satisfied. Numerical results showed that with the use of this model, the pricing model can play a key role in ensuring QoS guaranteed services. Moreover, it was shown that prioritizing traffic classes leads to important revenue gains. These gains become more significant at low values of the user SNR.

Chapter 6

Critical Evaluations

The contributions of this thesis can be found in Chapters 3, 4, and 5. In Chapter 3, a rate-adaptive resource allocation problem under the SCSI assumption is considered. Works that have considered continuous sum-rate maximization problems under the SCSI assumption have focused on ergodic rate maximization [12],[34].

In [30]-[31], goodput (outage capacity) maximizing rate-adaptive resource allocation problems were solved under the partial CSI assumption (i.e users regularly feedback the channel state). A closed form expression for the data-rate that maximizes the goodput is obtained. However, this expression is only useful when reliable channel estimates are considered. In this thesis, a closed form expression is derived for the data-rate that maximizes the goodput of a Rayleigh fading channel when the transmitter only knows the user SCSI. This expression allows for the resource allocation problem to be simplified. Also, it is shown that the throughput gains that can be realized through optimal subcarrier loading grow with the value of the average user SNR. However, these gains will not exceed the value of e (where e is the base of the natural logarithm). Factors that affect the performance difference between ICSI and SCSI for goodput maximizing rate-adaptive resource allocation problems are also investigated. In [30]-[31], comparisons are made that show the difference of perfect CSI based resource allocation with partial CSI schemes. However, these comparisons did not account for the impact of neither multiuser diversity nor of the QoS constraint on the difference between the two schemes.

Exploiting multiuser diversity to efficiently perform OFDMA resource allocation has been well documented in the literature beginning with [22]. In Chapter 3, resources are adaptively allocated to users under the assumption that all users do not adhere to the same power delay profile. It is shown that when there are many users whose channels are characterized by the vehicular PDP (i.e have a large delay spread), the multiuser diversity gains that can be realized are greater than when there are many users whose channels are characterized by the pedestrian PDP (i.e the delay spread is small). As far as the author is aware, no study has explicitly compared the result of a rate-adaptive resource allocation problem under different power delay profile assumptions.

In Chapter 4, this thesis considers the practical scheme of only allowing a discrete set of modulation and coding schemes to be used for transmission rather than just the theoretical continuous rate. Outage capacities are again considered, as an ergodic capacity may be inappropriate as a measure for spectral efficiency because it applies to the case when each codeword is long enough to experience a sufficient number of different channel states. Using MMSE estimation theory, an expression for the probability of MCS selection level mismatch is derived for the case of outdated CSI. This expression is then used in an MCS level selection rule which maximizes the throughput subject to the BER/BLER constraint. The result is then exploited in the development of a scheduling rule which effectively deals with feedback delay. Using this rule, the throughput of a system can be maximized under the predetermined BER/block error rate constraints which are obtained through link layer simulation results. Similar work was presented in [33] which considers scheduling for outdated channel states. The authors combine a given set of CQI values into one scalar effective CQI using the EESM scheme. In this paper, Assaad and Ayoub select an MCS level in accordance with the following formula:

$$m^* = \arg\max_{m \in M} E[(1 - BLER(m))]R_m, \tag{6.1}$$

where $BLER(m)$ is the block error rate associated with a particular MCS level m, R_m is its spectral efficiency and $E[]$ denotes expectation. However, as the uncertainty regarding the user CSI follows a conditional distribution (i.e $h|\hat{h},\ f_d\tau$), an ergodic capacity may be an inappropriate measure because it applies when each codeword is long enough to experience a sufficient number of different channel states. Also, with the use of the proposed MCS level selection scheme instantaneous BER/BLER constraints can be met.

In Chapter 4, a feedback reduction scheme is also proposed. It is designed to select between two different types of channel feedback mechanisms for each user. These are: a) statistical CSI based feedback where users only feedback the mean of their SNR distribution to the BS and b) periodic feedback of the instantaneous channel states. This thesis shows that periodic ICSI feedback significantly increases the system sum-capacity. Therefore, it is useful to allow the users that are most likely to be scheduled to send their ICSI to the BS. Such schemes were proposed in [11],[40]-[42], where a threshold value is set which users must exceed in order to be allowed to feedback their ICSI to the BS. The value of this threshold is set so that the users most likely to be scheduled, feedback their CSI. However, with these schemes, the number of users that are most likely to be scheduled will vary significantly between different timeslots which results in variable-bit rate ICSI feedback channels. In practical systems, the bandwidth allocated for ICSI feedback purposes is fixed. Using the proposed scheme, the number of users who feedback their ICSI is also fixed. Moreover, the specific users who are the most likely to be scheduled will vary between the different timeslots. To frequently revoke and reallocate the resources reserved for feedback by means of radio resource control is not feasible on a fast time scale [43]. Periodic feedback schemes require only one signaling message for the allocation of a CQI channel and one for its release [39]. Thus, it is useful to only allow a certain set of users to periodically feedback their ICSI. It was shown in this thesis that using the proposed scheme there is a very low probability that signaling bandwidth for the allocation/release of channels reserved for ICSI feedback will be consumed. Moreover, relevant works have not differentiated between statistical and instantaneous CSI. In particular, it is assumed that all of the users are scheduled using instantaneous CSI.

