TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

Stanislav Petromilov Nonchev

Fairness-Oriented and QoS-Aware Radio Resource Management in OFDMA Packet Radio Networks:
Practical Algorithms and System Performance

Julkaisu 1138 • Publication 1138

Tampere 2013

Tampereen teknillinen yliopisto. Julkaisu 1138
Tampere University of Technology. Publication 1138

Stanislav Petromilov Nonchev

Fairness-Oriented and QoS-Aware Radio Resource Management in OFDMA Packet Radio Networks: Practical Algorithms and System Performance

Thesis for the degree of Doctor of Science in Technology to be presented with due permission for public examination and criticism in Tietotalo Building, Auditorium TB109, at Tampere University of Technology, on the 7th of June 2013, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2013

Supervisor

Mikko Valkama, Dr. Tech., Professor
Department of Electronics and Communications Engineering
Tampere University of Technology
Tampere, Finland

Pre-examiners

Boris Bellalta, Ph.D., Assistant Professor
Department of Information and Communication Technologies
Universitat Pompeu Fabra (UPF)
Barcelona, Spain

Evgeny Osipov, Ph.D., Associate Professor
Department of Computer Science, Electrical and Space Engineering
Luleå University of Technology
Luleå, Sweden

Opponent

Zekeriya Uykan, Dr.Tech., Assistant Professor
Department of Control and Automation Engineering
Dogus University
Istanbul, Turkey

ISBN 978-952-15-3080-7 (printed)
ISBN 978-952-15-3097-5 (PDF)
ISSN 1459-2045

Abstract

During the last two decades, wireless technologies have demonstrated their importance not only in people's personal communications but also as one of the fundamental drivers of economic growth, first in the form of cellular networks (2G, 3G and beyond) and more recently in terms of wireless computer networks (e.g. Wi-Fi,) and wireless Internet connectivity. Currently, the development of new packet radio systems is evolving, most notably in terms of 3GPP Long Term Evolution (LTE) and LTE-Advanced, in order to utilize the available radio spectrum as efficiently as possible. Therefore, advanced radio resource management (RRM) techniques have an important role in current and emerging future mobile networks.

In all wireless systems, data throughput and average data delay performance, especially in the case of best effort services, are greatly degraded when the traffic-load in the system is high. This is because the radio resources (time, frequency and space) are shared by multiple users. Another big problem is that transmission performance can vary greatly between different users, since the channel state depends heavily on the communication environment and changes therein. To solve these challenges, new major technology innovations are needed.

This thesis considers new practical fairness-oriented and quality-of-service (QoS) -aware RRM algorithms in OFDMA-based packet radio networks. Moreover, using UTRAN LTE radio network as an application example, we focus on analyzing and enhancing the system-level performance by utilizing state-of-the-art waveform and radio link developments combined with advanced radio resource management methods. The presented solutions as part of RRM framework consist of efficient packet scheduling, link adaptation, power control, admission control and retransmission mechanisms. More specifically, several novel packet scheduling algorithms are proposed and analyzed to address these challenges.

This dissertation deals specifically with the problems of QoS provisioning and fair radio resource distribution among users with limited channel feedback, admission and power control in best effort and video streaming type traffic scenarios, and the resulting system-level performance. The work and developments are practically-oriented, taking aspects like finite channel state information (CSI), reporting delays and retransmissions into account. Consequently, the multi-user diversity gain with opportunistic frequency domain packet scheduling (FDPS) is further explored in spatial

domain by taking the multiantenna techniques and spatial division multiplexing functionalities into account.

Validation and analysis of the proposed solutions are performed through extensive system level simulations modeling the behavior and operation of a complete multiuser cell in the overall network. Based on the obtained performance results, it is confirmed that greatly improved fairness can be rather easily built in to the scheduling algorithm and other RRM mechanisms without considerably degrading e.g. the average cell throughput. Moreover, effective QoS-provisioning framework in video streaming type traffic scenarios demonstrate the effectiveness of the presented solutions as increased system capacity measured in terms of the number of users or parallel streaming services supported simultaneously by the network.

Preface

This dissertation is the result of my studies conducted between years 2006 and 2012 at the Department of Communications Engineering at Tampere University of Technology, Tampere, Finland and at the Electrical Engineering Department at Qatar University, Doha, Qatar. The research work started in mid 2006 within the "LTE Quasi-Static Radio System Simulation and Analysis" project carried out by Department of Communications Engineering, Tampere University of Technology. This project was originally sponsored by and conducted in close collaboration with Nokia Research Center, and later continued with Nokia Devices and Renesas Mobile Corporation. The latter part of the thesis work has been conducted under the project "Energy and Cost Efficiency for Wireless Access (ECEWA)" funded by the Finnish Funding Agency for Technology and Innovations (Tekes).

I would like to express my sincere gratitude to my supervisor Prof. Mikko Valkama for his encouragement, guidance and constant support. I would also like to give warm thanks to my co-supervisor Dr. Ridha Hamila for mentoring my research work and for sharing his profound knowledge. Special thanks are also due to Prof. Markku Renfors for giving me the initial opportunity to work at the Department of Communications Engineering and creating a good atmosphere for doing research.

I am very grateful to my colleague researchers and to everybody else at the Department of Communications Engineering for making it such a nice place to work. Special thanks go also to Mrs. Tarja Erälaukko, Mrs. Ulla Siltaloppi, and Ms. Elina Orava, for being able to promptly help with practical and administrative matters. Furthermore, I would like to express my appreciation for working on challenging topics with my close colleagues Mr. Juha Venäläinen and Dr. Fan Yong.

I wish to thank to Assistant Professor, Dr. Boris Bellalta from Universitat Pompeu Fabra and Associate Professor, Dr. Evgeny Osipov from Luleå University of Technology, the reviewers of this thesis, for their valuable comments and recommendations. I am also grateful to Assistant Professor, Dr. Zekeriya Uykan from Dogus University, for his willingness to act as the opponent in the defence.

I gratefully thank the Tampere Doctoral Programme in Information Science and Engineering (TISE) for funding my research visits and Nokia Foundation for financial support in the form of a grant award.

I would also like to thank my office mates, friends and colleagues Toni Levanen, Danai Skournetou, Tero Isotalo, Nikolay N. Tchamov, Svetozar Broussev, Kostadin Dabov, Vladislav Uzunov and Mihail Georgiev for the encouraging working atmosphere as well as for continuous helpfulness. At the same time, I am very thankful to all the personnel of TAMP Oy - Kari Virtanen, Lyubomir Zhivkov, Zdravko Nikolov and Donka Nikolova for the encouraging working atmosphere and creating a pleasant place to work.

Last but not least, I am immensely indebted to my parents and grandparents, especially my mother Todorka, my father Petromil and my brother Evgeny, for their love and support from a thousand miles away.

Finally, I wish to thank my wife Magdalena and my daughter Antonia-Teya for being my main inspiration, and whose love, support and encouragement have accompanied throughout my life.

Tampere, May 2013

Stanislav Nonchev

Contents

List of Publications

The main contributions of this compound dissertation are contained in the following original publications [P1]-[P11].

[P1] S. Nonchev, J. Venäläinen and M. Valkama, "New frequency domain packet scheduling schemes for UTRAN LTE downlink," in *Proceedings of ICT-MobileSummit 2008 Conference* (ICT-2008), Stockholm, Sweden, June 2008, 8 pages.

[P2] S. Nonchev and M. Valkama, "Efficient packet scheduling schemes for multiantenna packet radio downlink", in *Proceedings of Fifth Advanced Int. Conf. Telecommunications* (AICT-2009),pp. 404-409, Venice, Italy, May 2009.

[P3] S. Nonchev and M. Valkama, "Efficient power-aware packet scheduling for multiantenna packet radio systems," in *Proceedings of First International Conference on Networks and Communications* (NETCOM-2009), pp. 344-348, Chennai (Madras), India, December 2009.

[P4] S. Nonchev and M. Valkama, "Advanced packet scheduling in soft frequency reuse scenarios for multiantenna packet radio systems," in *Proceedings of Future Network and MobileSummit 2010,* pp. 1-8, Florence, Italy, June 2010.

[P5] S. Nonchev, M. Valkama, R. Hamila and M. Hasna "On the performance of advanced QoS-aware packet scheduling for multiantenna packet radio systems" in *Proceedings of 3rd IEEE International Conference on Broadband Network and Multimedia Technology 2010* (IC-BNMT-2010), pp. 209-304, Beijing, China, October 2010.

[P6] S. Nonchev, M. Valkama and R. Hamila, "Effect of high-velocity scenarios on the performance of MIMO LTE packet scheduling," in *Proc. 8th International Multi-Conference on Systems, Signals and Devices* (SSD-2011), pp. 1-6, Sousse, Tunisia, March 2011.

[P7] S. Nonchev and M. Valkama, "QoS-oriented packet scheduling for efficient video support in OFDMA-based packet radio systems," in *Proceedings of the 4th International*

Conference on Multiple Access Communications (MACOM-2011)*,* pp. 168-180, Trento, Italy, September 2011.

[P8] S. Nonchev, M. Valkama and R. Hamila, "Advanced packet scheduling for efficient video support with limited channel feedback on MIMO LTE downlink," in *Proceedings of IEEE GLOBECOM Workshop on Enabling Green Wireless Multimedia Communications,* pp.766-771, Houston, TX, December 2011.

[P9] S. Nonchev and M. Valkama, "A new fairness-oriented packet scheduling scheme with reduced channel feedback for OFDMA packet radio systems," *International Journal of Communications, Network and System Sciences* (IJCNS), vol. 2, no.7, pp. 608-618, October 2009.

[P10] S. Nonchev and M. Valkama, "Efficient packet scheduling schemes with built-in fairness control for multiantenna packet radio systems," *International Journal on Advances in Networks and Services (IJANS)*, vol.2, no. 2&3, pp, 182-194, 2009.

[P11] S. Nonchev and M. Valkama, "Advanced radio resource management for multiantenna packet radio systems," *International Journal on Wireless and Mobile Networks (IJWMN)* vol 2.2, pp. 1-14, 2010.

The author of the thesis is the main contributor of all the work presented in [P1]-[P11]. The thesis author has invented all the technical solutions, including the proposed scheduling metrics and overall radio resource management mechanisms, and also implemented all the system-level performance simulations. The role of the co-authors has mostly been in contributing to the writing and final appearance of the articles.

List of Abbreviations

16QAM	16 Quadrature Amplitude Modulation
3G	Third Generation
3GPP	Third Generation Partnership Project
64QAM	64 Quadrature Amplitude Modulation
AC	Admission Control
ACK	Acknowledgement
aGW	Access Gateway
AMC	Adaptive Modulation and Coding
AMPS	Advanced Mobile Phone Service
ATB	Adaptive Transmission Bandwidth
BLER	Block Error Rate
CAPEX	Capital Expenditures
CBR	Constant Bit Rate
CC	Chase Combining
CCI	Co-Channel Interference
CDF	Cumulative Distribution Function
CQI	Channel Quality Information
CSI	Channel State Information
dB	deciBel
DFT	Discrete Fourier Transform
DVB	Digital Video Broadcast
eNode-B	Evolved Node B
EPC	Evolved Packet Core
EESM	Exponential Effective SINR Metric
E-UTRAN	Evolved Universal Terrestrial Radio Access Network
FD	Frequency-Domain
FDD	Frequency Division Duplex
FDPS	Frequency-Domain Packet Scheduling
FFT	Fractional Power Control
FTP	File Transfer Protocol

GBR	Guaranteed Bit Rate
GGSN	Gateway GPRS Support Node
GPRS	General Packet Radio Service
GPS	Generalized Process Sharing
GSM	Global System for Mobile Communication
GW	Gateway
HARQ	Hybrid Automatic Repeat reQuest
HSDPA	High Speed Downlink Packet Access
HSPA	High Speed Packet Access
HSUPA	High Speed Uplink Packet Access
ICI	Inter-Cell Interference
IDFT	Inverse Discrete Fourier Transform
IETF	Internet Engineering Task Force
IFFT	Inverse Fast Fourier Transform
ILLA	Inner Loop Link Adaptation
IM	Instant Messaging
IP	Internet Protocol
IR	Incremental Redundancy
ITU	International Communication Union
KPIs	Key Performance Indicators
LA	Link Adaptation
LTE	Long Term Evolution
MAC	Medium Access Control
MCS	Modulation and Coding Scheme
MIMO	Multiple Input Multiple Output
MME	Mobility Management Entity
MMS	Multimedia Message Service
MRC	Maximal Ratio Combining
MU	Multi-User
NACK	Negative Acknowledgement
NMT	Nordic Mobile Telephony
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OLLA	Outer Loop Link Adaptation

PAPR	Peak-To-Average Power Ratio
PARC	Per-Antenna-Rate-Control
PC	Power Control
PF	Proportional Fair
PRB	Physical Resource Block
PS	Packet Scheduler
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RE	Resource Element
RNC	Radio Network Controller
RRM	Radio Resource Management
SAE	System Architecture Evolution
SC-FDMA	Single-Carrier Frequency Division Multiple Access
SDM	Spatial Domain Multiplexing
SINR	Signal-to-Interference-plus-Noise Ratio
SNR	Signal-to-Noise Ratio
SU	Single-User
TD	Time-Domain
TDPS	Time-Domain Packet Scheduling
TTI	Transmission Time Interval
UE	User Equipment
UL-SCH	Uplink Shared Channel
UMTS	Universal Mobile Telecommunications System
UTRAN	Universal Terrestrial Radio Access Network
VOD	Video On Demand
VoIP	Voice over Internet Protocol
WCDMA	Wideband Code Division Multiple Access
WLAN	Wireless Local Area Network

Chapter 1

Introduction

1.1 Research Area and Motivation

In recent years, the growing demands for wireless and mobile services have led to considerable research towards the development and integration of technologies that allow fulfilling the long-dreamed goal of "anything, anytime, anywhere" communication. As a concrete example, as part of the Third Generation Partnership Project (3GPP) Release'5 standardization work on Wideband Code Division Multiple Access (WCDMA) [1]-[3], the concept of High Speed Packet Access (HSPA) has been introduced achieving wireless data rates in the order of 10 Mbps [4], [5]. Current evolution is driving bit rates up from today's figure of around 10 Mbps in 3/3.5G HSPA towards 100 Mbps in Evolved Universal Terrestrial Radio Access Network (E- UTRAN) [6], [7] and eventually towards 1Gbps in future mobile networks (IMT Advanced) [8]. To support these goals, all established standardization bodies (3GPP, IEEE, etc.) have recently set up coherent design objectives to emerging radio system developments. These are: flexible radio system and waveform bandwidths in the range of 20-100 MHz, greatly improved spectral efficiency, improved system capacity and coverage, and packet switched domain only services [8]-[10].

In general, the purpose of radio resource management (RRM) is to ensure planned coverage for each service, required connection quality, reasonable inter-user fairness, and to optimize the system resource usage [11] in terms of time, frequency and spatial domains. The use of sophisticated RRM algorithms, such as channel-aware packet scheduling and link adaptation, enables efficient utilization of these air interface and hardware resources and enables system performance improvements. On the other hand, most of the emerging RRM enhancements require channel state and other side information to be communicated between base stations and terminals, which in turn considerably increase the signaling overhead in the system [5], [6], [12], [13]. Finding a good practical balance between the amount of side-information signaling and the resulting RRM performance improvements is the main problem area, to which this dissertation work is focusing on. The main emphasis is on RRM solutions, and especially packet scheduling methods, that take the inter-user fairness into account such that reasonable bit rates can also be obtained close to cell edges. Another big challenge in future radio systems is related to the Quality of Service requirements of the emerging new wireless data and multimedia services [14], like video streaming. To satisfy various QoS requirements of different services, like stringent delay constraints, joint optimization and design of the physical layer and RRM layer mechanisms is needed. This is also one of the main themes in this dissertation work, especially in the later part of the work and associated publications [P5], [P7], [P8].

At physical layer, one essential enabling tool towards reaching the above goals and in increasing the radio system performance is to use several transmitter and receiver antennas (in both base stations and mobile terminals) known as the multiple-input multiple-output (MIMO) technique, in order utilize the spatial domain in addition to more traditional time and frequency domains. In the seminal works reported e.g. in [15], [16], it was shown that the capacity of a MIMO link can be up to N times larger than the single-antenna link capacity where N is the minimum of the numbers of transmit and receive antennas. In general, the term MIMO includes traditional beamforming, diversity techniques and spatial multiplexing techniques [16]. To reach the ambitious target peak data rates of 3GPP Long Term Evolution (LTE)/LTE-Advanced, MIMO in terms of spatial multiplexing and diversity is exploited to increase the peak data rate of users. While currently the main focus in LTE/LTE-Advanced standardization is on closed-loop precoded MIMO schemes, the work in this thesis focuses more on relatively simple open-loop type schemes where explicit channel response is not known at the transmitter side for precoding purposes. This is partially because most of the reported work was already completed in 2006-2009 during which the open-loop schemes were considered as the core baseline technology.

At system level, the spatial domain enabled by multiple antennas can be utilized in terms of either the single-user (SU) or multi-user (MU) MIMO approach. From the RRM point of view, the former (SU) restricts the use of one time-frequency-space resource element to individual user equipment, while the latter (MU) enables multiuser sharing of individual time-frequency space resource elements. In this thesis, we consider both SU-MIMO and MU-MIMO developments since practically all radio systems are anyway inherently multiuser systems.

1.2 Related Work and Earlier Contributions

The performance of different packet radio networks has been widely investigated in the literature. Good selected examples are e.g. [1]-[5] and the references therein. Earlier WCDMA and HSPA network performance results [5]-[14] have been considered as the first step towards the new UTRAN LTE standard in 3GPP. The major outcomes reported in [19]-[22] pointed out the direction and gave the preliminary knowledge of what should be expected from LTE in terms of system capacity. The problem of efficient resource usage arose and was further addressed by proposition of RRM techniques reported e.g. in [23]-[26]. In [27], [28] the study considered various network aspects, traffic models, cell scenarios, and packet scheduling strategies, thus revealing the tradeoff between system throughput, cell coverage and user fairness.

The problem of choosing the right scheduling strategy in OFDMA- based packet networks gained much attention, e.g. [11], [17], and the supported studies in this direction indicated replacement of the "blind" scheduling strategies (like classical round robin, RR) by more advanced throughput/fairness oriented approaches (such as proportional fair, PF) [2], [5], [6]. In [29],[P1] decoupled time-frequency domain PS strategy is introduced revealing the potential of Proportional Fair scheduler. It was shown that time domain (TD) - frequency domain (FD) division and processing gives better control over user fairness, as well as considerable gains in throughput and coverage compared to opportunistic time domain packet scheduling (TDPS) alone. Consequently, in [30] the performance of spatial division multiplexing (SDM) multiple-input multiple-output (MIMO) techniques combined with FDPS is demonstrated to enhance the system capacity. Thus, new challenge appeared to RRM framework as the extended spatial domain functionality requires extended modifications to RRM entities. Moreover, it was found that there are always interactions of MIMO with other performance gain mechanisms in the system and the gain from different functionalities cannot be added to each other as if they were addressed separately. This is also one central theme in this dissertation work, through realistic system-level performance studies. Various system-level verification results on different proposed MIMO schemes with different complexity, signaling requirements and gain mechanisms have been reported in [31]-[33], [P2], [P10], [P11]. The major outcome is in selecting PF-based scheduling algorithms as a primary strategy in future network evaluations. Thus, most of the proposed scheduling strategies originate from PF and target better resource utilization and further system performance optimizations.

The solutions proposed in [33]-[40] considered the role of UE feedback as the key element in RRM functionality. Several Channel Quality Indicator (CQI) design cases were proposed together with corresponding performance analysis in OFDMA systems [36], [37]. In these works, it was clearly indicated that terminal measurement, estimation imperfection, compression and delays of user reports have significant impact on the system performance. Consequently, reducing the signaling overhead by introducing different reduced feedback schemes is demonstrated to be another key factor, e.g., in [39]. These are also considered in details in the system performance evaluations presented in this thesis.