When CAC functionality is accurate in a system with asymmetric fading statistics, it is likely that the users experiencing good channel conditions can meet their data-rate requirements when only their SCSI is known. This is useful, since the overhead costs of the ICSI scheme are much greater than the overhead costs of the SCSI scheme. With SCSI based feedback schemes, a single CQI value indicating the average user SNR can be fed back to the BS [43]. Also, the mean value of the SNR does not need to be updated as frequently as the ICSI values. The CAC unit will limit the number of ongoing connections so that the data-rate requirements of all users can be met. Also, in multiservice networks, the aggregate rate allocated to subscribers of each service cannot exceed the prescribed partition of the network throughput, particularly when subscribers have diverse QoS requirements [86]. In such a scenario and assuming that all users access the same service, a fairness oriented scheduling algorithm (i.e proportional fair, queue and channel aware) will assign more channels to the users with the lowest geometry. Otherwise, the requirements of those users will not be met and CAC functionality will not be accurate. Therefore, in a system where only a given number of users feedback their ICSI, enabling the group of users with the poorest geometry to do so will lead to more channels will be loaded with ICSI/CCSI. This will increase the system sum-capacity and improve the overall fairness. In this thesis, the simulations show that under these assumptions, the users with the best geometry can easily meet their data-rate requirements when the BS only knows the user SCSI. Therefore, those users will not benefit from frequency selective CQI reporting methods. Hence, it is useful to exclude such users from occupying the limited resources reserved for ICSI feedback.

In Chapter 5, the focus is on a fairness constrained optimal revenue bandwidth allocation strategy for multiservice mobile WiMAX networks. The bandwidth is partitioned in accordance with the CP scheme which is useful since it ensures the fairness of different priority connections. The work builds on [47], where the authors proposed a fairness constrained optimal revenue CP scheme for fixed WiMAX networks. In fixed WiMAX networks, the network capacity is constant; and therefore, the partitioning does not have to be dynamically adjusted. However, the inherently dynamic nature of mobile wireless networks means that the network capacity is not constant over time. Thus, in this case, the user channel conditions need to be accounted for when the bandwidth is partitioned. This allows the time-varying nature of the network capacity to be accounted for. In Chapter 5, a scheme is proposed to partition the available bandwidth in accordance with a CP structure whilst accounting for the user channel states and the cost of each different traffic class. The aim is to optimize the revenue whilst meeting the fairness constraints. The cost is used to prioritize the different traffic classes. The scheme also ensures that the data-rate of any ongoing connections are met.

In [55], the authors allowed new connections of a traffic class to be admitted into the network when the sum of the loads of each contributing bearer exceeds a certain threshold value. However, there is no information on how this threshold value should be set. In this thesis, the cost of each traffic class and the average SNR of all users present in the cell is used to adaptively determine the threshold value. It is shown that the cost of each class can play an important role in prioritizing services when the network is heavily overloaded.

The authors proposed a pre-emptive congestion control scheme for LTE in [58]. With the use of this method, high priority requesting bearers can displace lower priority connected bearers. This is clearly not a convenient policy. The proposed complete partitioning scheme presented in Chapter 5 accounts for the user QoS of the ongoing connections .

Chapter 7

Conclusions and Future Work

The next generation of Broadband Wireless Access networks requires very high spectral efficiencies so that bandwidth intensive services can be offered to users. The success of these networks is therefore strongly tied to the performance of corresponding resource allocation schemes. The research in this thesis focuses on resource allocation and Call Admission Control schemes for OFDMA networks where full/partial or statistical CSI is available at the transmitter. Following, is a summary of the key research contributions of the thesis and a proposal of further works:

7.1 Major Research Outcomes

In Chapter 3, adaptive subcarrier assignment for downlink multiuser OFDMA systems, when only statistical CSI is available at the transmitter side, has been studied. The problem addressed is maximizing the sum-capacity of the system subject to user data-rate requirements. In this chapter, the following important conclusions are drawn:

- When subcarriers are optimally loaded using the developed relationship between the average signal-to-noise ratio (SNR) and the Lambert-W function, important throughput gains can be realized. The throughput gains grow with the user SNR. It is shown through theoretical analysis that these gains cannot exceed the value of e as the user SNR approaches infinity. These gains are because of the SCSI model which does an excellent job of predicting the actual values of the Rayleigh fading channels.