Extensive performance analysis of advanced RRM functionalities includes traffic scenarios, QoS guarantees, user velocities etc. Considering PS as a key element in QoS provisioning, the basic PF principle does not fully consider the user's QoS requirements in its scheduling operation. Different variations of PF scheduler have been proposed to enhance the QoS-awareness, such as the guaranteed bit rate (GBR) or the maximum allowed packet delay, e.g. in [41], and also in this dissertation work in [P5], [P7]. Another proposed variation uses required activity detection (RAD) to adjust the user priorities based on scheduling history [42]. The performance of QoS-aware packet schedulers in real time and non-real time traffic has been extensively studied in [43]. The state of the art packet schedulers consider mixed type traffic scenarios and service differentiation based on queue state information parameters, which are not considered in this thesis work. Instead, video streaming traffic is used to estimate the number of UEs that can be supported by the system with employed QoS constrains [P8].

There is also a vast collection of works in the literature that address scheduling and resource allocation from a more theoretical perspective, assuming e.g. perfect instantaneous channel knowledge. Examples of such work are found in [44]-[46]. In this dissertation work, as already stated, the focus is on advanced packet scheduling and other RRM functionalities in a practical mobile cellular radio network context, and thus the forthcoming presentation also focuses on practical perspectives instead of deep theoretical issues.

1.3 Research Objectives and Outcomes

Future multiuser packet radio systems provide new opportunities to enhance system spectral efficiency and to obtain high data rates, low latency and packet optimized cellular network. These challenges are directly bound to RRM functionality comprising fast link adaptation (LA) including adaptive modulation and coding (AMC), dynamic packet scheduling (PS), power control (PC), admission control (AC) and reliable transmission – retransmission schemes e.g. HARQ. Increasing the average sector throughput, inter-user fairness, and/or coverage requires fast interaction between PS and LA entities as well as accurate channel-state (feedback) signaling. Another closely-related aspect is the utilization of advanced MIMO techniques and thereby providing spatial domain expansion to RRM functionality. On the other hand, QoS provisioning requires QoS aware PS strategy and AC.

This research focuses on the above advanced radio resource management techniques in future multiuser packet radio systems and examines the resulting radio system performance at cell level. The main RRM mechanisms considered here are fast channel-dependent scheduling, link adaptation and re-transmission methods under reduced practical feedback information. Moreover, we also consider the elementary use of multiple transmit and receive antennas, combined with advanced RRM techniques. RRM functionalities that take into account the physical layer waveform design as well as the different QoS requirements of various wireless services are becoming increasingly important in emerging radio communications systems. Thus, both best effort type full buffer traffic as well as video streaming type traffic scenarios are considered in this work. Furthermore, lots of emphasis is put on most of the developments in inter-user fairness such that service areas like cell edges where signal-to-interference-plus-noise-ratios (SINR) are low can also receive reasonable levels of service experience. The issue of heavily varying bit rates in cellular networks has recently received vast media attention, and the Finnish authorities have imposed more strict regulations on mobile operators as to how they advertise their network performance. Indeed, marketing based on the (theoretical) highest bit rates only is not considered anymore appropriate but the focus is more on the range of realistic bit rates that can be expected when receiving service from the network. This can also be seen as one incentive for mobile operators and network vendors to develop solutions to improve the fairness and cell-edge performance of their networks, being thus very well in-line with the contents of this thesis work. In order to make the studies realistic, the 3GPP LTE cellular radio framework and design considerations are mostly used in the developments and analysis [14].

The main outcomes and contributions of this dissertation work include the following

(1) Development of novel channel-aware packet scheduling (PS), link-adaptation (LA), channel-state (feedback) signaling and re-transmission mechanisms for future multiuser OFDMA-based packet radio systems. Special focus is put on efficient packet scheduling algorithms with increased resource allocation and inter-user fairness in case of best effort type full buffer data. [P1], [P3], [P5], [P8]-[P11].

(2) System performance analysis of advanced RRM algorithms with single and multi-antenna transmission schemes and reduced feedback information. Moreover, different cell

scenarios are used for further performance evaluation of the proposed multiuser PS algorithms [P2], [P3], [P9], [P10].

(3) Development of novel power-aware packet scheduling schemes and performance evaluations under different soft frequency reuse (SFR) scenarios. [P4], [P5], [P11]

(4) System performance analysis with advanced RRM algorithms under high-velocity scenarios. Specifically, different user velocities are considered in the performance evaluations and the impact over packet scheduler functionalities is investigated [P6].

(5) Development of novel QoS-aware packet scheduling schemes and performance evaluation for real-time video traffic scenarios. More precisely, this includes development of new packet scheduling schemes for video streaming scenarios with emphasis on QoS guarantees, like delay constraints, and user fairness [P7], [P8].

(6) Development of extensive quasi-static cellular system network simulator to analyze and verify the system-level performance of the developed RRM and interference management algorithms. Special focus is on obtaining realistic comparative performance figures in emerging OFDMA-based radio system context and especially LTE.

1.4 Scientific Methods Employed

While classical radio link level work, like waveform and coding oriented developments, and associated performance analysis can in many cases be addressed analytically by pen and paper, the modeling of a complete mobile radio network or even a single cell with one base-station and multiple user equipment is extremely complex. The central methodology in system level work is appropriate modeling of key entities, like packet scheduling, link adaptation, packet transmission and re-transmission, and especially the complex interactions between them. In this work, propagation models are used for characterizing the impacts of the core radio environment, and when combined with essential transmitter and receiver signal processing elements and interference scenarios, SINR expressions can be formulated. The behavior of the SINR's of different UEs over time, frequency and space, when moving in the serving cell and being controlled by the e-NB then form the basis for algorithm-level work in optimizing the behavior of, e.g., packet scheduler and link adaptation. Furthermore, incorporating the QoS requirements of different wireless applications, in terms of, e.g., minimum bit rate and/or maximum delay requirements, through appropriate service modeling is one central aspect. Due to the massively complex interactions of the core processes, analytical closed-form performance analysis of the whole network is not feasible, but system-level simulations are utilized instead, where all the functionalities and central processes and their evolution over time are modeled in software. This also allows then varying many central parameters, like number of users, cell size, exact scheduling metrics, etc., and observing the impact at the network level. Furthermore, statistics of the individual users can also be collected and examined. In this thesis work, we mostly aim at conforming to baseline performance evaluation requirements set by 3GPP for LTE [14].

1.5 Novelty and Contributions

The main contributions of this study are the design and analysis of a downlink RRM framework taking into account the complex interaction of PS, LA, and HARQ entities combined with practical finite-rate feedback reporting schemes and MIMO Tx/Rx techniques in LTE-downlink-type OFDMA networks. Advanced multiuser fairness-oriented, power-aware and QoS-aware packet scheduling algorithms exploiting the finite-rate user feedback in terms of CQIs are proposed, studied and analyzed. Extensive system simulator development, taking into account complex interactions and practical limitations, is carried out during the PhD study to evaluate the performance of the proposed algorithms including mathematical modeling considerations as well as software design, implementation and testing. Such practically oriented system-level performance analysis of the proposed RRM methods is tailored towards the LTE DL network model as a practical application scenario. Consequently, implementation of the proposed RRM methods is seen to rest on a solid basis and is expected to be fairly straight-forward through appropriate software updates to existing RRM mechanisms in eNode-Bs.

1.6 Thesis Outline

The rest of this dissertation is organized as follows:

- Chapter 2, *Mobile Network Evolution towards LTE-Advanced:* This chapter presents a brief overview of the mobile network system evolution with emphasis on 3GPP LTE.
- Chapter 3, *RRM in LTE-Downlink type OFDMA Packet Radio Systems – Fundamentals and Proposed Advanced Methods:* This chapter presents a general description of the downlink RRM functionalities in LTE. Furthermore, a description of PS and its interaction with HARQ, LA including AMC are detailed. The more advanced methods proposed in this thesis work are also described.

- Chapter 4, *System-Level Performance Evaluation Methods and Simulation Tool:* This chapter describes the system simulation techniques and tools for evaluating network performance. Furthermore, the detailed description of key modeling aspects and parameters are presented and discussed.
- Chapter 5, *Selected Research Outcomes and Examples:* This chapter provides examples of the most essential findings and outcomes within the framework of proposed packet scheduling algorithms and system level performance. Illustrations of the obtained network performance with the proposed methods are given, analyzed and discussed.
- Chapter 6, *Conclusions:* This chapter provides a summary of the overall study with concluding remarks.

Chapter 2

Mobile Network Evolution towards LTE-Advanced

2.1 Mobile Network Evolution

Mobile network evolution is the successful outcome of cooperation among numerous partners in 3GPP, 3GPP2, IEEE, Internet Engineering Task Force (IETF) and others. Today, the global dominant paths have been developed mainly by 3GPP and IEEE, as presented in Figure 1.

Evolution began several decades ago with the early deployments of analog cellular system concept. In the 1980s the first generation mobile technology appeared, supporting voice only services. The analog Nordic Mobile Telephony (NMT) was the first international mobile communication system introduced in Nordic countries in 1981, simultaneously with the introduction of analog Advanced Mobile Phone Service (AMPS) in North America. The bulky size of the user equipment, the inconsistent voice quality and the high service cost were some of the main drawbacks. After the introduction of the 'roaming' concept and digital technology, it was realized that new mobile standards are needed.

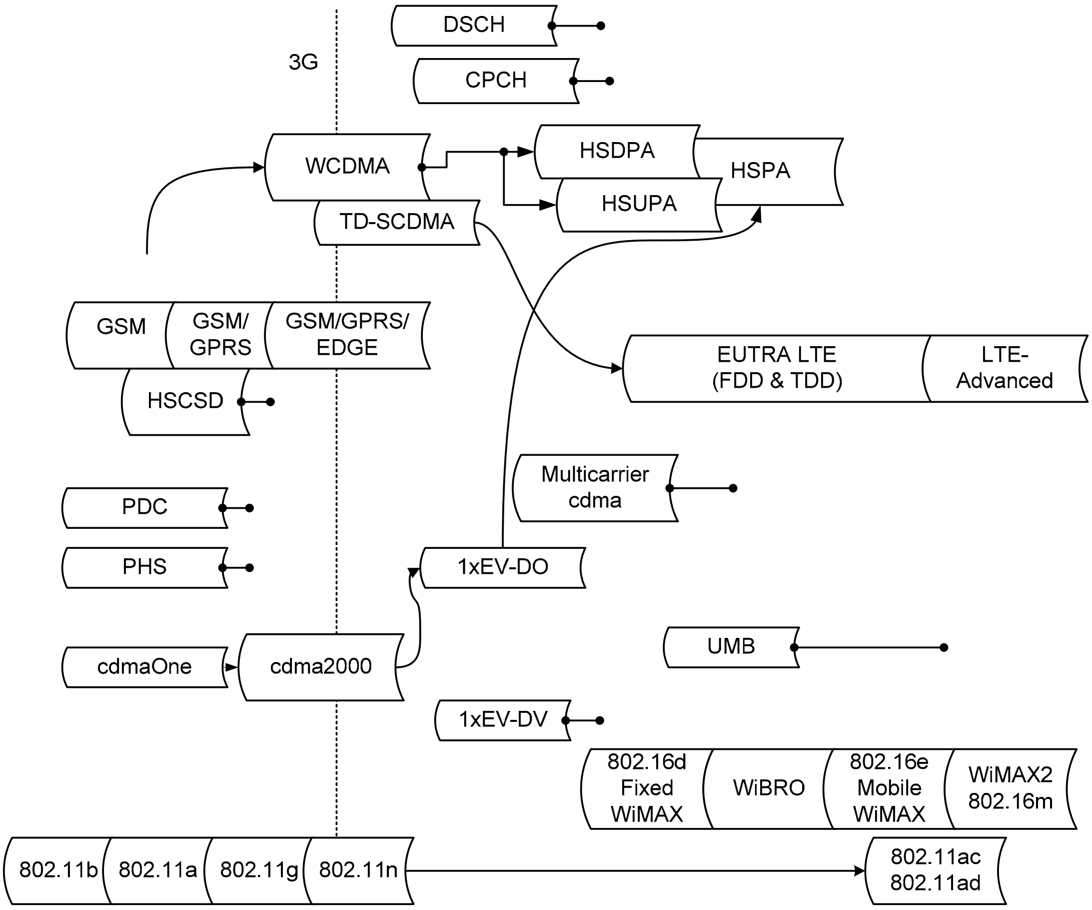


Figure 1: Mobile communication standards and systems.

2.1.1 Second Generation (2G and 2.5G) Systems

In the early 1990s the 2nd generation (2G) mobile communication system was introduced as Global System for Mobile Communications (GSM). Here the major advances were the newly provided data services enabling text messaging (SMS) and elementary e-mail as well as reduced-size equipment and lower costs. However, introducing new services and applications in 2G was fairly limited due to the low data rates being only in the order of 10 kbps or so.

Later on, with the introduction of packet switched services, the radio access system was upgraded to the General Packet Radio Service (GPRS) architecture referred to as 2.5G. Users were able to be permanently online and paying only for actual data used. Thus, GPRS offered data services like multimedia messaging, e-mail downloads or transaction based commerce applications.

The GSM air interface soon turned out to be too limited for practical network data and thus the core of the radio interface was upgraded further towards Enhanced Data rates for Global Evolution (EDGE) and Enhanced GPRS, respectively [1]-[5]. By incorporating 8PSK modulation together with fast link adaptation across GPRS air interface, user data rates were significantly increased. EDGE enabled provisioning of new services like e-newspaper, image and sound file transfer, IP-based video telephony and improved end-user experience over e-mail downloads and web browsing. At this point, most of the global mobile communications were enabled by the 2/2.5G technologies but the new applications required much higher data rates than these networks were able to offer. Radio network evolution based on end-user experience for the different applications requirements is illustrated in Figure 2 using two key network performance parameters – bit rate and latency.

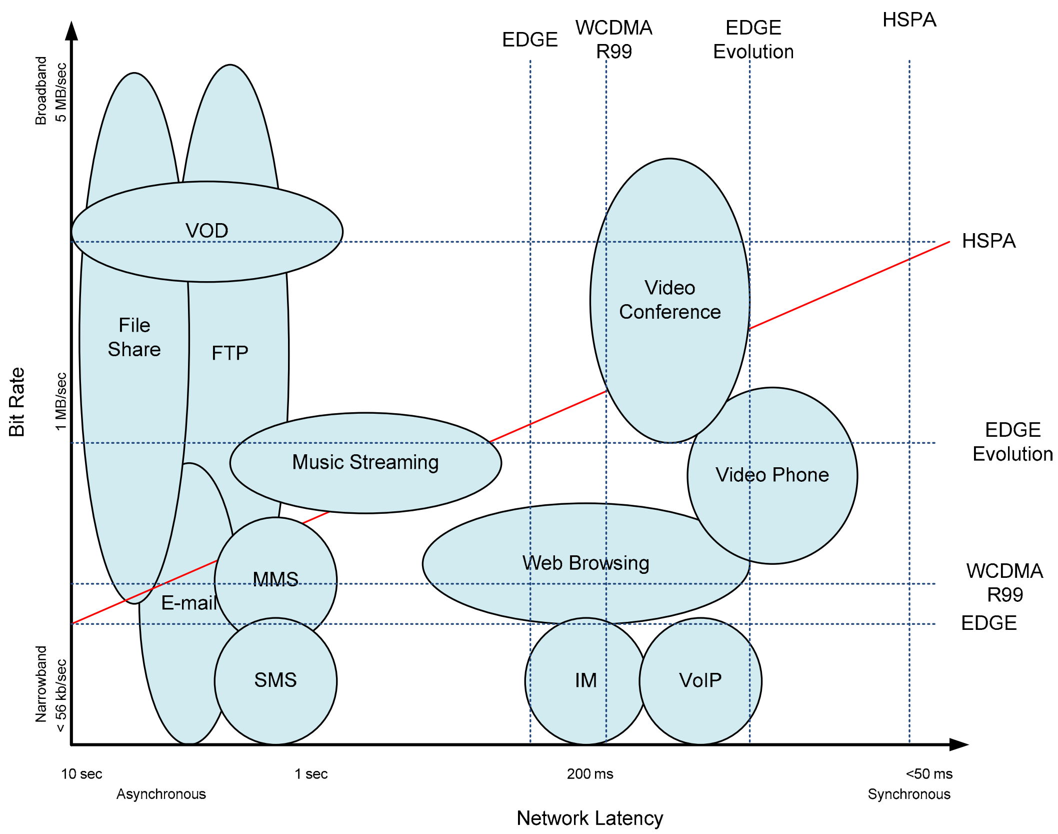


Figure 2: Radio network evolution and end-user experience with different applications requirements.

2.1.2 Third Generation (3G and 3.5G) Systems

The next step came in 1999 with the introduction of Wideband Code Division Multiple Access (WCDMA) air interface for Third Generation systems (3G). Named as Release '99 and standardized by 3GPP, the WCDMA was the most global radio technology in commercial use. The UMTS specified the technical requirements [47] for different operational environments, mobility and architectures, setting the base of the global International Communication Union

(ITU) IMT-2000 standard [48]. The general requirements for high system throughput were met with theoretical data rates of 2 Mbps, while the practical data rates typically range between 0.3 – 1 Mbps. The services provided by WCDMA were similar to those of EDGE, with the additional support of delay-sensitive applications such as VoIP.

Recognition of the multiuser network limits with higher and bursty data rates together with power consumption concerns gave rise to the next evolutionary step to the High Speed Packet Access (HSPA) concept defined in 3GPP Releases 5 – High Speed Downlink Packet Access (HSDPA) and 6 - High Speed Uplink Packet Access (HSUPA). HSPA boosted system capacity and increased user data rates considerably [49], [50]. HSDPA offers 14.4 Mbps peak data rates, while HSUPA reaches 5.76 Mbps. In reality, depending on factors like cell load, the actual peak data rates are again typically much lower than these. With these technologies, however, users can experience better Internet and intranet access, faster file downloads and employment of elementary streaming applications.

The HSDPA concept consists of a new downlink time shared channel that supports a 2-ms transmission time interval (TTI), adaptive modulation and coding (AMC), multi-code transmission, and fast physical layer hybrid ARQ (H-ARQ). The link adaptation (LA) and packet scheduling (PS) functionalities are executed directly from the Node B, which enables them to acquire knowledge of the instantaneous radio channel quality of each user. This knowledge allows advanced packet scheduling techniques that can profit from a form of selection (multiuser) diversity.

The next update from 3GPP came as Release 7 and included the support of higher order modulation - 64 Quadrature Amplitude Modulation (64QAM) and the use of dual-stream Multiple Input - Multiple Output scheme [51]. As a result, the peak data rates reached 42 Mbps and HSPA also provided more efficient and new protocol architecture, physical channels and RRM algorithms. These techniques are in commercial deployments today.

2.1.3 3GPP Long Term Evolution (LTE)

From the radio link spectrum efficiency point of view, HSPA with its enhancements is relatively close to the theoretical maximum defined by Shannon’s law in 5 MHz carrier bandwidth, and assuming single antenna transmission. Further improvements are achieved only by introducing wider carrier bandwidths and advanced antenna systems. On the other hand, going for higher bandwidths with spread spectrum techniques would lead to problems with, e.g., intersymbol interference (ISI) and complex equalization would evolve. The bandwidth scalability requirements would be another challenge.

Clearly, the next upgrade required further change in radio technology and came within 3GPP as Release 8 - Long Term Evolution (LTE), whose radio access is called Evolved UMTS Terrestrial Radio Access Network (E-UTRAN) [52]. LTE targets long term competitiveness and is expected to substantially improve end-user throughputs and sector capacity, reduce user plane latency and provide full mobility. Also, LTE is designed to provide support for IP-based traffic with end-to-end Quality of service (QoS), VoIP support and better integration with other multimedia services. The target aimed at was downlink peak rates of at least 100 Mbps, and uplink of at least 50 Mbps. LTE Release 8 was closed in Spring 2009 and the work continued towards Release 9 increasing the sophistication of LTE [53]. Initial trial deployments of LTE started in December 2009. Release 9 ended in spring 2010 enabling introduction of many new different services and meeting the demands of future LTE users. Currently, development is focused on LTE – Advanced (LTE Release 10 and beyond) to further increase the system performance of radio access networks and to respond to the demands of rapidly growing traffic. More detailed discussion of LTE is given in Section 2.2.