Then, SCSI and ICSI based adaptive resource allocation is performed, and their results are compared. It is shown that even though the subcarriers are optimally loaded when SCSI is used, important differences between the two cases still exist. In particular, the following conclusions are reached:

•For the given parameters, the ICSI scheme yields 2-2.5 times higher values of the system throughput than SCSI based resource allocation. When a QoS constraint is considered, these values are slightly higher as more subcarrier reallocation operations are needed to satisfy the users. As the QoS constraint increases, more subcarrier reallocation operations are needed, leading to a greater difference between the two schemes.

•The results show that the difference between the ICSI and SCSI based resource allocation schemes is highly dependent on the number of active users present in the cell, the QoS constraint, and the SNR per subcarrier.

Finally, Chapter 3 studies the impact of the power delay profile on adaptive subcarrier assignment. It was shown through simulations, that when the instantaneous channel realizations are known, adaptive subcarrier assignment leads to higher gains under the vehicular power delay profile assumption. However, it is noted that under significant user mobility, acquisition of perfect CSI is impossible.

In Chapter 3, the results showed that adaptive resource allocation based on perfect CSI can significantly improve the performance of OFDMA systems. However, in real systems, accurate CSI is impossible because of noisy channel estimates, channel feedback delays, and processing delays. Therefore, only imperfect/outdated CSI and SCSI can be used to perform resource allocation. Chapter 4 differentiates between these two cases and performs resource allocation for the more practically relevant case where only a discrete number of modulation and coding levels are available. The presence of a call admission control unit is also assumed. Here, the following conclusions are drawn:

- Optimal selection of the MCS level leads to higher values of the system throughput.
- Even when the correlation between the current and outdated CSI is low, correlated CSI always leads to higher values of the system throughput when a constraint on the target BER requirement is imposed.

Since in a fast fading environment excessive overhead is required for the scheduler to continuously work with CCSI, a strategy was presented to efficiently utilize the limited resources reserved for feedback purposes. It was shown through simulation that when CAC functionality is accurate, allowing only the users with the lowest average SNR to utilize the uplink resources reserved for feedback leads to higher values of the system throughput and better fairness. Also, it was shown that great savings in terms of signalling bandwidth consumed for allocating/deallocating the bandwidth resources reserved for ICSI feedback can be realized.

In Chapter 4, the presence of the CAC unit was considered when comparing the system throughput achieved with correlated CSI to the system throughput achieved with statistical CSI. It is important to consider the CAC unit as in the 4G environment multiple services need to be supported and CAC functionality distributes the network throughput between these services allowing for QoS guarantees to be provided. Moreover, it limits the number of admitted flows. In Chapter 5, it is noted that the CAC functionality in OFDMA based networks is also strongly tied to the user SNR. A mechanism is developed which uses the cost of each service to differentiate and the user channel gains to partition the available bandwidth in accordance with a complete partitioning (CP)-structure. The following conclusions are drawn:

- The developed cost based CP CAC scheme admits users of expensive services (priority services) into the network even under conditions of extreme network overload.
- Under a Complete Sharing (CS) CAC scheme, the number of users of the prioritized services admitted into the network are significantly fewer.
- When the average user SNR improves, the difference (in terms of the number of users

of prioritized users admitted into the network) between the CS and CP scheme decreases.

- This novel cost based CP CAC scheme leads to important revenue gains for the network operators as the more expensive services are prioritized.

7.2 Further Work

The proposed schemes for resource allocation and Call Admission Control in this thesis provide insight into some important challenges. However, there still exists a number of open issues to extend this research which deserve further investigation. Below are listed some important open questions derived from the research carried out during this thesis:

In Chapter 3, the problem of resource allocation when only statistical CSI is available at the transmitter for Rayleigh fading channels was investigated. It was shown that when subcarriers are optimally loaded using the developed closed form expression that links the Lambert-W function with the average user signal-to-noise ratio the difference with the ICSI case can be minimized. It would be interesting to develop closed form expressions that link the average SNR with the optimum goodput for other channel models (i.e Ricean, Nakagami). Then, the difference with the ICSI based resource allocation scheme for these different channel models can be studied. It would be interesting to observe whether the difference between the two cases is similar to the result obtained for Rayleigh fading.

In Chapter 4, a comparison is made between CSI that is correlated with its current state and statistical CSI. It is useful to take into account the overhead due to signalling and pilot transmission when performing this comparison. This will allow for a deeper insight into the tradeoff between CSI overhead and CSI accuracy. The presence of a CAC unit is considered in the work of Chapter 4 where it is shown that when CAC functionality is accurate allowing the users with the lowest average SNR to feedback their CSI leads to more efficient utilization of the resources reserved for overhead purposes. It is important to consider the presence of the CAC unit in the problem formulation when developing heuristic algorithms for OFDMA resource allocation problems. This means imposing a limit on the optimum aggregate data rate of each class of users. The CAC unit constraint will ensure that the network throughput is not overused by a particular class of subscribers. The solution of such problems will lead to joint CAC/scheduling schemes.