Driven by the market and competitiveness, there has also been intense research to further improve the data rates in HSPA standard. Release 8 (HSPA+) increased HSDPA peak data rate up to 42 Mbps by introducing dual-carrier operation with 10 MHz total bandwidth. Soon after, Release 9 added MIMO support in downlink doubling the peak data rates and introduced 10 MHz bandwidth with 16 QAM in uplink achieving 23 Mbps peak data rates. Similarly as Release 8, Release 10 increased the bandwidth to 20 MHz and doubled the obtained data rates to 168 Mbps in DL. Development of HSPA continues in parallel with LTE –Advanced. A natural next step for HSPA Release 11+ is to further extend the supportable bandwidths to 40 MHz with 4x4 MIMO support achieving theoretical peak rates of 672 Mbps in downlink. In uplink direction, MIMO support and 64 QAM is added achieving 70 Mbps.

Other technologies, like WLAN/WiFi (IEEE 802.11 series) featuring high data rates and low latency connections have also tried to extend towards mobile communications [54]. The technology works well in small cell areas and very low mobility, but its extensions to wide area networks have not succeeded due to mobility, coverage and cell-edge interference problems. Later, the WiMAX (802.16e IEEE standard) appeared, focusing mainly on overcoming the shortcomings of WLAN/WiFi [55], [56]. Even though increased mobility and high data rates were achieved, it resulted in a complex specification and solutions more suitable for fixed wireless links, fixed wireless broadband or complementary technology to cellular networks like WLAN.

The evolution of standards and systems impact on the expansion of the coverage and mobility of practical deployments are illustrated in Figure 3. Moreover, the classical performance measures of coverage and voice capacity have been further extended and converted to new ones in terms of achieved user throughputs with low latency and satisfying certain QoS constraints.

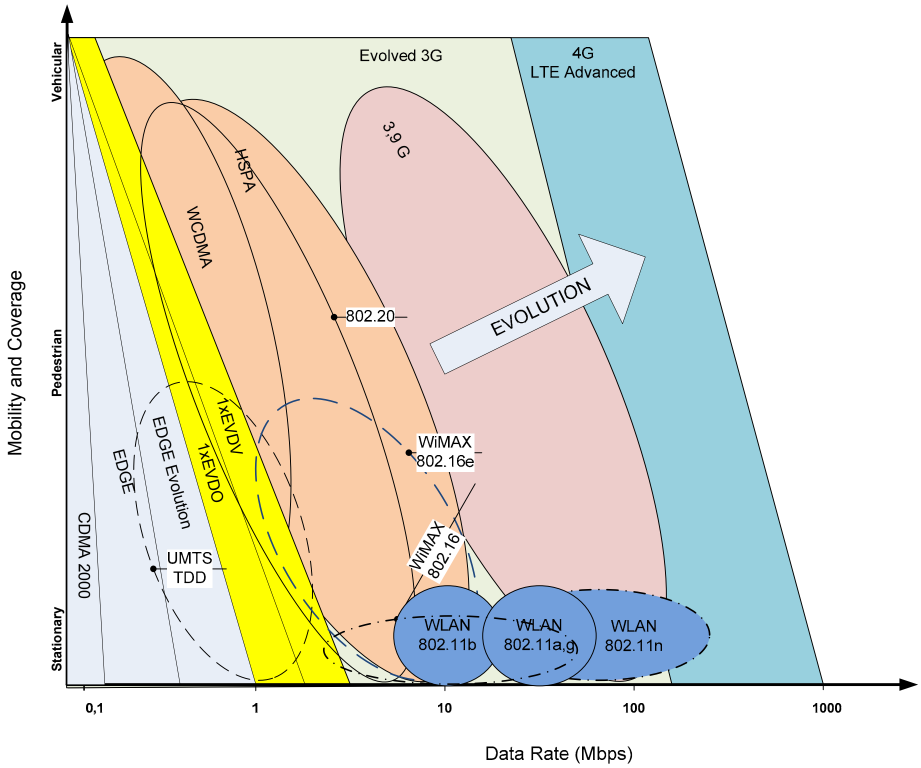


Figure 3: Scope of cellular system evolution.

2.2 3GPP Long Term Evolution

2.2.1 Introduction

As outlined above, LTE is the next major step in mobile radio communications after HSPA and is introduced in 3GPP Release 8. The project started in December 2004 in Study Item (SI) titled "Evolved UTRA and UTRAN". The main objective of this SI was to develop a framework for the evolution of the UTRAN towards a high-data-rate, low-latency and packet-optimized radio-access technology. The SI was needed to certify that the LTE concept could meet the requirements specified in [57], such as the following:

- Increased peak data rates: 100 Mbps downlink and 50 Mbps in uplink
- Scalable bandwidth of 20 MHz, 15 MHz, 10 MHz, 5 MHz, 3 MHz and 1.4 MHz
- IP optimized – packet switched domain only
- Reduced Latency of 50-100 ms for C-plane and less than 10 ms for U-plane

- Improved spectral efficiency (up to four times compared to HSPA Release 6)
- Multi-antenna configuration
- Improved coverage and capacity
- Improved mobility

Another major step in LTE deployment was the introduction of new simplified network architecture to meet the requirements set by the increased data traffic volume, low latency and cost effective operation. Thus, the standard is considered as a milestone towards LTE-Advanced and the full IMT- Advanced capability [58].

2.2.2 System Architecture

System Architecture Evolution (SAE) is designed to optimize network performance by increasing data plane efficiency and minimize the number of nodes facilitating IP-based services. SAE Gateway (GW) is introduced replacing earlier intermediate nodes such as Radio Network Controller (RNC), the Serving GPRS Support Node (SGSN) and the Gateway GPRS Support Node (GGSN). The second node in SAE architecture user plane is the LTE base station (eNode-B) connected to the core network over the so-called S1 interface. The LTE flat system architecture significantly reduces the number of associated nodes in the connections as well as the inter-node data traffic delays. Furthermore, the central control functions are now distributed between Mobility Management Entity (MME) and eNode-B connected through a new logical interface called X2. The eNode-B functionality is greatly extended compared to 3G Node-B. The LTE Radio Access Network (RAN) architecture is presented in Figure 4.

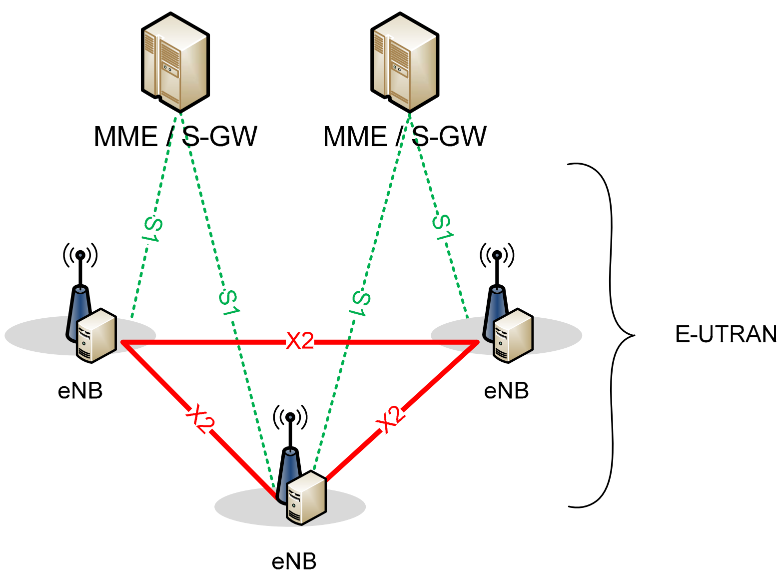


Figure 4: LTE Radio Access Network (RAN) architecture [52].

2.2.3 Overview of LTE air interface

3GPP chose Orthogonal Frequency Division Multiplexing (OFDM) based radio access technology for LTE. OFDM meets the requirements for spectrum flexibility, increased robustness against frequency selective fading or narrowband interference and seamlessly providing cost-effective solutions. The technology is also well-suited for multiantenna operation, is proven to provide high spectral efficiency and is well-established in standards like IEEE 802.11a/b/g, IEEE 802.16, HiperLAN-2, Digital Video Broadcast (DVB) and others. Through the Orthogonal Frequency Division Multiple Access (OFDMA) principle, where different UEs are simultaneously scheduled but on different subsets of subcarriers, multicarrier modulation also enables efficient frequency domain scheduling and link adaptation.

2.2.3.1 Downlink (DL)

The core multicarrier modulation principle consists of using many narrow and mutually orthogonal subcarriers in parallel, which can in principle be data modulated independent of each other. This enables, e.g., channel-dependent adaptive modulation and coding across the subcarriers, and also simplifies channel equalization on the receiver side. In LTE, the baseline subcarrier spacing is 15 kHz, and the available data modulations are QPSK, 16QAM, and 64QAM. A group of subcarriers is called physical resource block (PRB) and consists of 12 subcarriers. In a multi-user setting, meaning OFDMA, individual UE bandwidths are allocated as multiples of PRBs and also the modulation and coding parameters can be adjusted at PRB level, based on the reported channel qualities.

In a practical implementation, an Inverse Fast Fourier Transform (IFFT) block is used for the actual baseband waveform generation from the blocks of subcarrier symbols. Transmitter then inserts a Cyclic Prefix (CP) extension to the time domain multicarrier symbol, which is longer than the channel impulse response. Such a CP eliminates Inter-Symbol Interference (ISI) and also enables simple one-tap equalizers, per subcarrier, on the receiver side.

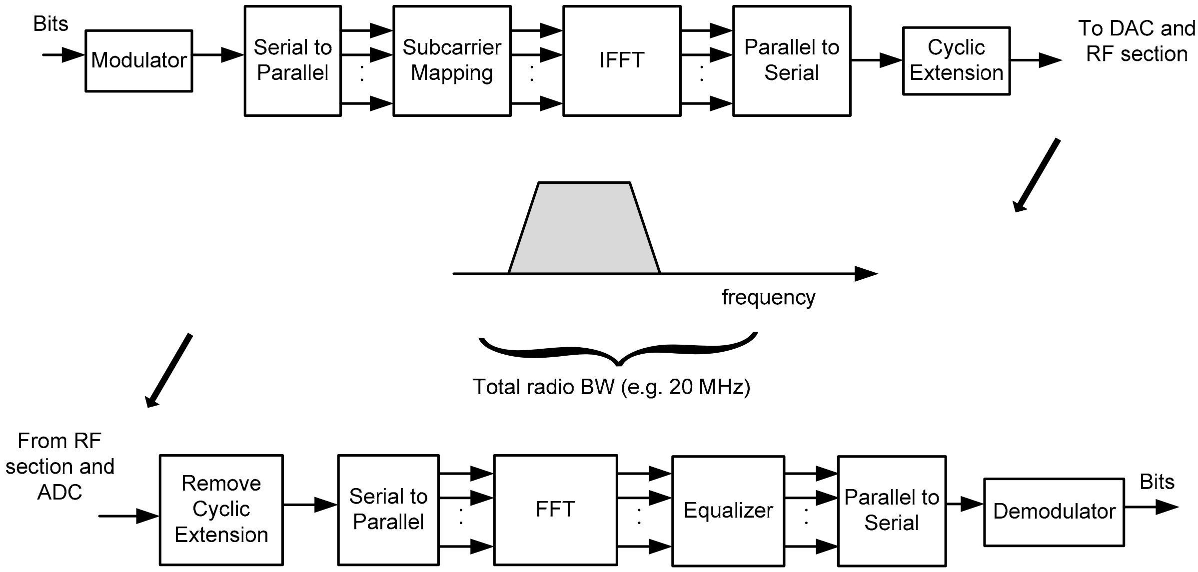


Figure 5: OFDMA baseband transmitter and receiver [2].

At the receiver side as shown in Figure 5, the signal is moved from time domain representation to frequency domain representation by a Fast Fourier Transform (FFT) block. Since cyclic insertion and removal is done at the transmitter and the receiver respectively, ISI, caused by the multipath propagation, is no longer a problem unlike in WCDMA. On the other hand, the receiver has to deal with the impact of a frequency selective multipath channel at subcarrier level, which essentially means correcting for the frequency dependent phase and amplitude changes of the received subcarriers. Thus, subcarrier-wise equalization is performed after FFT to remove the channel impact for each subcarrier utilizing the estimated channel response. The downlink reference symbols, which are placed in both time and frequency domains, are used to estimate the channel frequency response at data subcarriers by interpolating the impact of channel at reference subcarrier symbols.

2.2.3.2 Uplink (UL)

In UL direction, 3GPP decided to use single-carrier frequency division multiple access (SC-FDMA), instead of plain OFDMA, in order to reduce the Peak-to-Average Power Ratio (PAPR) while still enabling efficient frequency–domain equalization at the receiver side stemming from the subcarrier structure. In SC-FDMA, an extra Discrete Fourier Transform (DFT) block is used to move subcarrier data symbols from time domain to frequency domain. Hence the names DFT-spread OFDMA and DFT-precoded OFDMA are also commonly used in this context. After the mapping of resources in the frequency domain, the data symbols are converted to time domain symbols by using IFFT. As in an OFDMA system, the CP is inserted periodically to avoid ISI between SC-FDMA blocks and to enable subcarrier-wise simple equalization.

At the receiver side, cyclic prefix removal and FFT are applied, as shown in Figure 6. After that, subcarrier level channel equalization is deployed. The reference symbols used for channel estimation are located in the middle of each slot. After equalization, Inverse Discrete Fourier Transform (IDFT) operation at the receiver side removes the DFT "precoding" and transforms the signal back to the time domain.

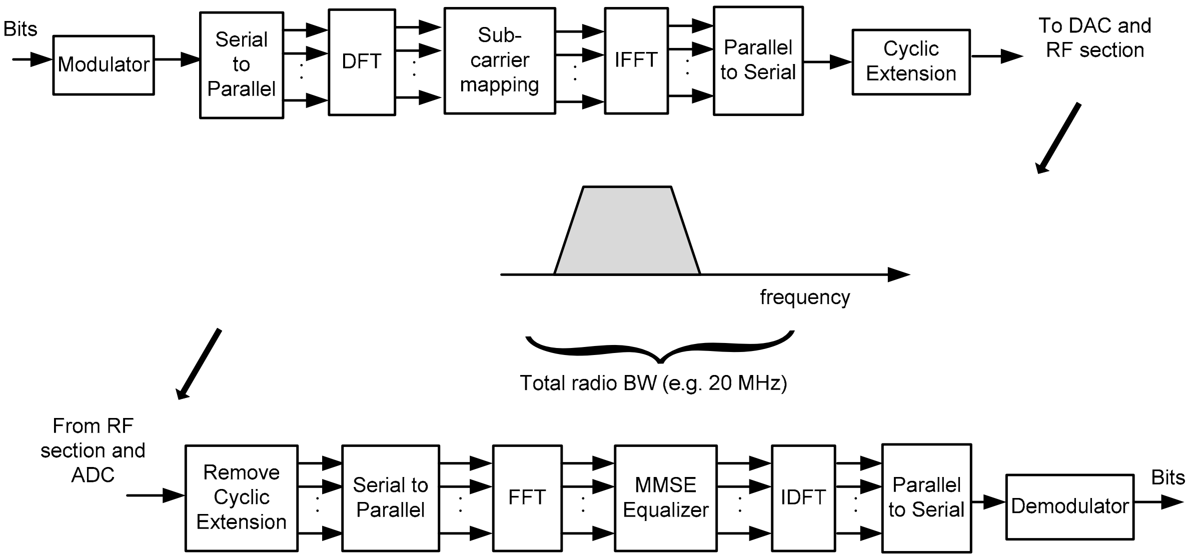


Figure 6: SC-FDMA baseband transmitter and receiver [2].

2.2.3.3 Multi-antenna/MIMO techniques

In the last decade, multi-element antenna arrays have been adopted for reliable communications and higher data rates compared to Single Input Single Output (SISO) systems. The major considered drawback in the deployment of the Multiple Input Multiple Output (MIMO) systems is the increased hardware complexity on one side and the cost on the other side due to multiple expensive RF chains (e.g., low noise amplifiers, mixers and analog to digital converters on the receiver side). Conversely, the increasing demand for higher data rates and the reduction of the capital expenditures (CAPEX) make MIMO technology more favorable from the operator's point of view. Therefore, the first release of LTE standards covers up to 4 antennas while LTE-A will support 8 antennas in order to achieve IMT-A targets.

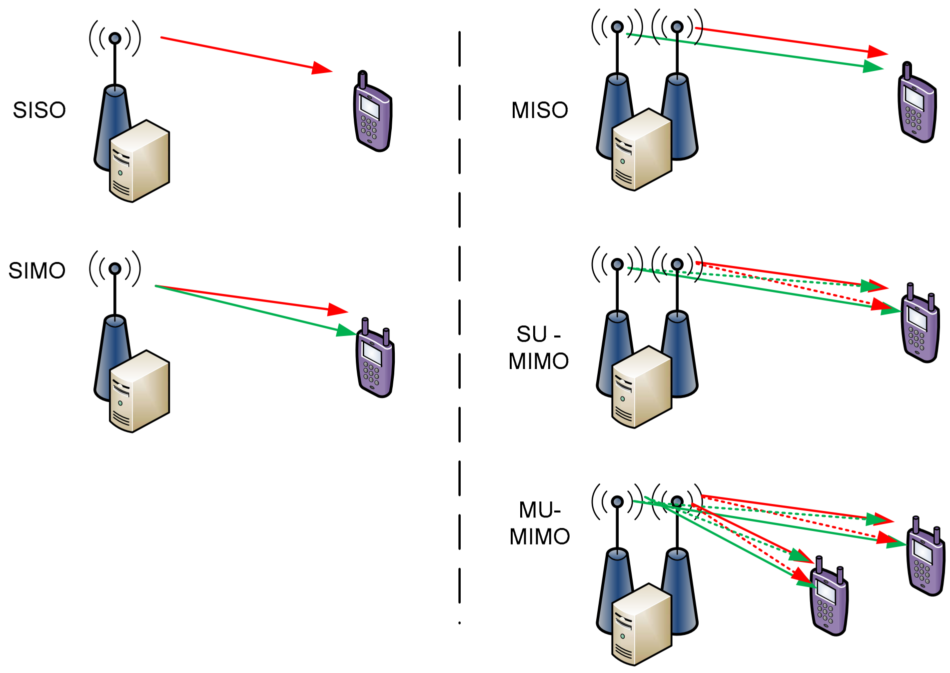


Figure 7: Basic models in multi-antenna systems.

In general, by deploying multiple antennas on the transmitter and receiver sides, the spatial domain also enters the picture, in addition to the classical time and frequency domain. In short, multiple transmit and receive antennas can be used in order to

- create directional properties to transmitting and receiving devices through, e.g., classical antenna array beamforming
- improve the radio link reliability through various diversity methods
- increase the radio link spectral efficiency through spatial multiplexing principles, called single-user (SU) MIMO
- multiplex many user devices into same time-frequency chunk, called space-division multiple access or multiuser (MU) MIMO

In classical beamforming, antenna arrays have directional properties such that the radiated useful signal power is used more efficiently. This is often called array gain. In addition, when combining signals properly on the receiver side with multiple receiver antennas, the useful signal component is combined coherently while noise is combined incoherently. This gives gain against noise in terms of average SNR.