In Chapter 5, it was shown that the use of a CP based CAC scheme is able to prioritize services. In this work, the cost was used to differentiate and prioritize the different services. Other methods of prioritizing the different services can be addressed. Moreover, the impact of the CAC scheme on the overall system throughput needs to be analyzed. This is because the network throughput will be distributed amongst the different services which could result in underutilization of the available spectrum. Therefore, hybrid CP/CS (Complete Sharing) mechanisms should be investigated.

In addition, since the resource allocation and CAC schemes studied in this thesis consider a single cell without any intercell interference, it would be also of great interest to extend this research to a multicellular environment. Another extension of the thesis work could be to study the performance of these implementations with real-time traffic models.

Appendix A

Proof of Eq. 3.5

The goodput is written as:

$$G(\rho) = \rho(\exp[-(\bar{\gamma}^{-1})(2^{\rho} - 1))]) \tag{A.1}$$

In order to find the value of ρ that yields the maximum values of G derivative of G with respect to ρ is set to zero. By setting $g(\rho) = -(\bar{\gamma}^{-1})(2^{\rho} - 1)$ A.1 can be written as:

$$\frac{dG}{d\rho} = \exp(g(\rho)) + \rho \exp(g(\rho)) \ln(2) g(\rho) = 0 \tag{A.2}$$

which reduces to:

$$1 + \ln(2)\rho g(\rho) = 0 \tag{A.3}$$

By replacing $g(\rho)$ with its original value the following expression is obtained:

$$\frac{\bar{\gamma}}{(2^{\rho}) \ln(2)\rho} = 1 \tag{A.4}$$

Using $2^{\rho} = \exp(\ln(2)\rho)$ this can be written as:

$$\frac{\bar{\gamma}}{(\exp(\ln(2)\rho)) \ln(2)\rho} = 1 \tag{A.5}$$

Setting $y = \ln(2)\rho$ this can be written in the form $y \exp(y) = \bar{\gamma}$. By applying the Lambert W function we get $y = W(\bar{\gamma})$. Replacing with the original value of y will result in Eq 3.5.

Appendix B

Calculation of the optimal loading limit

Here, it is shown that as the user SNR tends to ∞ the ratio of the optimal goodput that can be achieved on a subcarrier to the goodput that can be achieved on the same subcarrier without optimal loading is equal to e. The limit is written as:

$$\lim_{\gamma\to\infty} \exp\left[-(\gamma^{-1})(2^{\frac{W(\gamma)}{\ln(2)}}-1)\right]\frac{W(\gamma)}{\ln(2)}\frac{1}{\log_2(1+\gamma)\exp\left[-(\gamma^{-1})(2^{\log_2(1+\gamma)}-1)\right]}$$

(B.1)

Using the Lambert function identity, (A.6) can be written as:

$$\lim_{\gamma\to\infty} \exp\left[-\frac{\exp^{W(\gamma)}-1}{\gamma}\right]\left(\frac{W(\gamma)}{\ln(\gamma)}\right)\left(\frac{\ln(\gamma)}{\ln(1+\gamma)}\right)\frac{1}{\exp\left[-(\gamma^{-1})(2^{\log_2(1+\gamma)}-1)\right]}$$

(B.2)

Simple algebraic manipulations and application of Del'Hospital's rule to the second and third factor of (A.7) yield:

$$\lim_{\gamma\to\infty} \exp\left[-\frac{1}{W(\gamma)}-\frac{1}{\gamma}\right]\left(\frac{W(\gamma)}{1+W(\gamma)}\right)\left(\frac{\gamma+1}{\gamma}\right)\frac{1}{\exp\left[-(\gamma^{-1})(2^{\log_2(1+\gamma)}-1)\right]}$$

$$=\lim_{\gamma\to\infty}\frac{1}{\exp\left[-(\gamma^{-1})(2^{\log_2(1+\gamma)}-1)\right]}$$

$$=e. \quad (B.3)$$

This result also implies that the ICSI to SCSI ratio is equal to one as the user SNR approaches infinity.

Appendix C

Publications

Key results of this thesis can be found in the publications listed below. All of these publications have undergone the peer review process:

•L.Sivridis, J.Choi, and Y.Li, "A cost based resource allocation policy for multiservice mobile WiMAX networks", *Proceedings of IEEE International Symposium on Wireless Pervasive Communications*, pp. 489-494, Modena, Italy, May 2010.

•L.Sivridis, J.Choi, and Y.Li, "Call Admission Control in WiMAX networks", WiMAX Book, INTECH (Open access publisher), 2011. (accepted).