In spatial diversity, on the other hand, the purpose is to improve the radio link reliability by creating multiple independent signal paths between the transmitting and receiving devices. This has the impact of improving the effective distribution of the link SNR/SINR in fading channels and is called diversity gain. Without channel knowledge on the transmitter side, space-time or space-frequency coding (through e.g. Alamouti code [59]) can in practice be deployed to create spatial transmit diversity. Even more simple scenario is receiver diversity where receiving the transmit signal through multiple parallel antennas and receivers, together with appropriate signal combining, yields diversity. LTE downlink supports both space-frequency transmit coding as well as receiver diversity, depending on the transmission mode, starting already from Release 8. If, on the other hand, channel knowledge is already available in the transmitter, the transmit signal can be pre-processed such that the parallel antenna signals combine coherently when arriving in the receiver antenna. This will also yield diversity gain, and is sometimes called closed-loop beamforming. LTE downlink supports closed-loop beamforming starting from Release 9.

In spatial multiplexing, in turn, the idea is to increase the spectral efficiency and transmission throughput by transmitting simultaneously multiple parallel bit streams over a given time-frequency resource. In principle, multiple parallel streams can be coded and modulated independently of each other. Such separate MCS's for different streams or sets of streams are typically called codewords, especially in LTE terminology. LTE downlink supports a maximum of two codewords and 1-2 streams or layers in 2x2 and 4x2 cases, and 1-4 streams or layers in 4x4 case, respectively. When arriving in the receiver, such overlay signals will obviously interfere with each other and stream separation through spatial equalization is needed. Linear minimum mean squared error (LMMSE) spatial equalizer is one of the baseline receivers in LTE performance evaluations, but more advanced receivers stemming e.g. from Maximum Likelihood (ML) detection are also possible. Exact choice of receiver is not specified in the standard and is up to the device manufacturer to decide. In general, if the transmitter has channel knowledge, it can also preprocess the transmit signals to partially assist stream separation in the receiver. This is typically called closed-loop precoding and closed-loop spatial multiplexing. With or without pre-coding, the number of spatially multiplexed streams is upper-limited by the rank of the core propagation channel matrix. The rank, in turn, can at best be the minimum of the number of transmit and receive antennas.

Spatial multiplexing capabilities of MIMO channels can also be used for UE multiplexing. When all the streams belong to one receiver or UE, the corresponding time-frequency slot is used for only one user and hence called SU-MIMO. Then if two or more users are multiplexed on the same time-frequency resource, true spatial domain multiple access takes place and is called MU-MIMO. LTE downlink Release 9 supports maximum of 2 users in MU-MIMO mode.

2.2.4 Summary of LTE Downlink Parameterization and Resource Structure

Table 1 below summarizes the fundamental physical layer parameters for LTE downlink in 5 MHz, 10MHz and 20 MHz cell bandwidth cases.

Table 1: Main physical layer parameters in UTRAN LTE downlink for 5 MHz, 10MHz and 20 MHz cell bandwidth cases

Parameter		**Settings**		
Total transmission BW		5 MHz	10 MHz	20 MHz
Sub-frame duration		1 ms		
Sub-carrier spacing		15 kHz		
Core sampling frequency		7.68 MHz	15.36 MHz	30.72 MHz
FFT size		512	1024	2048
Number of occupied sub-carriers		300	600	1200
Number of OFDM symbols per time slot	Normal CP	7		
	Extended CP	6		
Normal CP length (μs)		5.21 μs 4.67 μs		
Extended CP length (μs)		16.67 μs		

Figure 8 also illustrates the basic physical resource structure [33]. The smallest time and frequency resource is the Resource Element (RE). It consists of 1 OFDM symbol in TD and 1 sub-carrier in FD. Depending on CP length, one time slot has 6 or 7 OFDM symbols where 7 symbols is the basic scenario. Each TTI has duration of 1 ms and then contains 14 OFDM symbols. Physical Resource Block is the smallest resource allocation unit for UE data transmission spanning over 1 ms in TD and 12 adjacent sub-carriers equally spaced at 15 kHz in FD.

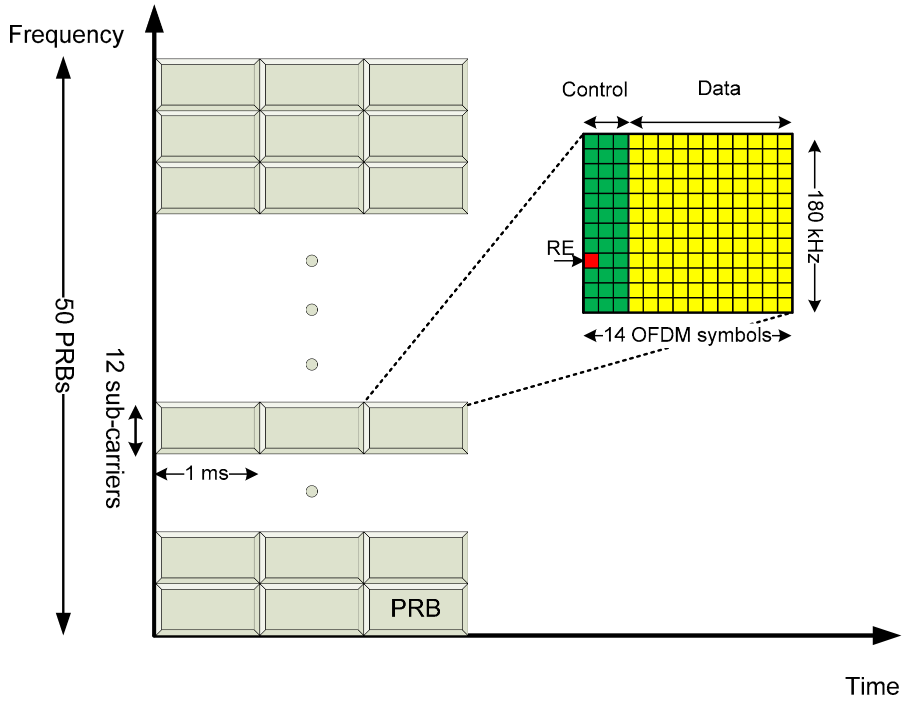


Figure 8: Physical resource structure in E-UTRA downlink.

Chapter 3

RRM in LTE-Downlink type OFDMA Packet Radio Systems – Fundamentals and Proposed Advanced Methods

3.1 Introduction

The RRM functionality targets at the efficient use of radio resources (time, frequency, space) available in a mobile cellular network when multiple users are sharing the resources. In general, it includes scheduling, link adaptation, handover, power control, load control and admission control. In this chapter, selected features of RRM are covered, with emphasis on those which are central to this thesis work.

In general, the efficiency of the RRM process depends on physical resource properties like time and frequency resource scalability and resource allocation resolution. Moreover, RRM has to operate with relatively low signaling overhead, meaning in practice only coarse channel knowledge, but also simultaneously support efficiently many different scenarios (different services, different QoS criteria, different cell sizes, etc) under radio channel dynamics and the involved mobility. In addition to advanced waveforms and radio link level methods, efficient RRM process can be seen as one of the biggest contributors to increased capacities and system-level performance in mobile cellular radio systems. A conceptual illustration of RRM framework in single cell is presented in Figure 9.

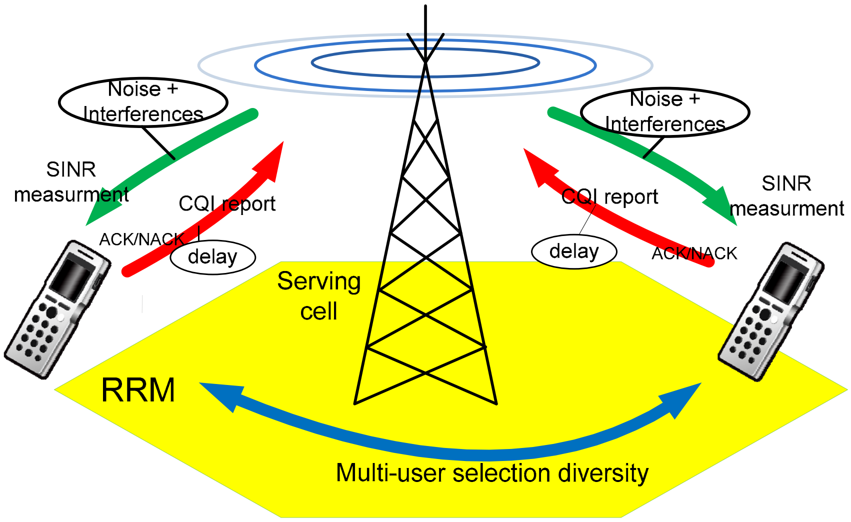


Figure 9: RRM framework in serving cell.

This chapter provides an overview of the main downlink radio resource management functionalities and their interaction in LTE type OFDMA packet radio systems. Section 3.2 presents the general time-frequency domain scheduling framework. Section 3.3 focuses on the Link Adaptation (LA) functionality including Adaptive Modulation and Coding (AMC), Outer Loop Link Adaptation (OLLA) and Inner Loop Link Adaptation (ILLA). Hybrid Automatic Repeat reQuest (HARQ) modeling is described in Section 3.4 while CQI reporting is described in Section 3.5. Sections 3.6 and 3.7 present an overview of the proposed advanced scheduling metrics, covering both best-effort and QoS-oriented scenarios.

3.2 Packet Scheduling

Packet scheduling is the key component of RRM functionality. The task of the Packet Scheduler is to select the most suitable users, per given time window, to access the channel in order to optimize system performance parameters (throughput, coverage, QoS and delay constrains, etc.). Considering systems affected by time and frequency selective fading, the Packet Scheduler can exploit the multi-user diversity by assigning each user to those resources which are experiencing favorable conditions for transmission. To efficiently utilize the limited radio resources, the scheduler thus considers the state of the channel, if available, when selecting the users to be scheduled. Such schedulers are typically called channel-dependent or channel-aware schedulers. This is illustrated at principal level in Figure 10. Searching for optimized solutions of the resource allocation problem for orthogonal multiple access system requires joint optimization over all the available domains (time, frequency and spatial), which typically has very high computational complexity. In other words, a complete search over all possible combinations of transmit parameters and users would be required. Several developments taking such an approach towards maximizing the system capacity for OFDMA system are proposed, e.g., in [29], [41]. Considering subcarrier level granularity, the optimal multi-user resource allocation within one TTI is, however, not a reasonable solution due to complexity reasons [26].

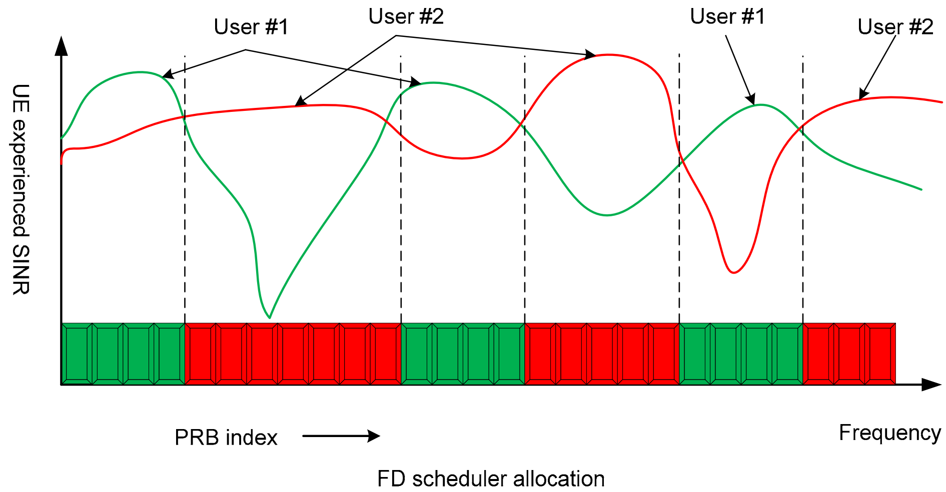


Figure 10: Channel aware frequency domain scheduling principle.

In practice the entities which interact with PS to allocate resources in LTE are QoS parameters, HARQ manager, and LA as shown in Figure 11. To efficiently utilize the limited radio resources, the scheduler should consider the current state of the channel when selecting the user to be scheduled, by utilizing e.g. the ACK/NACK signaling information and CQI reports [5], [37]. Depending on the selected CQI reporting scheme, the accuracy and resolution of the channel quality information can easily vary considerably. In OFDMA based radio systems, like LTE, the CQI information is not necessarily available for all the individual subcarriers but more likely for certain groups of subcarriers only [63].These entities are further introduced in this chapter.

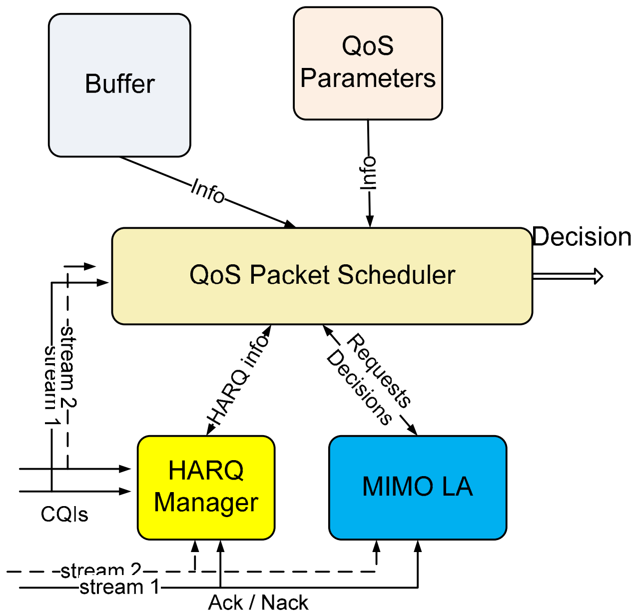


Figure 11: RRM entity interaction.

3.2.1 General Scheduler Principles

The classical scheduler, such as *Round Robin* (RR) [60] is a well known and widely used packet scheduling algorithm. It is very simple and fast giving every user the same amount of channel allocation time in sequential order, without utilizing any channel quality information. Since the scheduling order is decided in advance, the performance in a fast fading environment is naturally not optimal. In TD, the RR scheduler can select user based on priority, QoS or fairness. In FD, RR then maps the RBs to users in a sequential order independent of their radio channel quality. The aim of this algorithm is to assign equal resources to users, i.e. RBs are equally allocated to each user.

A good example of a well performing scheduler is the Proportional Fair scheduler [61], which is one of the key points in this thesis. The Proportional Fair scheduler offers an interesting balance between radio resource allocation fairness among users and overall cell throughput [64]-[66]. It exploits the short-term channel quality variations by scheduling the users at the top of their fades. The scheduler does not choose the user with the best instantaneous channel quality globally but rather the user with the best relative instantaneous channel quality relative to its own recent history. The relative instantaneous channel quality can be derived, e.g., from the relation between instantaneous supportable data rate and the average offered data rate of the user. This approach makes it possible to offer channel resources to users with poor radio conditions (cell edge) and still maintain good data rates for users with good radio conditions.

Another well known scheduling strategy is Max-SINR algorithm [62]. It selects the user with the highest SINR at a given TTI, which is equivalent to the best feasible instantaneous data rate. In this approach, the overall resources are distributed highly unfairly and eventually only the users with very good channel conditions are served. The Max-SINR algorithm maximizes the cell throughput at the expense of user fairness.

In general, the scheduling utility function could be formed from many arguments such as CQI, SINR, Throughputs and higher layer protocol constrains (QoS, delays and other bearers). In most cases it consists of weighting algorithms (time, throughput, etc.) and additional strict priorities. In more formal terms, this corresponds to selecting the user i' from scheduling candidate set of users, indexed here with i, at a time instant (TTI) n to the resources (sub-band) k, for whom:

$$i' = \arg\max_{i} \left\{ P_{i,k}(n) \right\} \tag{3.1}$$

where $P(n)$ denotes the scheduler specific priority metric calculation utility function. Therefore, the user with highest priority metric among all users is granted with resource k during this period n.

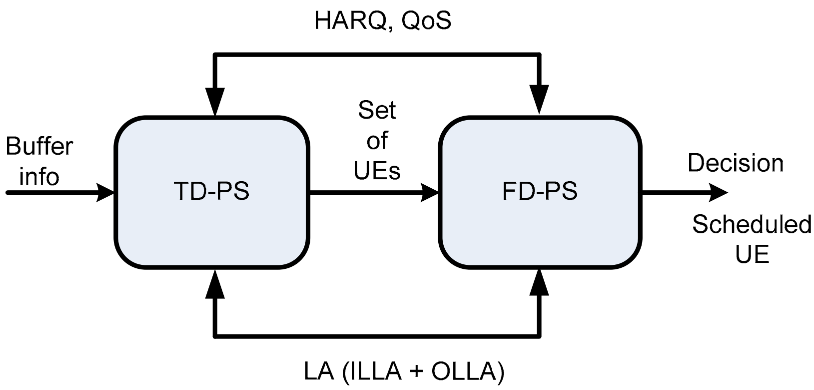


Figure 12: Two-step PS principle.

3.2.2 Novel Two-Step Scheduling Approach

The novel scheduler design approach used in several studies for LTE is a decoupled Time-Domain (TD) - Frequency-Domain (FD) scheduler, as illustrated in Figure 12 [29], [67], [68]. The PS works in two consecutive steps: 1) time domain step and 2) frequency domain step. Such a simplified scheduling principle is beneficial from the complexity point of view, since the FD step considers a reduced number of UEs for frequency multiplexing in each TTI. Thus in the first step, inside each TTI n, all the UE's, say I_{TOT}, are ranked according to a certain scheduling priority metric. In the next step, out of this ranked list of UE's, the first I_{BUFF} ($< I_{TOT}$) UE's with the highest priority metric are picked for the actual frequency domain multiplexing or scheduling stage. The value of I_{BUFF} is set according to the potential channel constraints as well as the available number of PRBs. The role of TD scheduler is to provide the primary mechanism for controlling the QoS, while the role of FD scheduler is mostly to optimize the spectral efficiency per TTI. In such a case the overall scheduler performance will be somewhat sub-optimal due to the limited user diversity at the FD scheduler but the complexity is greatly reduced. Additionally, there can be dependency between the TD and the FD schedulers in many cases, especially with those TD schedulers which depend on the average delivered throughput to users in the past (i.e. dependent on the FD scheduler decision).

In the following, we formulate the core TD and FD metrics used in this work as principal reference metrics against which the more advanced metrics introduced in Sections 3.6 and 3.7 are compared in performance evaluations. The principal scheduling function of the time-domain proportional fair is given by

$$\gamma_i^{td}(n) = \frac{R_i(n)}{T_i(n)} \tag{3.2}$$

where $i = 1, 2, \ldots, I_{\text{TOT}}$ is the UE index, $R_i(n)$ is the estimated throughput of UE i calculated over the full bandwidth allocation for TTI n, and $T_i(n)$ in turn is the corresponding average throughput delivered to the UE i during the recent past and can be obtained by

$$T_i(n) = \left(1 - \frac{1}{t_c}\right) T_i(n-1) + \frac{1}{t_c} R_i^{'}(n-1) \tag{3.3}$$

In (3.3.), t_c controls the averaging window length over which the average throughput is calculated and $R'_i(n\text{-}1)$ denotes the actually realized throughput to the UE i at the previous TTI.

The scheduling function of the frequency-domain proportional fair step with physical resource block index k is then given by

$$\gamma_{i,k}^{fd}(n) = \frac{R_{i,k}(n)}{T_i(n)} \tag{3.4}$$

where $R_{i,k}(n)$ denotes the estimated throughput to the UE i for the *k-th* PRB and $T_i(n)$ is again the corresponding average throughput delivered to the UE i during the recent past given in (3.3).

Finally, access to each PRB resource is granted for the particular user with the highest metric for the corresponding PRB, however, here in the FD step this is evaluated only for those I_{BUFF} users who passed the TD step.

3.3 Link Adaptation

The Link Adaptation (LA) algorithm adapts the modulation and channel coding rate to the instantaneous channel conditions to obtain a proper compromise between the spectral efficiency and reliability in wireless systems [69], [70]. In general, LA may include adaptive modulation and coding (AMC) and power control. Link adaptation functionality requires knowledge of the channel state from the UE side. In the CQI feedback, errors and reporting delay always have a direct impact on the LA process.