•L.Sivridis, X.Wang, and J.Choi, "Impact of CSI on Radio Resource Management for the OFDMA downlink", *Journal of Communications* (Special Issue on Practical Physical Layer Techniques for 4G Systems and Beyond), Academy Publisher, August 2011. (accepted for publication).

•L.Sivridis, X.Wang, and J.Choi, "Radio Resource Management for fast fading environments", *Journal of Communications* (Special Issue on Advances in Wireless Communications and Networks), Academy Publisher, October 2011. (accepted for publication).

•L.Sivridis, X.Wang, and J.Choi, "Impact of varying channel model mixtures on Radio Resource Management for the OFDMA downlink", *International Journal of Communication Networks and Information Security*, KUST Press, August 2011. (accepted for publication).

Bibliography

[1] J.E. Padgett, C.G. Gunther and T.Hattori, "Overview of wireless person communication systems, " *IEEE Communications Magazine*, vol.33, no.1, pp. 28-41, 1995.

[2] J.G. Proakis, Digital Communications, New York: McGraw-Hill, 1995.

[3] A.J. Viterbi, CDMA Principles of Spread Spectrum Communication, Reading: Addison-Wesley, 1995.

[4] R. Love, A. Ghosh, X. Weimin, and R. Ratasuk, "Performance of 3GPP high speed downlink packet access (HSDPA)," in *Proceedings of the 50th IEEE Vehicular Technology Conference (VTC-2004-Fall)*, pp. 3359-3363, September 2004.

[5] ITU-R Recommendation M.1645, "Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000," 2003.

[6] B.Evans and B.I.Wong, Resource Allocation in Multiuser Multicarrier Systems, Springer Science and Media, 2008.

[7] A.Goldsmith, Fundamentals of Wireless Communications, Cambridge University Press, 2005.

[8] R.V.Nee and A.d.Wild, "Reducing the Peak-to-Average power ratio of OFDM," in *IEEE Vehicular Technology Conference (VTC'98)*, pp. 2072-2076, Ottawa, Canada, May 1998.

[9] N. Al-Dhahir and J. M. Cioffi, "Optimum finite-length equalization for multicarrier transceivers," *IEEE Transactions on Communications*, vol. 44, no. 1, pp. 56-64, 1996.

[10] A. Goldsmith, "Adaptive modulation and coding for fading channels," in *Proceedings of IEEE Information Theory and Communications Workshop*, pp. 24-26, Kruger National Park , South Africa, June 1999.

[11] D.Gesbert and M-S.Alouini, "How much feedback is multi-user diversity really worth?," in *IEEE International Conference on Communications (ICC'04)*, vol.1, pp. 234-238, Paris, France, 2004.

[12] F.Brah, J.Louveaux, and L.Vandendorpe, "CDIT-Based Constrained Resource Allocation for Mobile WiMAX Systems," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, Article ID 425367, 8 pages, 2009.

[13] S.Y.Tang, P.Muller, and H.Sharif, WiMAX Security and Quality of Service: An End-to-End Perspective, Hoboken, NJ: Wiley, 2010.

[14] M.Bohge and A.Wolisz, "The impact of Channel State Information processing delay in Optimally Scheduled OFDMA networks," in *International OFDM Workshop (InOWo09)*, pp. 282-232, Hamburg, Germany, September 2009.

[15] H. Minn and V. Bhargava, "An Investigation into Time-Domain Approach for OFDM Channel Estimation," *IEEE Transactions on Broadcasting*, vol. 46, pp. 240-248, 2000.

[16] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1277-1294, 2002.

[17] S.Y Park, D.Park, and D.Love, "On Scheduling for Multiple-Antenna Wireless Networks Using Contention-Based Feedback," *IEEE Transactions on Communications*, vol. 55, no.6, pp. 1174-1190, 2007.

[18] J.Choi, T.Kwon, and Y.Choi, "Call Admission Control for Multimedia Services in Mobile Cellular Networks: A Markov Decision Approach," in *IEEE Symposium on Computers and Communications (ISCC 2000)*, pp. 594 - 599, Antibes, France, July 2000.

[19] J. Hayes, "Adaptive feedback communications," *IEEE Transactions on Communications*, vol. 16, pp.29 - 34 , Feb. 1968.

[20] S. T. Chung and A. Goldsmith, "Degrees of Freedom in Adaptive Modulation: A Unified View", in *Proceedings of IEEE Transactions on Communications*, vol.49, no.9, pp. 1561-1571, 2001.

[21] A. Goldsmith and S.-G. Chua, " Variable-rate variable-power MQAM for fading channels," in *IEEE Transactions on Communications*, vol. 45, no. 10, pp. 12181230, Oct. 1997.

[22] R. Knopp and P. Humblet, Information capacity and power control in single cell multiuser communications, *in IEEE International Conference on Communications*, vol. 1, 1995.

[23] D.Kivanc and G.Liu, "Computationally efficient bandwidth allocation and power control for OFDMA," *IEEE. Transactions on Wireless Communications*, vol.2, no.6, pp. 1150-1158, 2006.

[24] J.Jang and K.B.Lee, "Transmit Power Adaptation for Multiuser OFDM Systems," *IEEE Journal on Selected Areas on Communications*, vol.21, no.10, pp. 171-178, 2003.

[25] Y. Zhang and K.B. Letaief, "Multi-user Adaptive Sub-carrier and bit allocation with adaptive cell selection for OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 3, no. 5, pp. 1566-1575, 2004.

[26] M. R. Souryal and R. L. Pickholtz, "Adaptive modulation with imperfect channel information in OFDM," in *Proceeding of IEEE International Conference on Communications*, vol. 6, pp. 1861-1867 , June 2001.

[27] Y.Sigen, R.Blum, and L.Cimini, "Adaptive OFDM with Imperfect CSI," *IEEE Transactions on Wireless Communications*, vol.5, no. 11, pp. 3255-3265, 2006.

[28] Y.Yao and G.B.Giannakis, "Rate-Maximizing power allocation in OFDM based on partial CSI knowledge," *IEEE Transactions on Wireless Communications*, vol.4, no.3, pp. 1073-1083, 2005.

[29] I. C.Wong and B. L. Evans, " Optimal resource allocation in the OFDMA downlink with imperfect channel knowledge," in *IEEE Transactions on Communications*, vol. 57, pp. 232241, Jan. 2009.

[30] S. Stefanatos and N. Demetriou, "Downlink OFDMA Resource Allocation under partial Channel State Information," in *Proceedings of IEEE International Conference on Communications (ICC'09)*, pp. 1-5, Dresden, Germany, June 2009.

[31] S. Stefanatos, C.Papathanasiou, and N.Dimitriou, "Downlink mobile OFDMA with minimum user rate requests,", in *Proceedings of IEEE Global Telecommunications Conference*, pp. 1-5, Honolulu, U.S.A, December 2009.

[32] A.Ahmad and A.Assad, " Margin adaptive resource allocation in downlink OFDMA system with outdated channel state information, " in *Proceedings of IEEE 20th International Symposium on Wireless and Personal Communications*, pp. 1868-1872, Melbourne, Australia, September 2009.

[33] H.Ayoub and M.Assad, " Scheduling in OFDMA Systems with Outdated Channel Knowledge, " in *Proceedings of IEEE International Conference on Communications*, pp. 1-5, Cape Town, South Africa, May 2010.

[34] I.C.Wong and B.L Evans, " Adaptive downlink OFDMA resource allocation, " *Proceedings of 2008 42nd Asilomar Conference on Signals Systems and Computers*, pp. 2203-2207, Pacific Grove, U.S.A, October 2009.

[35] S.Choudhury and J.Gibson, " Information transmission over fading channels, " *Proceedings of IEEE 2007 Global Telecommunications Conference*, pp. 3316-3321 , Washington D.C, U.S.A, 2007.

[36] T.Kim and J.T-Lim, " Reduced feedback for capacity and fairness tradeoff in multiuser diversity, " in *IET Communication*, pp. 945-950, vol.5, no.7,2011.

[37] T.Kim and J.T-Lim, " Capacity and fairness tradeoff in multiuser scheduling system with reduced feedback", in *IEEE Communications Letters*, pp. 841-843 , vol.13, no.11, 2009.

[38] G.Hwang and F.Ishizaki, "Design of a fair scheduler exploiting multiuser diversity with feedback information reduction" in *IEEE Communications Letters*, pp. 124-126, vol.12, no.2, 2008.

[39] Cohen.R and Grebla.G, " Efficient allocation of CQI channels in broadband wireless networks ", in *Proceedings of IEEE Infocom*, pp. 96-100, Orlando, U.S.A, 2011.

[40] Hoon.K, Youngnam.H " An Opportunistic Channel Quality Feedback Scheme for Proportional Fair Scheduling", in *IEEE Communications Letters*, pp. 501-503, vol.11, no.3, 2007.

[41] S.Kittipiyakul, A. Boonkajay, and T. Aphichartsuphapkhajorn, "Performance of low-feedback-rate, gradient-based OFDMA subcarrier allocation with partial chan-

nel information", in *Proceedings of IEEE Symposium on Wireless and Pervasive Communications*, pp. 511-521, Modena, Italy, 2010.

[42] J. Jeon, K.Son, H-W Lee, and S, Chong, " Feedback Reduction for Multiuser OFDM Systems ", in *IEEE Transactions on Vehicular Communications*,pp. 160-169, vol. 59, no.1, 2009.

[43] S.Sesia, I.Toufik and M.Baker, LTE-The UMTS Long Term Evolution-From Theory to Practice, John Wiley and Sons, 2009.