In the LTE DL system, Link Adaptation is performed by adaptive modulation and coding (AMC) in both time and frequency domains. Therefore, in frequency domain AMC, modulation and coding scheme (MCS) is selected on PRB basis and might be different for each PRB. The supported modulation schemes in downlink are QPSK, 16QAM and 64QAM. Code rate adaptation is based on adapting and tuning the amount of redundancy in the used Turbo codes. Clearly, the MCS choice depends on received signal quality (SINR) and the choice of coding rate depends on the expected BLER of the decoded transport block and corresponding BLER target.

In a high SINR condition setting a high modulation order may also imply a high coding rate which together try to maximize the instantaneous throughput. In a low SINR condition, on the other hand, setting lower order modulation combined with a low coding rate considerably increases the link robustness and reliability but will obviously decrease the throughput. Overall the approximate range of coding rates is from 0.076 up to 0.93. For simplicity, we have limited the supported coding rates in our studies to 0.33 to 0.8.

Link adaptation requests and decisions are based on cooperation between outer and inner algorithms. The outer loop link adaptation algorithm is needed for scenarios where the CQI feedback from the UEs is subject to errors and reporting delays, among other information for packet scheduling decisions. The actual block error rate for transmissions will tend to be higher than the original anticipated target BLER. In order to achieve target values, OLLA algorithm is needed and should be used. OLLA module is used to maintain the 1st transmission target BLER by adding an adaptive offset to the available CQI reports for the UE based on the ACK/NACKs [28], [71]. The OLLA is shown to be very effective in combating the CQI error. On the other hand, the inner loop link adaptation algorithm makes recommendations to the packet scheduler based on the received feedback from the users. Its functionality is based on adjustments of the received CQIs according to an offset parameter provided by the outer link adaptation algorithm, before using them for important decisions. Based on that, “LA decisions” refers to computing the modulation and coding scheme which gives the best throughput while maintaining the BLER target.

It should be pointed out that there is one OLLA algorithm per user in the cell, so that the ILLA algorithm depending on the UE feedback will use different offsets. Moreover, the flexible MCS adaptation per PRB is a difficult and challenging task to achieve due to signaling constraints, thus only one MIMO mode for one UE within a TTI is supported in this study, which reduces implementation complexity.

3.4 HARQ

HARQ is a transmission scheme that combines an error detection/correction with a retransmission mechanism of the erroneous packet [72]-[74]. Every UE has an individual HARQ entry, which operates the physical layer retransmission functionalities. Typically the received packet is not discarded even when transmission has failure. On the contrary, information is stored and later combined with incoming retransmission(s) on the receiver side. HARQ is based on Stop-And-Wait (SAW) protocol. It keeps on transmitting the current transport block until a positive acknowledgement has been received, before initiating new transmission. The number of HARQ processes plays an important role in scheduling because if the user cannot receive data for every TTI, the number of new transmission UEs to be scheduled is reduced. In order to benefit more of the gain from multi-user diversity in the system, the number is set to 3.

Currently, there are mainly four types of HARQ schemes.

- Conventional ARQ

In this type of ARQ, the packet data is always encoded with a FEC code, such as Turbo codes. In the receiver, if the received packet was found erroneous after decoding, a retransmission of the same packet is requested while the erroneous one is discarded.

- HARQ with Chase Combining (CC)

This type of HARQ scheme adopts a code combining method to reduce the frequency of retransmission. The erroneous packet is not discarded, but is stored at the receiver and combined with the retransmitted packet. The retransmission is the same packet data as the first transmission.

- HARQ with full Incremental Redundancy (IR)

This type of HARQ scheme uses a parity retransmission strategy where instead of sending simple repeats of the coded packet, additional redundant information is incrementally transmitted during the retransmissions. This additional redundancy is combined with the previous received packets for subsequent decoding. The “full” IR means that the retransmissions are not identical with the first transmission but instead additional redundancy bits for error correction are sent.

- HARQ with partial IR

This type of HARQ scheme is a special case of HARQ with full IR. The retransmitted packet is self-decodable, i.e. each packet contains all the information bits necessary for correct reception of the data with different parity bits. So at the receiver, the retransmitted packet may be combined with the previous version if available, or directly fed into the decoder.

Based on earlier studies, for example [75], [76], IR provides better coding gain compared with CC for coding rates higher than ½. The advantages of the CC method are low complexity realization and less memory consumption per UE. Thus, it is preferable to use the CC method with lower coding rates.

Figure 13 presents the operation of 3-Channel SAW protocol used for HARQ, which is the serving one in this dissertation work.

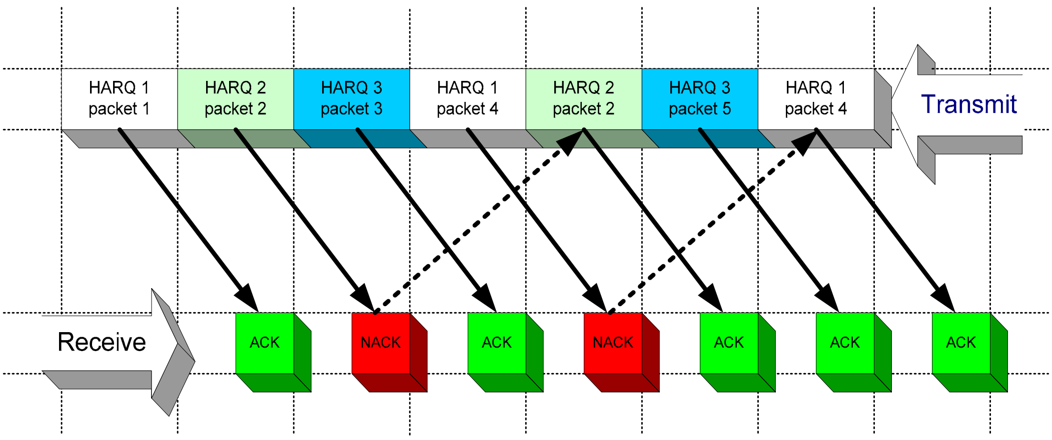


Figure 13: Example of 3-Channel SAW HARQ protocol functionality.

3.5 CQI Reporting

As discussed earlier in Section 3.2, the CQI report needs to provide information on time-frequency resolution of PRBs in order to achieve efficient FDPS. Moreover in Section 3.3, the CQI report is also used for LA functionality. On the other hand, MIMO functionality requires extended stream-wise SINR measures, which causes increased periodic reporting from each UE and thus increased load in uplink channel. We further assume that CQIs for both single- stream and dual-stream MIMO are reported from the UE to eNode-B. Moreover, reports outside the channel coherence time are useless and therefore fairly frequent reporting is needed. Shorter reporting period leads to higher accuracy at the expense of increased reporting overhead. Consequently, the channel coherence time is a function of UE velocity and scheduling gain is significantly reduced at high mobility. More details will be introduced later in Chapter 5.

Clearly, the CQI signaling requirement will increase linearly with the number of scheduling units and the number of UEs. With limited uplink resources, the CQI bandwidth reduction is crucial for the system design. The technique proposed in [34] for time domain PS can reduce the CQI signaling by letting the UE report the CQI only when the channel quality exceeds a certain threshold. The variant of this for FDPS is called “Best- m” scheme in which only the full CQIs for the *m* selected PRBs are reported. To further compress the required signaling, another efficient threshold-based CQI scheme proposed in [36] reduces the CQI signaling overhead for FDPS significantly. These techniques are discussed in the following subsections.

3.5.1 Full CQI Reporting

In a general OFDMA radio system, the overall system bandwidth is assumed to be divided into v CQI measurement blocks. Then quantizing the CQI values to say q bits, the overall full CQI report size is

$$S_{full} = q \times v \tag{3.5}$$

bits which is reported by every UE for each TTI [1], [3], [11].

3.5.2 Best-m CQI Reporting

One simple approach to reduce the reporting and feedback signaling is obtained as follows. The method is based on selecting only m < v different CQI measurements and reporting them together with their frequency positions to the serving cell [11], [33] We assume here that the evaluation criteria for choosing those m sub-bands for reporting is based on the highest SINR values (hence the name best-m). The resulting report size in bits is then given by

$$S_{best-m} = q \times m + \left\lceil \log_2 \left(\frac{v!}{m!(v-m)!} \right) \right\rceil \tag{3.6}$$

3.5.3 Threshold based CQI Reporting

This reporting scheme is a further simplification and relies on providing information on only the average CQI value above a certain threshold together with the corresponding location (sub-band index) information. First the highest CQI value is identified within the full bandwidth, which sets an upper bound on the used threshold window. All CQI values within the threshold window are then averaged and only this information is sent to the BS together with the corresponding sub-band indexes. On the scheduler side, the missing CQI values can then be treated, e.g., as the reported averaged CQI value minus a given dB offset (e.g. 5 dB, the exact number is again a design parameter). The number of bits needed for reporting is therefore only

$$S_{threshold} = q + v \tag{3.7}$$

The threshold-based scheme is illustrated graphically in Figure 14 [36], [40].

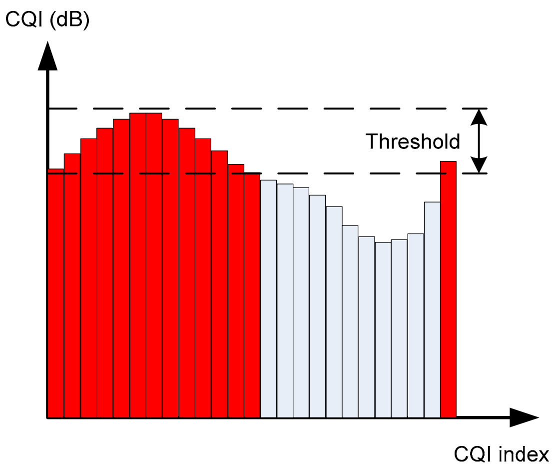


Figure 14: Basic principle of threshold-based CQI reporting.

3.6 Fairness-Oriented Channel-Aware Scheduling and Proposed Metrics

Obtaining a scheduler with increased fairness in resource allocation requires modifications in scheduling priority metric calculation. As a concrete example in this thesis, the utilization of already available CQI reports into PS metric is introduced. Thus, required changes are made in both time and frequency domains.

In basic single stream type Modified PF, the proposed scheduling metric [P1],[P9] is calculated as

$$\bar{\gamma}_i^{td}(n) = CQI_i(n)\left(\frac{T_i(n)}{T_{tot}(n)}\right)^{-1} \tag{3.8}$$

where $CQI_i(n)$ denotes the full bandwidth channel quality report for UE i at TTI n and $T_i(n)$ is as defined in (3.3). $T_{tot}(n)$ is the averaged throughput over the past and over the scheduled users and is obtained by

$$T_{tot}(n) = \left(1 - \frac{1}{t_c}\right)T_{tot}(n-1) + \frac{1}{t_c}\frac{1}{I_{BUFF}}\sum_{i\in\Omega(n-1)} R_i^{'}(n-1) \tag{3.9}$$

In (3.9), $R'_i(n\text{-}1)$ denotes the actual *delivered* throughput for UE i at the previous TTI. The corresponding derivation in FD is obtained by

$$\bar{\gamma}_{i,k}^{fd}(n) = \left(\frac{CQI_{i,k}(n)}{CQI_i^{avg}(n)}\right)^{\alpha 1}\left(\frac{T_i(n)}{T_{tot}(n)}\right)^{-\alpha 2} \tag{3.10}$$

The $CQI_{i,k}(n)$ is the *channel quality* report of user i for sub-band k ($k \in K_{tot}$) at TTI n and $CQI_i^{avg}(n)$ is the corresponding average CQI over the past and over the sub-bands, and can be calculated using

$$CQI_i^{avg}(n) = \left(1 - \frac{1}{t_C}\right)CQI_i^{avg}(n-1) + \frac{1}{t_C}\frac{1}{K_{TOT}}\sum_{k=1}^{K_{TOT}} CQI_{i,k}(n) \tag{3.11}$$

Coefficients $\alpha 1$ and $\alpha 2$ in (3.10) are adjustable parameters and used to conform the influence of either user feedback or throughput term in priority metric. This additional degree of tuning capability is similar to choosing a different forgetting factor $1/t_c$ in ordinary PF scheduler (3.3). Finally, the access to each PRB resource is then granted for the particular user with the highest metric in (3.11) for the corresponding PRB.

Intuitively, the proposed scheduling metrics described above are composed of two elements, affecting the overall scheduling decisions. The first dimension measures the relative instantaneous quality of the individual user's radio channels against their own average channel qualities while the second dimension is related to measuring the achievable throughput of individual UE's against the corresponding average throughput of all scheduled users. In this way, and by understanding the power coefficients α1 and α2 as additional adjustable parameters, the exact scheduler statistics can be tuned and controlled to obtain the desired balance between the throughput and fairness. This will be demonstrated later in Chapter 5.

The dual-stream type Modified PF scheduling metric proposed in this dissertation work, and formulated next, takes into account the extra spatial domain functionality from MIMO technique. Considering SU- and MU-MIMO operation the new stream-wise priority metric building on (3.10) is calculated as

$$\overline{\gamma}_{i,k,s}(n) = \arg\max_{i}\left\{\left(\frac{CQI_{i,k,s}(n)}{CQI_{i}^{avg}(n)}\right)^{\alpha 1}\left(\frac{T_i(n)}{T_{tot}(n)}\right)^{-\alpha 2}\right\} \tag{3.12}$$

The individual users channel quality reports are now calculated for each stream s and therefore the average CQI calculation takes into account the maximum number of streams, say S_{TOT}. Thus, equation (3.11) is modified as

$$CQI_i^{avg}(n) = \left(1-\frac{1}{t_c}\right)CQI_i^{avg}(n-1) + \frac{1}{t_c}\frac{1}{K_{TOT}}\frac{1}{S_{TOT}}\sum_{s=1}^{S_{TOT}}\sum_{k=1}^{K_{TOT}} CQI_{i,k,s}(n) \tag{3.13}$$

A potential problem is that multiple MIMO modes can be selected for a single user, i.e., both dual-stream and single-stream mode may be selected for the same user on different PRBs, within one TTI. On the other hand, this is in contradiction with the implementation feasibility and constraint we made in Sections 3.3 and 3.5. The problem is prevented by the following simple and efficient approach proposed for these two cases. If there is a conflict, a comparison is performed of the total throughput from all the allocated single-stream PRBs for this user with that from all the allocated dual-stream PRBs. The UE is then forced to use the MIMO mode which gives better total throughput. Considering the MU-MIMO case, the decision is made for each dual-stream PRB based on partner UE allocation on the same frequency resource element. A special case here is if the single-stream mode is favored for a user in which case a check is then made for the other stream assignment. If the same user is selected on both streams, this user is forced to single-stream for this frequency resource element, stemming from the previous discussion. Otherwise, this stream is assigned to a partner user. Once the assignment of PRBs to users has been performed, the scheduler asks the LA to calculate the supported data rate for each user, also taking the selected MIMO mode into account. Even though the obtained system performance is suboptimal, the algorithm design is efficient and suitable for practical implementations from a complexity perspective.

The next step in metrics evolution originates from considering soft frequency reuse (SFR) scheme in OFDMA network deployments. The SFR scheme reserves part of the frequency band for the cell-edge users and uses the power bound specified for it by the power mask. The rest of the unallocated sub-bands are dedicated to the near-to-BS users. The challenge of obtaining UE fairness is mitigated by introducing modified power aware multi-stream PF (MPMPF) scheduling technique [P3], [P4], [P11]. The PS priority metric function is extended by the ratio of individual steam-wise sub-band power over the maximum transmission power for any sub-band. Thus, the new metric is defined as

$$\overline{\gamma}_{i,k,s} = arg\,\underset{i}{max}\left\{ \frac{P_{k,s}}{P_{\max}} \left(\frac{CQI_{i,k,s}(n)}{CQI_i^{avg}(n)} \right)^{\alpha_1} \left(\frac{T_i(n)}{T_{tot}(n)} \right)^{-\alpha_2} \right\} \tag{3.14}$$

System-level performance illustrations of the proposed metrics will follow in Chapter 5, while the original publications contain yet more detailed examples and performance studies.

3.7 QoS- and Channel-Aware Scheduling and Proposed Metrics

One of the main targets in next generation networks is to provide seamless access to voice and multimedia services with end-to-end Quality of Service (QoS). The efficient QoS control requires QoS-aware PS and AC. This thesis mainly focuses on implementation of QoS-aware PS, while simple AC algorithm is used to emphasize the PS framework. Further, this framework is shown to meet the QoS targets, taking into account Constant Bit Rate (CBR) traffic scenarios with Guaranteed Bit Rate (GBR) constraints.

The actual metric calculation of the proposed QoS-aware multi-stream PF (QoS-MSPF) is defined as [P5]

$$\overline{\gamma}_{i,k,s} = arg\,\underset{i}{max}\left\{ \eta_i(n) \left(\frac{CQI_{i,k,s}(n)}{CQI_i^{avg}(n)} \right)^{\alpha_1} \left(\frac{T_i^{se}(n)}{T_{i,k,s}(n)} \frac{T_i(n)}{T_{tot}(n)} \right)^{-\alpha_2} \right\} \tag{3.15}$$

The $\eta_i(n)$ is, in turn, a QoS specific factor defined as [42]

$$\eta_i(n) = 1 + \xi e^{\left[-\psi\left(T_i(n) - GBR_i\right)\right]} \tag{3.16}$$

The parameters ξ and ψ are used to control user traffic requirements. GBR_i corresponds to the desired user throughput in this context.

Similarly to the previously discussed metrics in Section 3.6, the scheduling metric in (3.15) is essentially composed of different parts or elements affecting the overall scheduling decisions. The first term takes into account the QoS requirements defined in (3.16). The second ratio is the individual user over average CQI, interpreted in a stream-wise manner and the average CQI is calculated as in (3.13). The third ratio consists of estimated throughputs of individual UE's and achievable over total throughputs.

Considering next actual video traffic models, the used QoS constraints are mainly packet size, arrival rate and packet delay. Therefore, in QoS-MSPF scheduling metric, only new QoS delay function factor $\delta_i(n)$ is replacing $\eta_i(n)$ in equation (3.15), which is defined as:

$$\delta_i(n) = \frac{\max(d_i / B_i(n))}{d_{\max}} \tag{3.17}$$

Here d_i is the delay of the packet of user i in the transmit buffer $B_i(n)$ at time instant n, and $d_{\max}$ is the maximum delay allowed. Consequently, the equation (3.15) will transform to:

$$\overline{\gamma}_{i,k,s} = \arg\max_i \left\{ \delta_i(n) \left(\frac{CQI_{i,k,s}(n)}{CQI_i^{avg}(n)} \right)^{\alpha_1} \left(\frac{T_i^{se}(n)}{T_{i,k,s}(n)} \frac{T_i(n)}{T_{tot}(n)} \right)^{-\alpha_2} \right\} \tag{3.18}$$

Again, intuition is similar as earlier. That is, the first term is QoS-related while the other two emphasize a tunable compromise between fairness and channel quality.

Chapter 4

System-Level Performance Evaluation Methods and Simulations

4.1 Link Level and System Level Analysis

In general, simultaneous analytical treatment and analysis of all the processes involved in the operation of a wireless mobile cellular radio network is not a feasible task. One common approach is thus to divide the modeling and analysis to radio link and radio system levels, and to use numerical simulations combined with elementary process modeling. Such an approach is a practical trade-off between the analysis accuracy due to the modeling simplifications and the problem complexity. Moreover, performance gains that may be achieved on a single radio link (link level), do not necessarily always translate into the same or relative gains at system level, where multiple base stations communicate with multiple users and share the common radio resources. Thus understanding the performance limitations, and possible interactions, can be more clear when the individual link and the overall system can be analyzed first separately and then combined.