[44] Yahya A.S, Al-Harthi, Y.S., Tewfik, A.H., Alouini, M.-S, "Multiuser diversity with quantized feedback", in *IEEE Transactions on Wireless Communications*, pp. 330-337, vol.1, no.6, 2007.

[45] Sanayei.S and Nosratinia.N, " Exploiting multiuser diversity with only 1-bit feedback", in *Proceedings of IEEE Wireless Communications and Networking Conference*, pp. 978-983, New Orleans, U.S.A, 2005.

[46] B. Rong, Y. Qian, and K. Lu, "Integrated downlink resource management for multiservice WiMAX networks," in *IEEE Transactions on Mobile Computing*, vol. 6, no. 6, pp. 621-632, Jun. 2007.

[47] B.Rong and Y.Qian, "Call Admission Control Optimization in WiMAX Networks," *IEEE Transactions on Vehicular Technology*, vol.57, no.4, pp. 2509-2521, 2008.

[48] Khemiri, S., Boussetta, K., Achir, N., and Pujolle, G. " WiMAX bandwidth provisioning service to residential customers " in *Proceedings of International Conference on Mobile Wireless Communications Networks*, pp. 116-120, Paris, France, September 2007.

[49] Niyato. D and Hossain.E "Connection admission control algorithms for OFDM wireless networks" in *Proceedings of IEEE Global Telecommunication Conference*, pp. 2455-2459, December 2005.

[50] B. Rong, Y. Qian, and K. Lu, " Downlink call admission control in multiservice WiMAX networks," in *Proceedings IEEE ICC*, pp. 5082-5087, June 2007.

[51] Niyato, D. and Hossain, E. "Call admission control for QoS provisioning in 4G wireless networks: issues and approaches" in *IEEE Network Magazine*, pp. 5-11, vol.19, no.5, October 2005.

[52] S-H Oh, and Y-H. Kim," Policy-based Congestion Control in WCDMA Wireless Access Networks for End-to-end QoS", in *Proceedings of the Joint International Conference on Optical Internet and Next Generation Network*,pp. 153-156, Jeju, Korea, July 2006.

[53] K. A. Ali, H. S. Hassanein, and H. T. Mouftah," Directional Cell Breathing Based Reactive Congestion Control in WCDMA Cellular Networks", in *Proceedings of IEEE Symposium on Computers and Communications*,pp. 685-690, Aveiro, Portugal, July 2007.

[54] Fantacci, R. ; Mennuti, G. ; Tarchi, D. "A priority based admission control strategy for WCDMA systems" in *Proceedings of IEE International Conference on Communications*, pp. 334-337,Seoul, Korea May 2005.

[55] R. Kwan and R. Arnott, et al.," On Radio Admission Control for LTE Systems," in *Proceedings of IEEE Vehicular Technology Conference (VTC)*, pp. 1-5, October 2010.

[56] S.Jeong J. Han W. Jeon "Adaptive connection admission control scheme for high data rate mobile networks" in *Proceedings of IEEE Vehicular Technology Conference*, pp. 2607-2611, Stockholm ,Sweden, 2005.

[57] M.Qian, Y. Huang, J. Shi, Y. Yuan. L. Tian and E. Dutkiewicz, "A Novel Radio Admission Control Scheme for Multiclass Services in LTE Systems", in *Proceedings of IEEE Globecom*, pp. 1-6, Beijing, China, December 2009.

[58] R.Kwan, Arnott.R, Trivisonno. R, Kubota. M "On Pre-Emption and Congestion Control for LTE Systems" in *Proceedings of IEEE Vehicular Technology Conference* ,pp. 1-5, October 2010.

[59] S. Pietrzyk, OFDMA for Broadband Wireless Access, Artech House, London, UK, 2006.

[60] W. Rhee and J.M. Cioffi, "Increase in capacity of multi-user OFDM system using dynamic subchannel allocation," in *Proceedings of IEEE Vehicular Technology Conference (VTC 2000-Spring)*, pp. 1085-1089, Japan, Tokyo, May 2000.

[61] D. Tse and P. Viswanath, Fundamentals of Wireless Communication, Cambridge University Press, May 2005.

[62] E.Biglieri, J.Proakis, and S.Shamai, "Fading Channels Information Theoretic and Communication Aspects," *IEEE Transactions on Information Theory*, vol. 44, no 6, pp. 1747-1758, 1998.

[63] G .Sierksma, Linear and Integer Programming, New York: Marcel Dekker, 1996.

[64] G. Reklaitis, A. Ravindran, and K. M. Ragsdell, Engineering Optimization, Methods, and Applications, New York:Wiley, 1983.

[65] Recommendation ITU-R M.1225, "Guidelines for evaluation of radio transmission technologies for IMT-2000," 1997.

[66] COST Action 231, "Digital mobile radio towards future generation systems," European Commission, Brussels, Belgium, Final Report, 1999.

[67] M.Rahnema, UMTS Network Planning, Optimization, and Inter-Operation with GSM, John Wiley and Sons, 2008.

[68] Y. Zhang and K. Letaief, "Multiuser subcarrier and bit allocation along with adaptive cell selection for OFDM transmission," *Proceedings of IEEE International Conference on Communications (ICC 2002)*, pp. 861-865, New York, U.S.A., April 2002.

[69] D. Kivanc, G. Li, and H. Liu, "Computationally efficient bandwidth allocation and power control for OFDMA," *IEEE Transactions on Wireless Communications*, vol. 2, no. 6, pp. 1150-1158, 2003.

[70] L.Sivridis, J.Choi, and Y.Li, "A cost based resource allocation policy for multiservice mobile WiMAX networks," *Proceedings of IEEE International symposium on wireless and pervasive communications (ISWPC 2010)*, pp. 489-494, Modena, Italy, 2010.

[71] T. Kailath, A. H. Sayed and B. Hassibi, Linear Estimation, Prentice Hall, New Jersey, 2000.

[72] J.M.Holtzmann and A.Sampath, "Adaptive averaging methodology for handoffs in cellular systems," *IEEE Transactions on Vehicular Technology*, vol.44, pp 59-66, 1995.

[73] J.Marcum, "A statistical theory of target detection by pulsed radar: Mathematical appendix," RAND Corporation, California, Research Memo RM-753, 1948.

[74] T.Girici, C.Zhu, J.Agre, and A.Ephremides, "Practical resource allocation schemes for OFDMA wireless systems", in *Proceedings of IEEE Conference on Consumer Communications and Networking (CCNC 2008*,pp. 897-901, Las Vegas, U.S.A., 2008.

[75] M.Andrews, A.Kumaran, A.Ramanan, A.Stolyar, and P.Whiting, "Providing quality of service over a shared wireless link," *IEEE Communications Magazine*, vol.39, no.2, pp. 150 - 154, 2001.

[76] B.Sadiq, R.Madan, and A.Sampath, "Downlink Scheduling for Multiclass traffic in LTE", *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, Article ID 510617, 18 pages, 2009.

[77] H.Holma and A.Toskala, LTE for UMTS OFDMA and SC-FDMA based Radio Access, John Wiley and Sons, 2009.

[78] 3GPP TSG-RAN, " TR 25.814: Physical Layer Aspects for Evolved UTRA (Release 7)", v 1.3.1, 2006.

[79] IEEE 802.16 Broadband Wireless Access Working Group, "Multi-hop Relay System Evaluation Methodology (Channel Model and Performance Metric)," 2007.

[80] So.J and Cioffi.J, " Feedback reduction scheme for downlink multiuser diversity " in *IEEE Transactions on Wireless Communications*, pp. 668-672, vol.8, no.2, 2009.

[81] R.Jain, The art of computer systems performance analysis, John Wiley and Sons, 1991.

[82] F. Lai, J. Misic, S. Chanson, "Complete Sharing versus Partitioning : Quality of Service Management for Wireless Multimedia Networks," in *Proceedings of International Conference on Computer Communications and Networks (ICCCN'98)*, pp. 584, Lafayette, U.S.A., 1998.

[83] J.S. Kaufman, "Blocking in a shared resource environment," *IEEE Transactions on Communications*, vol.29, no.10, pp. 1474-1481, 1981.

[84] D. Mitra, M.I. Rieman, and J. Wang, "Robust dynamic admission control for unified cell and call QoS in statistical multiplexers," *IEEE Journal on Selected Areas in Communications*, vol.16, no.5, pp. 692-707, 1998.

[85] B. Kraimeche and M. Schwartz, "Circuit access control strategies in integrated digital networks", in *Proceedings of the IEEE INFOCOM '84*, pp. 230-235, San Francisco, U.S.A., 1984.

[86] M.Awad, V.Mahinthan, V.Mehrjoo, S.Xuemin and J.W.Mark, "A Dual Decomposition Based Resource Allocation for OFDMA Networks With Imperfect CSI," *IEEE Transactions on Vehicular Technology*, vol.59, no.5, pp. 2394-2304, 2010.

[87] WiMAX Forum, "Mobile WiMAX part I: a technical overview and performance evaluation," March 2006.

[88] G.Song and Y.G,Li: "Cross Layer Optimization for OFDM wireless networks-Part I: Theoretical Framework," *IEEE Transactions on Wireless Communications*, vol.4, no.2, pp.614-624, 2005.

[89] D. Jagerman, "Some properties of the Erlang loss function," *Bell Systems Technical Journal*, vol. 53, no. 3, pp. 525-551, 1974.

[90] A. Jensen, "Moe's Principle," Copenhagen, Denmark: Copenhagen Telephone Co. (K.T.A.S.), 1950.