In general, link-level analysis and simulations are suitable for developing, e.g. receiver structures or coding schemes and therefore it is not possible to demonstrate the effects of scheduling, traffic modeling, or inter-cell interference using such approaches. Moreover, simulating all the radio links at exact waveform level between the User Equipments (UEs) and eNodeBs is an unreasonable way of performing system level simulations due to the immense amount of computational power that would be required [77]. Therefore, a more practical way in system-level analysis and associated simulations is that the physical layer is abstracted by simplified models that capture its essential characteristics with high accuracy and with reasonable complexity. The operation of such approach is that the link-level processing is abstracted in lookup tables with signal-to-interference plus noise ratio (SINR) traces from link-level simulations, while the related system-level processing is modeled in a network overlay including system-level parameters, as illustrated in Figure 15. Usually, the performance of the individual radio links has been evaluated in terms of the block error rate (BLER) as a function of SINR, averaged over all channel realizations of one specific channel model.

In order to perform a realistic evaluation of the enhancements achieved by advanced RRM functionalities, system-level analysis and simulations are thus used in this work. Hence, modeling the performance of fast scheduling (in time, frequency, and/or space domain), fast link adaptation in the form of adaptive modulation and coding (AMC), or other advanced schemes such as hybrid ARQ, is impossible if only averaged parameters (e.g., pathloss and shadowing) are modeled at system level. Therefore, system-level simulations should also include modeling of small-scale/fast fading effects, at least to an extent that is relevant from RRM perspective. This is also the approach adopted in this work.

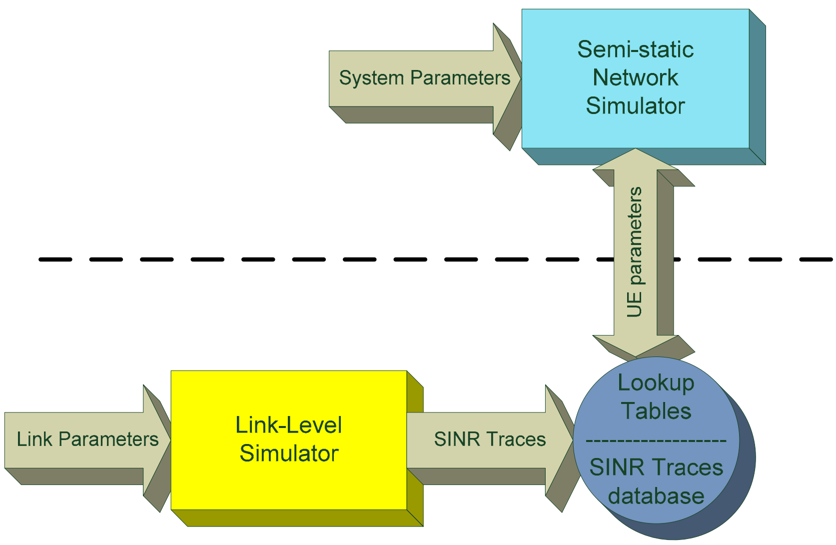


Figure 15: Link and system simulation methodology.

System-level simulations can in general be further divided into quasi-static (semi-static) and dynamic. In the quasi-static simulator, snapshot-like simulations are conducted for a relatively short time with close to static UEs and fast fading effects are taken into account by modeling the Doppler effects in the radio channels and propagation but excluding large-scale user movements and therefore also handovers. In dynamic simulators, in turn, user mobility is the key point and simulations are conducted for a longer time which then also typically imposes, e.g., handovers. In our work, the main emphasis has been on quasi-static performance analysis.

4.2 Network Simulator Modeling

The implemented quasi-static system simulator, developed and used within this thesis work, provides traffic modeling, multiuser scheduling, and link adaptation including Chase Combining HARQ under the UTRAN LTE downlink parameters and assumptions described in [3].

The overall simulation flow is as follows:

1. Initialization – Create all static objects (scenario based eNode B and UEs)
2. Warm up – Generate initial calls for initial network load
3. Simulation run – Calculation of system performance statistics (UE throughputs, SINRs, etc.)
4. Finish – save all data and clear the objects

The modeling details of modules such as Link Adaptation, CQI reporting, HARQ, Link to system mapping and traffic models are described in more details below.

4.2.1 Network Topology and Radio Propagation Environment

The used network topology consists of a hexagonal regular grid cellular setup, where the center three cells are surrounded by the two tiers of cells, as shown in Figure 16. There are nineteen cell sites in the simulation area, each consisting of three sectors per site, giving a total of fifty seven sectors.

Moreover, the 3-sector network topology is assumed for the Macro cell deployment and 1-sector network topology is assumed for the Micro cell deployment. The propagation modeling consists of the path loss, shadowing and fast fading. The path loss model for the Macro cell case includes a 20 dB outdoor-to-indoor penetration loss and 0 dB for the Micro cell case. Fast fading is simulated according to the Typical Urban (TU) power delay profile for different user speeds, and typical tapped delay line implementation with uncorrelated Rayleigh fading paths is deployed.

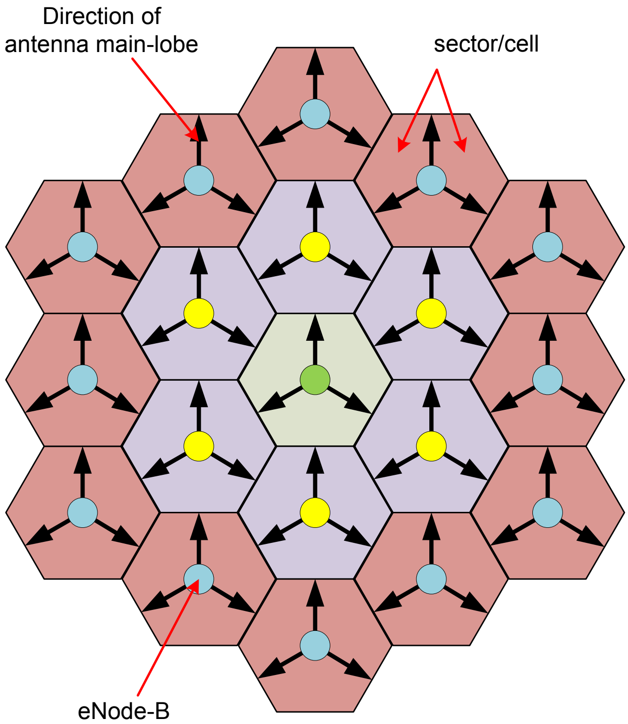


Figure 16: Assumed regular hexagonal cellular grid [3].

4.2.2 User Generation and Distribution

At the beginning of a simulation run, UEs are uniformly dropped into the entire cell. The minimum distance between the UE and eNode-B used in the simulations is 35 meters. If it happens that a user is dropped within this 35 meter circle buffer zone, then the user will be redropped again outside the zone. During one simulation run the location of the mobile is unchanged but fast fading is modeled according to classical Jake's Doppler spectrum, once per TTI (1ms).

4.2.3 Traffic Models

Infinite (full) buffer traffic model is considered to represent the behavior of best effort services. It is the simplest traffic model for system-level evaluation, in which the users always have data packets to transmit (if scheduled). Moreover, all users experience equal session time irrespective of their location within the cell and UEs close to the eNode-B download a much larger amount of data in comparison to those located near the cell edge (despite typical fairness measures). Specifically for this model, the cell and user throughput statistics are collected over a large number of simulation runs of the same duration.

In addition to full buffer scenario, streaming traffic scenarios are also considered and are used as a constant bit rate (CBR) source model. For each CBR user, a fixed amount of data packets is generated at the source with a constant packet size and constant inter-arrival time. The streaming condition requires that the amount of traffic is in equilibrium, i.e., the number of packets transmitted within a given delay constraint equals the average achieved bit rate. Therefore, video streaming and conversational video telephony are rarely used in the system analysis, even if their models exist in the literature. Video codecs typically generate regular frames with complete picture information and a sequence of intermediate frames with differentially coded information. Each video frame consists of a different number of variable packet sizes. The inter-arrival process is rather deterministic because of the constant frame rate, but the number and size of packets may vary.

The delay constraint may be fairly large for streaming (up to second), and traffic smoothing is required by a playout buffer at the receiver. Any transmit and scheduling delay variation and the consequent receive delay variation observed at the input of the playback buffer will not be present at the output of that buffer. Therefore, these variations do not decrease the user experience.

4.2.4 Detectors and SINR Modelling

One central element in the modeling is related to effective SINR of the individual radio links. This depends overall, e.g., on the propagation characteristics, neighboring cell interference levels, receiver noise floor, and deployed transmitter and receiver signal processing. Here, for simplicity, we assume a two-antenna transmitter and receiver case.

The actual effective SINR calculations rely on subcarrier-wise complex channel gains (estimated using reference symbols in practice) and depend in general also on the assumed receiver (detector) topology. Concerning the actual UE receiver topologies (spatial filters), maximum ratio combining (MRC) and LMMSE spatial equalization based receivers are assumed. The LMMSE receiver is sometimes also called interference rejection combining (IRC) receiver in the literature when explicit local interference covariance is used in building the LMMSE receiver. The MRC receiver is applicable in a TX antenna-selection based single-stream single-user case only, while the LMMSE receiver is then more general receiver topology that is also used in dual-stream SU and MU cases, respectively, and properly tailored depending on the transmission mode (1-stream SU, 2-stream SU or 2-stream MU). The detector structures and SINR modeling for different transmission modes are described in detail below.

4.2.4.1 Single-Stream SU Case

In this case, only one of the two BS transmit antennas is used to transmit one stream. At individual time instant (time-index dropped here), the received spatial 2x1 signal vector of UE i at sub-carrier c is of the form

$$\mathbf{y}_{\mathbf{i,c}} = \mathbf{h}_{\mathbf{i,c}} x_{i,c} + \mathbf{n}_{\mathbf{i,c}} + \mathbf{z}_{\mathbf{i,c}} \tag{4.1}$$

where $x_{i,c}$, $\mathbf{h}_{i,c}$, $\mathbf{n}_{\mathbf{i,c}}$ and $\mathbf{z}_{\mathbf{i,c}}$ denote the transmit symbol, 2x1 channel vector, 2x1 received noise vector and 2x1 inter-cell interference vector, respectively. Then the LMMSE detector $\hat{x}_{i,c} = \mathbf{w}_{i,c}^H \mathbf{y}_{i,c}$ is given by

$$\mathbf{w}_{i,c} = \sigma_{x,i}^2 (\sigma_{x,i}^2 \mathbf{h}_{i,c} \mathbf{h}_{i,c}^H + \mathbf{\Sigma}_{n,i} + \mathbf{\Sigma}_{z,i})^{-1} \mathbf{h}_{i,c} \tag{4.2}$$

where $\sigma_{x,i}^2$, $\mathbf{\Sigma}_{n,i}$ and $\mathbf{\Sigma}_{z,i}$ denote the transmit power (per the used antenna), noise covariance matrix and inter-cell interference covariance matrix, respectively.

Now the SINR is given by

$$\gamma_{i,c} = \frac{\left|\mathbf{w}_{i,c}^H \mathbf{h}_{i,c}\right|^2 \sigma_{x,i}^2}{\mathbf{w}_{i,c}^H \mathbf{\Sigma}_{n,i} \mathbf{w}_{i,c} + \mathbf{w}_{i,c}^H \mathbf{\Sigma}_{z,i} \mathbf{w}_{i,c}} \tag{4.3}$$

into which the detector in (4.2) can be substituted. The noise variables at different receiver antennas are assumed to be uncorrelated (diagonal $\mathbf{\Sigma}_{n,i} = \sigma_n^2 \mathbf{I}$) and independent of the user index i. The more detailed modeling of inter-cell interference (structure of $\mathbf{\Sigma}_{z,i}$) is given by

$$\mathbf{\Sigma}_{z,i} = \sum_{l=1}^{L_{int}} \sigma_{int,l,i}^2 \mathbf{g}_{l,i,c} \mathbf{g}_{l,i,c}^H \tag{4.4}$$

where L_{int} is the number of interfering base-stations, $\mathbf{g}_{l,i,c}$ are the corresponding path gain vectors and $\sigma_{int,l,i}^2$, denotes the nominal interferer transmit power per antenna and per interference source (l) [17].

The MRC detector, in turn, is given by

$$\mathbf{w}_{i,c}^{MRC} = \frac{\mathbf{h}_{i,c}}{\|\mathbf{h}_{i,c}\|^2} \tag{4.5}$$

In general, this represents only coherent combining of the received spatial samples and does not explicitly consider receiver interference conditions. The corresponding SINR is obtained by substituting (4.5) into (4.3).

4.2.4.2 Dual-Stream SU Case

In this case, both of the two BS transmit antennas are used for transmission, on one stream per antenna basis. At individual time instant, the received spatial 2x1 signal vector of UE i at sub-carrier c is now given by

$$\mathbf{y}_{i,c} = \mathbf{H}_{i,c}\mathbf{x}_{i,c} + \mathbf{n}_{i,c} + \mathbf{z}_{i,c} \tag{4.6}$$

where $\mathbf{x}_{i,c}$ and $\mathbf{H}_{i,c} = [\mathbf{h}_{i,c,1}, \mathbf{h}_{i,c,2}]$ denote the 2x1 transmit symbol vector and 2x2 channel matrix, respectively.

Now the LMMSE detector $\hat{\mathbf{x}}_{i,c} = \mathbf{W}_{i,c}\mathbf{y}_{i,c}$ is given by

$$\mathbf{W}_{i,c} = \Sigma_{x,i}\mathbf{H}_{i,c}^H(\mathbf{H}_{i,c}\Sigma_{x,i}\mathbf{H}_{i,c}^H + \mathbf{\Sigma}_{n,i} + \mathbf{\Sigma}_{z,i})^{-1} = \begin{bmatrix} \mathbf{w}_{i,c,1}^H \\ \mathbf{w}_{i,c,2}^H \end{bmatrix} \tag{4.7}$$

where $\Sigma_{x,i} = diag\left\{\sigma_{x,i,1}^2, \sigma_{x,i,2}^2\right\} = diag\left\{\sigma_{x,i}^2/2, \sigma_{x,i}^2/2\right\}$ denotes the 2x2 covariance matrix (assumed diagonal) of the transmit symbols. Note that compared to the single-stream case, the overall BS transmit power is now divided between the two antennas, as indicated above. Then the SINRs for the two transmit symbols are given by

$$\begin{aligned} \gamma_{i,c,1} &= \frac{\left|\mathbf{w}_{i,c,1}^H\mathbf{h}_{i,c,1}\right|^2 \sigma_{x,i,1}^2}{\left|\mathbf{w}_{i,c,1}^H\mathbf{h}_{i,c,2}\right|^2 \sigma_{x,i,2}^2 + \mathbf{w}_{i,c,1}^H\Sigma_{n,i}\mathbf{w}_{i,c,1} + \mathbf{w}_{i,c,1}^H\Sigma_{z,i}\mathbf{w}_{i,c,1}} \\ \gamma_{i,c,2} &= \frac{\left|\mathbf{w}_{i,c,2}^H\mathbf{h}_{i,c,2}\right|^2 \sigma_{x,i,2}^2}{\left|\mathbf{w}_{i,c,2}^H\mathbf{h}_{i,c,1}\right|^2 \sigma_{x,i,1}^2 + \mathbf{w}_{i,c,2}^H\Sigma_{n,i}\mathbf{w}_{i,c,2} + \mathbf{w}_{i,c,2}^H\Sigma_{z,i}\mathbf{w}_{i,c,2}} \end{aligned} \tag{4.8}$$

where again the detector filters from (4.7) can be substituted.

4.2.4.3 Dual-Stream MU Case

In this case, the transmission principle and SINR modeling are similar to the subsection above, but the two spatially multiplexed streams now belong to two different UEs, say $i^{(1)}$ and $i^{(2)}$. Thus the signal model in (4.6) is essentially repeated twice, once per receiving UE, and also the channel responses are then totally independent processes (one channel process per UE). The UE detector recovers, of course, only the particular stream that is transmitted to the receiving UE and neglects the other stream. The detectors and SINRs in (4.7) and (4.8) are then also interpreted accordingly.

4.2.5 CQI Model

As discussed in Sections 3.2, 3.3 and 3.5, channel quality measurements and reporting are essential for improved use of radio resources. In this work, the CQI concept stands for received SINR measurements and calculations, observed by the UE, including assumed receiver post-processing. The detailed expressions for these post-processing SINRs, in both single-stream and dual-stream transmission scenarios, are given in equations (4.3) and (4.8), as functions of complex spatial channel gains, noise and interference covariance and receiver post-processing.

In practice, the receiver uses reference or pilot symbols for extracting information about the channel state. In our work, we follow directly LTE Release 8 reference symbol structure, which means that reference symbols are antenna-specific in the sense that when one base-station antenna transmits a reference symbol (at a given subcarrier), the other antennas are muted. This is then repeated periodically over the other antenna elements. This means that UE is able to estimate the antenna-specific spatial channels (spatial channel from each transmit antenna to all receiver antennas), and also the noise and interference powers when appropriate alignment and processing are applied between cells. This then directly enables the SINR calculations, following equations (4.3) and (4.8) and thereon CQI reporting. The actual CQI reports are quantized SINR values, and are here assumed to be composed at PRB level and reported from UE to E-NodeB using a signaling channel.

In practice, the UEs cannot estimate perfectly the CQI and therefore the reported CQI is subject to measurement errors and to quantization effects since the CQI needs to be reported with a small finite number of bits. This was discussed conceptually and quantitatively in Section 3.5. The accuracy of the CQIs has a significant impact on the potential gain from using frequency domain link adaptation and packet scheduling. In this thesis work, to model measurement errors, we use the basic approach of adding a Gaussian distributed measurement error to the ideal CQI expressed in decibel scale. Furthermore, as explained in Section 3.5, we quantize the CQI to a finite resolution. Thus in the general case, the CQI for k-th PRB on stream s is modeled here as:

$$CQI(k,s) = Quantize\left\{Q; 10\log_{10}\left[CQI_{ideal}(k,s)\right] + \delta(k)\right\} \tag{4.9}$$

where $\delta(k)$ is a zero mean Gaussian distributed variable with standard deviation 1 dB representing the measurement errors in the CQI (received SINR) and Q is quantization step. The recommended value of $\delta(k)$, used also in the thesis work, is 1 dB [78].

As discussed in Chapter 3, the full CQI and alternative reduced feedback schemes are used in the actual work of this dissertation. Considering the full CQI reporting scheme in the case of LTE, with 10 MHz system bandwidth and grouping 2 physical resource blocks (PRBs) into 1 measurement block, it follows that we have totally 25 measurement blocks. Assuming further that quantization is carried with 5 bits, then according to (3.5) each UE is sending 125 bits for every 1ms (TTI length). Similarly, for best-m reporting scheme the required bits for best-10 blocks are 72 bits following from (3.6). Finally, a threshold-based CQI reporting scheme requires only 30 bits, if using the same parameters, which follows from (3.7). Consequently, MIMO mode selection will require 3 times more bits if always one CQI per possible mode is reported.

4.2.6 LA

In OFDMA systems, Link Adaptation is performed through AMC. UEs typically send different feedback reports and other information for the needs of the link adaptation mechanism. Decisions to be made are based on data packet acknowledgement (ACK/NACK), CQI information or a combination of them. The LA functionality is based on OLLA and ILLA algorithms, discussed in detail in Section 3.3.

The OLLA algorithm used in this work is based on the well-known outer loop control algorithm, used originally for adjusting SIR targets for dedicated WCDMA channels [72]. The main target is controlling the BLER target for the first transmission to users by adjusting the offset factor Λ to ILLA according to following rules:

- If an ACK is received for a first transmission, then decrease the offset factor Λ by Λ_{Down} decibels.
- If a NACK is received for a first transmission, then increase the offset factor Λ by Λ_{Up} decibels.

The ratio between the parameters Λ_{Down} and Λ_{Up} determines the BLER target that the algorithm will converge to, i.e.,

$$BLER = \frac{\Lambda_{Down}}{\Lambda_{Up} + \Lambda_{Down}} \tag{4.10}$$

The OLLA algorithm for adjusting Λ is recursive and requires the initial value of Λ which in our work is set to 0.5 dB [79]. A further imposed constraint is bounding the offset factor Λ in the range of [-1, 3] dB.

4.2.7 HARQ

The HARQ processes are explained in details in Section 3.5. The explicit scheduling of multiple HARQ processes per user has not been performed due to complexity reasons. Instead we include the effect of HARQ in terms of Chase Combining using a simple HARQ process model from [80], where the soft combined SINR after each transmission is given by

$$\{SINR\}_m = \eta^{m-1} \sum_{l=1}^{m} \left(SINR\right)_l \tag{4.11}$$

where $\{SINR\}_m$ represents the combined SINR after m transmissions, η denotes the chase combining efficiency and $(SINR)_l$ denotes the SINR of the l-th transmission.

The recommended value for η in [80] is 0.95, which is assumed in this study. Note that (4.11) is only valid when m is relatively small, e.g., around 3-4, which is also practical. In our study the HARQ process allows a maximum of three retransmissions per block before it is discarded, i.e., $m = 4$.

4.2.8 L2S Mapping

System-level performance estimation with reasonable accuracy requires an evaluation based on extensive simulations under a variety of scenarios. For complexity reasons, link-level simulation of all links between eNode-B and UEs is not feasible [58], [81] and therefore, separate link-level and system-level simulators are used. As discussed in Section 4.1, the link and system levels are connected through a link-to-system performance mapping function. The idea behind the mapping function is to predict the instantaneous BLER at system-level without performing detailed link-level processing steps. In practice, we use single link-level simulation results in terms of the BLER as a function of SINR to predict each UE's throughput.

The link to system mapping adopted here is based on the *Exponential Effective SINR Mapping* (EESM) [20], [82], [83]. The idea behind this method is to map the current channel conditions (which will involve frequency selective fading for a multi-path channel, for example) to an effective SINR value that may then be used directly with the averaged BLER curves to determine the appropriate block error rate. The effective SINR is defined as

$$SINR^{i}_{eff} = -\beta_i \ln\left[\frac{1}{N_u}\sum_{k}^{Nu}\left(-\frac{SINR_k}{\beta_i}\right)\right] \tag{4.12}$$

where parameter β_i is scaling factor which depends on the used modulation and coding scheme i. N_u is the number or useful sub-carriers within one RB and $SINR_k$ is $SINR$ at k-th sub-carrier

Consequently, EESM is used to calculate the actual throughput from the experienced SINRs per sub-carrier at the receiver. This means that for each transmission, the associated packet error is obtained.

4.2.9 KPI

The Key Performance Indicators (KPI) used in our study to evaluate the system performance are as follows:

- The average cell throughput $(\overline{TP}_{cell})$ is defined here as

$$(\overline{TP}_{cell}) = \frac{\text{total correctly received bits per cell}}{\text{simulation time}} \tag{4.13}$$

- The average user throughput $(\overline{TP}_{i})$ for user i is defined here as

$$(\overline{TP}_{i}) = \frac{\text{correctly received bits for user } i}{\text{session time for user } i} \tag{4.14}$$

- Coverage, denoted by $(TP_{coverage})$, is determined from the CDF curve of the average user throughput taken over all the completed sessions. Coverage is defined as the data rate corresponding to the 5% quantile in the CDF curve, i.e., 95% of the users experience a higher average data rate than the rate specified by the coverage parameter. This KPI indicates the data rate experienced by users around the cell edge. Further, it can be used to differentiate packet schedulers in terms of fairness in the distribution of throughput among users.

- Fairness – measured by Jain's fairness index [84] and given by

$$\Upsilon(TP_i, .., TP_{i_{tot}}) = \frac{\left(\sum_{i}^{i_{tot}} TP_i\right)^2}{i_{tot} \sum_{i}^{i_{tot}} TP_i^2} \tag{4.15}$$

  The result ranges from 0 (worst case) to 1 (best case), and it is maximum when all users achieve the same throughput.

- The outage probability (Po) is defined as the ratio of the number of users not fulfilling their traffic or QoS requirements to the total number of users admitted to the system. In our studies, this is applicable only in QoS-aware streaming studies.

4.2.10 LTE System Parameters

The main simulation parameters and assumptions are summarized in Table 2.

Table 2. DL system level parameters and assumptions

Parameter		**Assumption**
Cellular Layout		Hexagonal grid, 19 cell sites, 3 sectors per site
Inter-site distance		1.732 m
Distance-dependent path loss		$L=128.1 + 37.6\log_{10}(R)$, R in kilometers
Lognormal Shadowing		Similar to UMTS 30.03, B 1.4.1.4
Shadowing standard deviation		8 dB
Correlation distance of Shadowing		50 m
Shadowing correlation	Between cells	0.5
	Between sectors	1.0
Minimum distance between UE and cell		>= 35 meters
Antenna pattern [4] (horizontal) (For 3-sector cell sites with fixed antenna patterns)		$A(\theta) = -\min\left[12\left(\frac{\theta}{\theta_{3dB}}\right)^2, A_m\right]$ θ_{3dB} = 70 degrees, A_m = 20 dB
Carrier Frequency / Bandwidth		2000MHz / 10 MHz
Channel model		Typical Urban
UE speed		3km/h
Total BS TX power (Ptotal)		46dBm - 10MHz carrier
Traffic model		Full Buffer
Fast Fading Model		Jakes Spectrum
Simulation Step		0.5 msec
CQI reporting cycle		1 TTI
MCS feedback delay		2 TTIs
ACK/NACK delay		5 TTIs
Number of SAW channels		6
Maximum number of retransmissions		3
AMC BLER target		20%
Forgetting factor		0.001
Scheduling schemes		Round Robin, PF, MPF, MMPF, QoS-MSPF, MPMPF

Chapter 5

Selected Research Outcomes and Examples

This Chapter presents selected examples of the main research outcomes achieved during the dissertation work, in terms of the system-level performance of the proposed RRM methods, while more detailed and comprehensive examples and related discussions can be found in the original publications [P1]-[P11] that are included in this dissertation. The example outcomes shown here are based on the achieved system-level performance results obtained using the various proposed RRM schemes with emphasis on packet scheduling solutions.

5.1 Fairness-Oriented Scheduling with Single-Stream and Dual-Stream Transmission

The objective of this thesis work was firstly to analyze the OFDMA based cellular network performance and then to present solutions for improving the performance through intelligent RRM components. The obtained results when different MIMO modes and associated UE receivers are considered, in the framework of proposed RRM metrics described in Section 3.6, are primarily presented in publications [P1],[P2] and [P9]. This includes studying the performance of the LTE DL-type system using various receiver structures (single-antenna, dual-antenna MRC and dual-antenna LMMSE), various multiantenna modes (single-stream single user, dual-stream single user, and dual-stream multiuser) and proposed scheduling schemes. In this subsection, we show selected examples of these performance studies assuming 10MHz system bandwidth, Macro case 1 scenario and full buffer best effort traffic type. Detailed simulation parameters are as summarized in Table 2 in Chapter 4.

5.1.1 Single-Antenna Transmission Examples

Examples of the system performance evaluated in terms of average sector throughput and coverage assuming single-stream transmission and single-antenna receiver for different scheduling schemes are presented in Figure 17 and Figure 18 respectively. The power coefficient values used in the proposed PF-Modified scheduler are presented as index M, where M1 represents the first couple of the power coefficients α1 and α2, etc. To focus mostly on the role of the channel quality reporting, α2 is fixed here to 1 and different positive integer values are then demonstrated for α1.The exact values are given in Table 3 [P1]. A clear conclusion is that the coverage (cell edge performance) can be clearly increased using the proposed scheduling approach without essentially compromising the average cell or sector performance. This implies greatly improved fairness in the resource allocation and system performance.

Table 3: Different power coefficient combinations used in the proposed modified PF scheme

Coefficient	Value						
	M1	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**
α1	1	2	4	6	8	10	20
α2	1	1	1	1	1	1	1

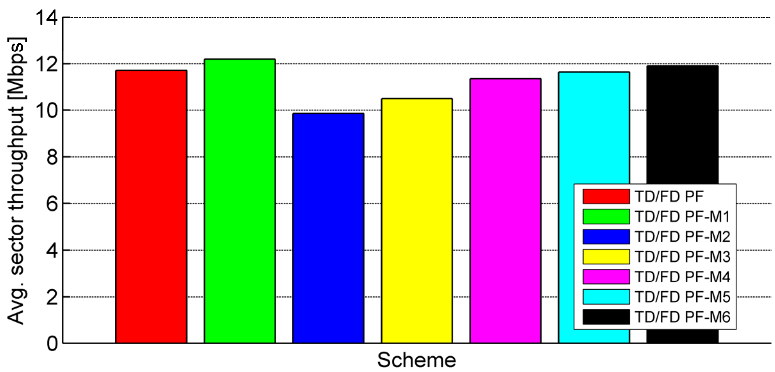


Figure 17: Obtained average cell throughput using the proposed modified PF metrics over the reference PF scheduler for single-stream transmission and single-antenna receiver. M1-M6 correspond to different scheduler parameterizations in the proposed metrics as given in Table 3.

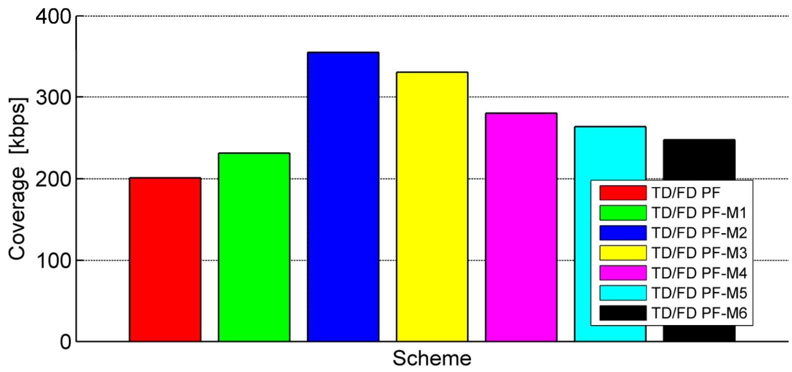


Figure 18: Obtained coverage using the proposed modified PF metrics over the reference PF scheduling scheme for single-stream transmission and single-antenna receiver. M1-M6 correspond to different scheduler parameterizations in the proposed metrics as given in Table 3.

The achievable individual UE data rate performance is illustrated in Figure 19, showing the cumulative distribution function (CDF) of UE throughput, assuming 10 active users in the cell and single-stream transmission. In this case, the UE receivers are dual-antenna MRC receivers [P9]. Clearly, at the 5% outage point, the coverage user throughput is around 300 kbps in the case with PF, while the coverage user throughput can be increased to 450 kbps by using the proposed scheduler indicated by a clear shift to the right. The results again reveal more efficient and fair resource utilization, enabled by the proposed scheduling metrics, together with the direct relation between the throughput and coverage in system performance evaluations. Moreover, such significant coverage improvements (in the order of 40-50%) are achieved at the expense of only very small throughput losses (3-5%) by the introduction of the new flexible (tunable) scheduling metric. Since the slope of the CDF defines the behavior of the algorithms we aim to achieve a steeper slope corresponding to achieving increased fairness. This is clearly visible in Figure 19.

Further examples of the obtained performance results are demonstrated in Figure 20 where the CDFs of scheduled PRBs per UE for the different scheduler scenarios are illustrated, again with dual-antenna MRC receiver. The conclusions made previously are also validated here. The proposed modified PF results in improved internal resource allocation (utilization of PRB's) in all example cases. Considering the 50% probability point for the resource allocation, and taking the case of M1 as an example, we have a gain of about 5%, while in the case of M2 the gain is already raised to 15% compared to ordinary PF. While the number of allocated PRBs is to some extent a system-internal issue, it nevertheless demonstrates better use of radio resources.

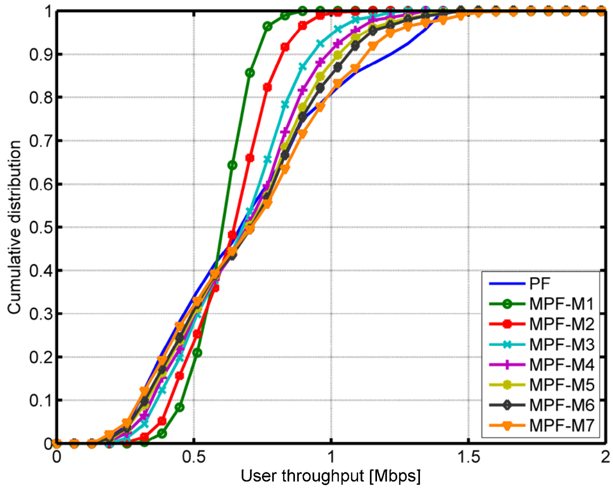


Figure 19: CDFs of user throughputs for different scheduling schemes with single-stream transmission and assuming dual-antenna MRC UE receivers. M1-M7 correspond to different scheduler parameterizations in the proposed metrics as given in Table 3.

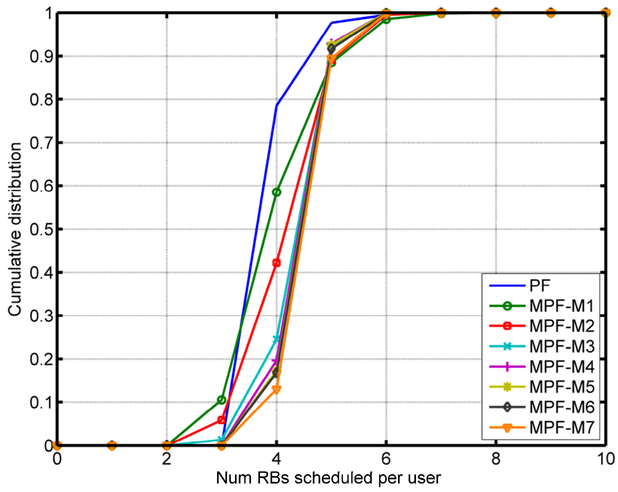


Figure 20: CDFs of scheduled PRBs per user for different schedulers. M1-M7 correspond to different scheduler parameterizations in the proposed metrics as given in Table 3.

Another important performance measure is the achieved user fairness measured with Jain's fairness index defined in (4.15). The Jain's fairness index [84] is generally in the range of 0…1, where the value of 1 corresponds to all users having the same realized throughput (maximum fairness). Fairness distribution for MRC receiver case can be observed in Figure 21 [P2], [P9]. The value on the x-axis corresponds to the used scheduler type, where 1 refers to the reference PF scheduler and 2-8 refer to the proposed modified PF schedulers with the different power coefficients given in Table 3. Increased user fairness is observed for all modified PF scheduler cases and the obtained gain is maximum of 17% compared to the reference PF scheduler.

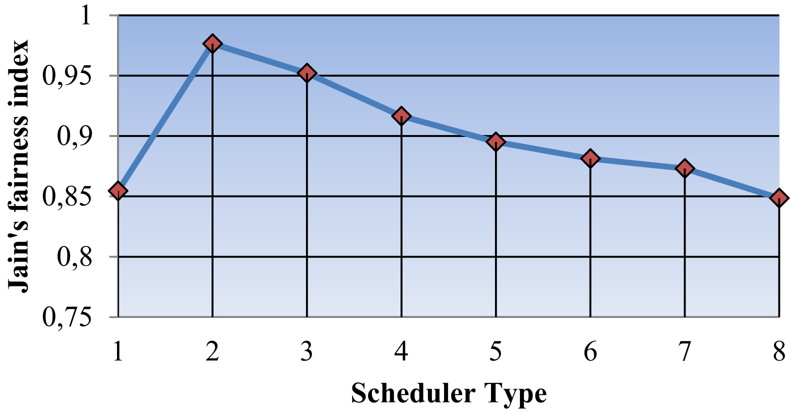


Figure 21: Jain's fairness index for dual-antenna MRC UE receiver case. Scheduler type 1 means ordinary PF, while 2-8 means proposed modified PF with different power coefficients [P9].

5.1.2 Multiantenna Transmission Examples

The next logical step in the early stages of this thesis work was the extension of the RRM principles from single antenna transmission to elementary multiantenna transmission, as described in Section 3.6. Such multiantenna transmission introduces an additional dimension, the space or spatial domain, which is then taken into account as well in the scheduling and resource management, as described in the earlier parts of this thesis. In effect, this results in MIMO channel-aware PS and LA units as part of the RRM process. Here we begin by giving an illustration of the performance of the link adaptation mechanisms(consisting of ILLA and OLLA algorithms, introduced in Chapter 3.3) that are used for removing CQI imperfections and estimating the supported data rates and MCSs. An example of LA functionality is presented in Figure 21 as an illustration of MCS selection probabilities for the different scheduler scenarios. The negligible decrease in higher order modulation usage (less than 3%) leads to an increase in the lower (more robust) ones for improving the cell coverage without essentially compromising the average cell or sector throughput. In all the simulated cases, the MCS distribution behavior shows a relatively similar trend following the choice of the power coefficients in the proposed packet scheduling.

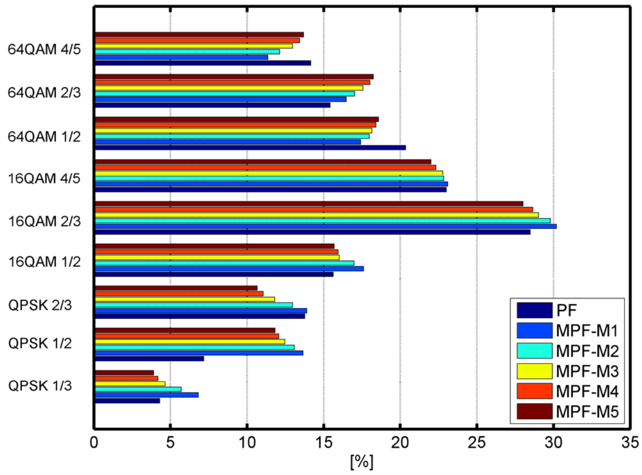


Figure 22: MCS distributions [%] for different scheduling principles in 2x2 MIMO scenario. The values of the tunable parameter for the proposed modified PF scheduler are as given in Table 4.

Next, the proposed MMPF scheduler performance was compared to a traditional scheduling scheme with SDM functionality, in terms of coverage and average throughput. This performance comparison is illustrated here in Table 4 below, while more detailed performance illustrations can be found in [P2].

Table 4: Obtained performance statistics compared to ordinary PF scheduler with different power coefficients (M1-M5) for the proposed multi-stream modified PF scheduler in Macro cell case scenario. Dual-antenna LMMSE UE receivers.

MMPF Scheduling Scenario	Coefficients α1, α2	Coverage Gain [%]	Throughput Loss [%]
M1	1, 1	63	20
M2	2, 1	51	16
M3	4, 1	36	13
M4	6, 1	28	9
M5	8, 1	21	7

It is concluded that the proposed MPF scheduling scheme can efficiently utilize the available system resources and can provide increased user fairness in an UTRAN LTE network also in elementary multinantenna transmission scenarios. Moreover, the LMMSE receiver principle, combined with advanced RRM schemes, can as a whole be considered a key element in improving the system performance in terms of providing higher bit-rates near the cell centre and throughput increase at cell borders by taking the intra- and inter-cell interferences into account. Therefore the combination of advanced receiver structures and suggested RRM algorithms is crucial for increasing system performance in emerging packet radio networks.

The following performance studies focus exclusively on multiantenna transmission, incorporating comparison of micro and macro cell scenarios, reduced feedback schemes, UE velocity, and soft frequency reuse.

5.2 Micro and Macro Cell Scenarios

This subsection illustrates the obtained performance results with advanced packet scheduling in practical example cases of Micro and Macro cell scenarios introduced earlier in Chapter 4.2.1 and [P2], [P10]. In the Macro case each cell site is divided into 3 sectors, inter-site distance is 1732m and the minimum distance between UE and cell is more than 35m. The Micro case consists of single sector, inter-site distance is 500m and the minimum distance between UE and cell is more than 10m.

The performance statistics in terms of throughput and coverage obtained for Micro and Macro cell scenarios are illustrated in Figure 23.The increased system throughput in the Micro case is due to decreased cell sizes and interference from neighboring sites that the UEs are experiencing. The MMPF scheduling scheme achieves coverage gains in the order of 75% at the expense of 23% throughput loss in the Macro case scenario presented in (a) and (b). In the Micro case scenario illustrated in (c) and (d) we obtain as much as 11% more coverage gain for the same throughput loss. The other cases shown with different power coefficient values correspond to different tradeoffs between average throughput and fairness in the scheduling performance. It is seen that the cell throughput loss is decreased stepwise by around 3% per index M, while the coverage gains decrease from 47% to 11%.

Figure 24, in turn, illustrates the Jain's fairness index per scheduling scheme for the Micro and Macro cell scenarios, calculated over all the I_{TOT} =10 UEs. It is clear that user fairness is significantly improved and the corresponding fairness gains, when measured using Jain's index, are in the range of 18%-32% in the Macro case and 25%-34% in the Micro case when compared to ordinary PF scheduling, which is again used as the reference scheme.

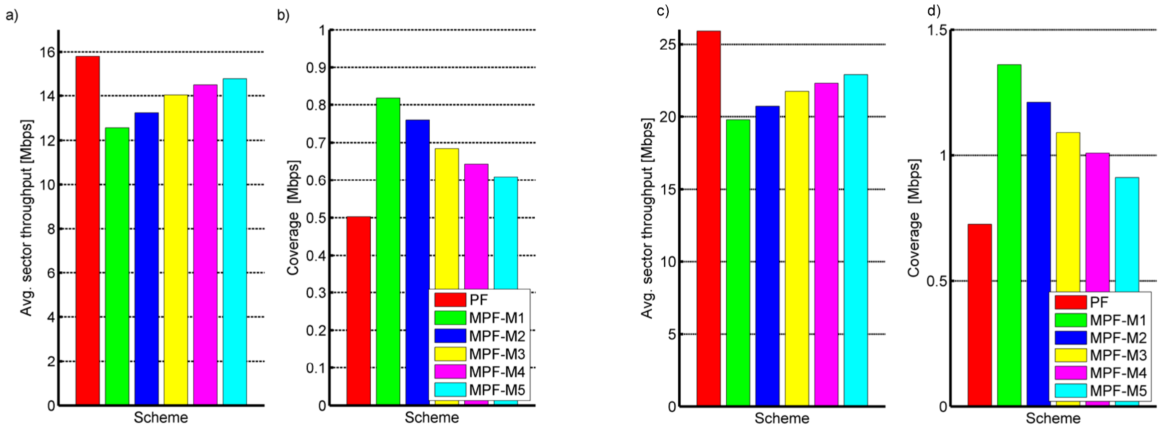


Figure 23: Average cell throughput and coverage gain over the reference PF scheduling scheme for Macro cell (a, b) and Micro cell (c, d). M1-M5 correspond to different scheduler parameterizations in the proposed metrics as given Table 3.

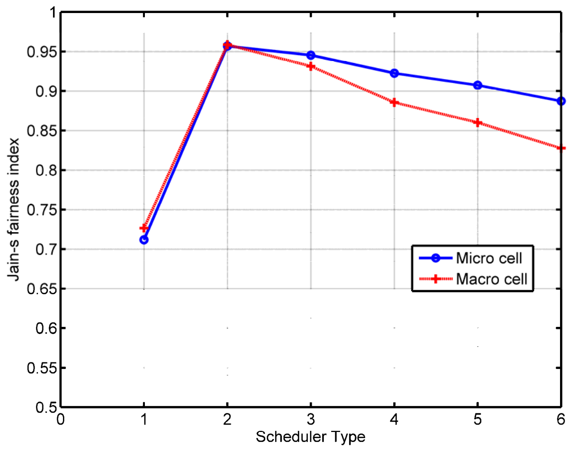


Figure 24: Jain's fairness index per scheduling scheme for Macro and Micro cases. Scheduler type 1 means ordinary PF, while 2-6 means proposed PF with adjustable power coefficients.

The performance numbers obtained for the simulated cases above with different cell scenarios are somewhat increased, compared to the earlier single-stream single-antenna cases, due to MIMO functionalities. The relative performance gains using the MMPF scheduling scheme follow the same trends in achieved throughput, coverage and fairness distribution as previously discussed. Clearly, higher user fairness and coverage are achieved at the expense of a small throughput decrease, as in the earlier single-stream single-antenna cases.

5.3 Impact of Reduced Feedback Schemes

The performance of a multi-antenna packet radio system with advanced RRM schemes and reduced feedback schemes is investigated in articles [P8]-[P10]. Within each time window (TTI) UE sends the formatted and compressed CQI report to eNode-B for each of the resource blocks. Each report is naturally subject to errors due to imperfect decoding of the received signal and delay, as discussed earlier in this thesis. Moreover, CQI needs to be reported with a small finite number of bits for reducing the signaling overhead, especially when multi-stream techniques are applied. The accuracy of the CQIs has a significant impact on the potential gain from using frequency/spatial domain link adaptation and packet scheduling in general.

The reduced feedback schemes are discussed in detail in Section 4.2.5. In this study, we consider different realistic CQI reporting schemes to thoroughly investigate the limits of achieved performance gains from advanced RRM functionality and proposed scheduling metrics. Therefore, the ideal case of full CQI knowledge per RB is compared to more practical schemes considering only partial knowledge of the feedback information. Moreover, MIMO in terms of spatial multiplexing requires additional feedback information for the needs of SU and MU operation as explained earlier. Thus, increased scheduling complexity and signaling overhead require additional constraints, as discussed in [P10].

Based on the 3GPP LTE framework, the simulation studies reported in Figure 25 reveal that a significant reduction in feedback signaling overhead can be achieved with some associated loss (within 15-20%) in cell throughput performance. This loss, however, is only around 7-15% when the proposed advanced scheduling schemes are used, while still providing similar coverage gains as with full CQI reporting. The obtained results are summarized in numerical format in Table 5.

The impact of reduced feedback is also seen as a limiting factor in the functionality of LA by possibly reducing the use of higher order modulations and coding schemes, which corresponds to decreased user throughput experience close to base-station. Consequently, we observe a slight increase of HARQ retransmissions but the 20% BLER target rate is still achieved in all simulated cases [P9], [P10]. Similarly, increased user fairness is achieved, as seen in the previous subsections, and the obtained fairness distributions follow the same trend for all CQI reduced feedback schemes, as illustrated in Figure 26.

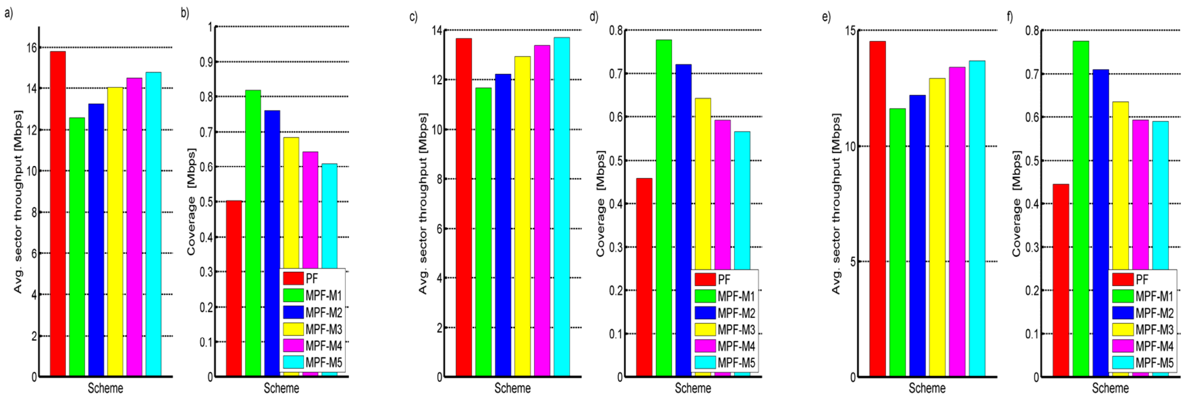


Figure 25: Average sector throughput and coverage for different scheduling schemes for Macro cell scenario with full CQI feedback (a, b), Best -m CQI feedback (c, d) and Threshold based CQI feedback (e, f).

Table 5: Obtained performance statistics compared to ordinary PF scheduler with different CQI reporting schemes and different power coefficients (M1-M5) for the proposed scheduler.

	Coverage Gain [%]			Throughput Loss [%]		
	full	best-m	threshold	full	best-m	threshold
M1	63	69	74	20	15	20
M2	51	57	60	16	11	16
M3	36	40	43	13	6	12
M4	28	29	33	9	2	8
M5	21	24	32	7	0	6

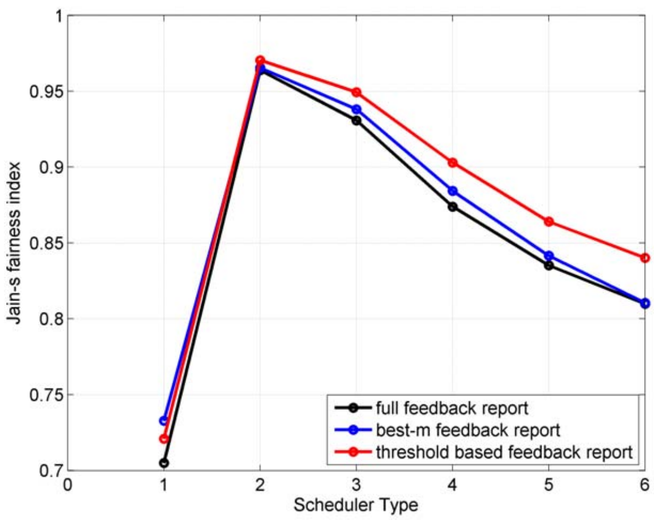


Figure 26: Jain’s fairness index per feedback reporting scheme.

It is concluded that the uplink signaling reduction or compression methods, as in the best-m and threshold-based CQI reporting, are essential in the overall system performance optimization. It is also concluded that due to the interaction between different performance gaining mechanisms in the RRM process, individual optimization may not lead to good overall performance. The understanding of these interactions, through the studies and results reported here, can, however, allow safe reduction of the signaling overhead while keeping most of the gain achievable from intelligent PS schemes.

5.4 Power-Aware Scheduling and Soft Frequency Re-use (SFR) Schemes

Mitigating inter-cell and co-channel interference problems in OFDMA cellular networks by utilization of SFR with different power mask patterns is investigated in articles [P3] ,[P4] and [P11].

Usually, UEs at the cell borders experience heavy co-channel interference (CCI), resulting in low signal to interference and noise ratio (SINR), which stems from frequency reuse in neighboring cells. A straightforward solution for this problem in OFDMA cellular networks is the utilization of soft frequency reuse schemes providing a capacity gain over the other hard frequency reuse schemes [3]. In general, the overall functionality of soft frequency reuse schemes is based on applying specific power masks (a fraction of the maximum transmission power level) over the whole system bandwidth. In soft frequency reuse (SFR), the overall frequency band, shared by all base stations (BS) (reuse factor is equal to 1) is divided into sub-bands with predefined power levels for cell-edge users and other users. Moreover, the power levels in neighboring cells are sub-band specific such that the highest transmission powers are used in non-overlapping sub-bands. Consequently, users located in geographical areas close to and farther away from the BS are allocated different powers, thus limiting the impact of neighboring cells. Furthermore, if the sub-band allocated to the cell-edge UEs is not fully occupied, it can still be used by the other UEs in the cell.

In [P3] the principal packet scheduling metrics are extended by adding transmit power considerations in scheduling decision and evaluated for SIMO and MIMO cases [P11]. This is formulated also in equation (3.14). The idea of incorporating the sub-band power ratio into the scheduling metric is that the sub-band power levels obviously affect the link adaptation and thereon the estimated supportable throughput as well as the actual delivered throughput. Thus by taking the power fluctuations into account, we seek yet higher fairness in the scheduling between truly realized UE throughputs, and thereon better cell-edge coverage. In this study, we assume that the LA unit can feed the PS with power allocation information on a PRB-per-PRB basis. Also, since the power is understood in a stream-wise manner, this gives somewhat higher priority to single-stream transmissions which also helps in increasing the coverage and cell-edge performance.

Figure 27 shows the system performance of combined soft frequency reuse schemes with advanced power-aware packet scheduling algorithms in terms of average cell throughput and coverage. The obtained result demonstrate the behavior of the proposed MPMPF scheduler by using different power coefficients and by comparing it against other PF scheduling algorithms with different SFR power masks. The power mask configurations used are PM1 (0dB, 4dB, 4dB) and PM2 (0dB, 1dB, 4dB) illustrated in Figure 28. The values in the brackets represent nominal transmission power values in dB.

Utilization of such technique offers significant coverage improvements in the order of 30% obtained at the expense of only 14% throughput loss.

Furthermore, the user fairness illustrated in Figure 29 is also greatly increased, when measured using Jain's fairness index. Clearly, the fairness distribution with MPMPF+SFR outperforms the used reference PF scheduler and PF+SFR for both analyzed cases. The received fairness gains are in range of 17%-31%. Compared to the PF+SFR case the corresponding gains are in the range of 15-17% for both cases.

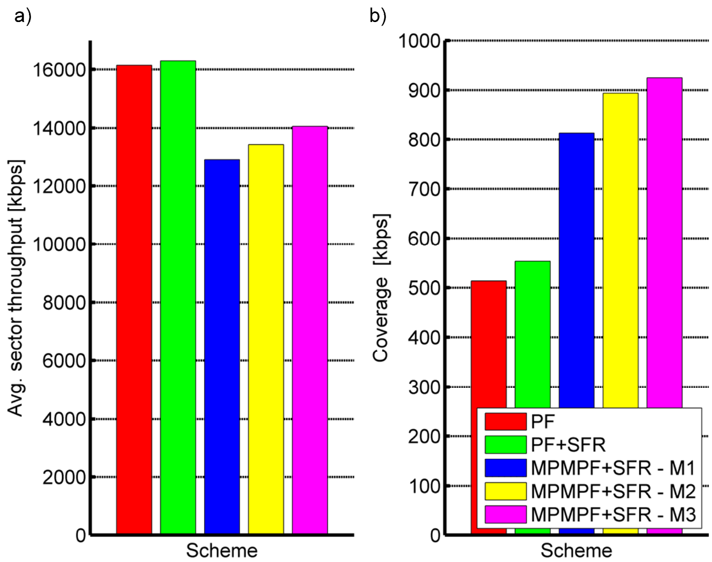


Figure 27: Average cell throughput (a) and coverage gain (b) over the reference PF scheduling scheme and PF +SFR scheme, with difference adjustable coefficients cases M1-M3 in the proposed scheduling metrics.

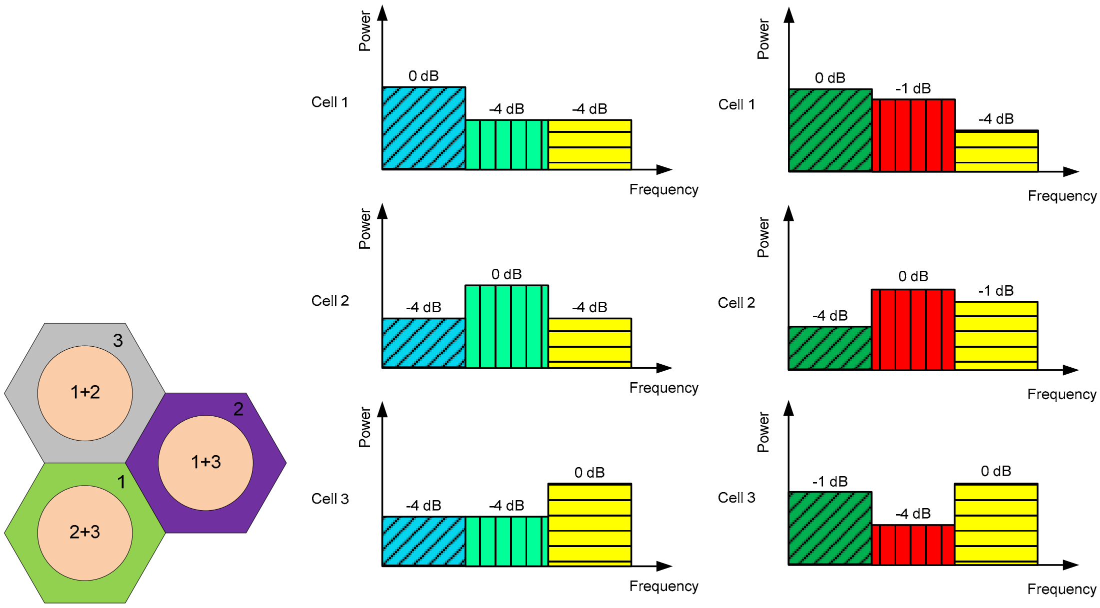


Figure 28: Soft Frequency reuse scheme with 3 sub-bands division.

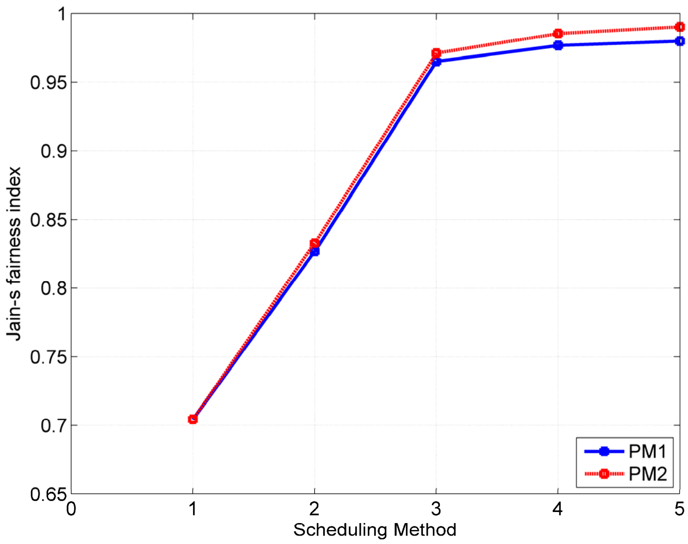


Figure 29: Jain's fairness index per scheduling scheme and different power masks. Scheduler type 1 means ordinary PF, while 2 means PF +SFR, 3-5 means MPMPF +SFR with power coefficients M1,M2 and M3 correspondingly.

5.5 UE Velocity Impact

Majority of our system simulation and performance studies have focused on low-mobility scenarios, mainly 3km/h. In article [P6], in turn, performance evaluations are carried out with different UE velocities. The obtained results indicate that increasing the velocity from 3 km/h to 30km/h and 120 km/h has a major impact on system throughput in general. The gains from link adaptation and advanced receiver structures, compared to low-mobility cases, are heavily reducing in higher velocity scenarios due to outdated channel state information and other feedback reports. In [P6], advanced PF scheduling was compared to basic round robin (RR) and basic PF in such high-velocity cases and the obtained performance gains are again analyzed.

The obtained performance numbers are summarized in Table 6, showing the results with both traditional PF and proposed MMFP, and comparing to classical RR scheduler. The advanced MMPF scheduling scheme achieves around 5-10% gain in system throughput and 25-35% gain in the coverage when compared to classical RR in high-velocity case of 120 km/h. Also compared to traditional PF, we are still able to demonstrate some coverage gain. The corresponding Jain's fairness indexes are again reported in Figure 30 where scheduler type 1 refers to RR, type 2 to PF, and 3 and 4 to MMPF. The outdated user feedback used in throughput estimation from PF scheduling algorithm leads to frequency domain scheduling gain loss. Therefore, the PF scheduler show slightly worse fairness compared to RR scheduler.

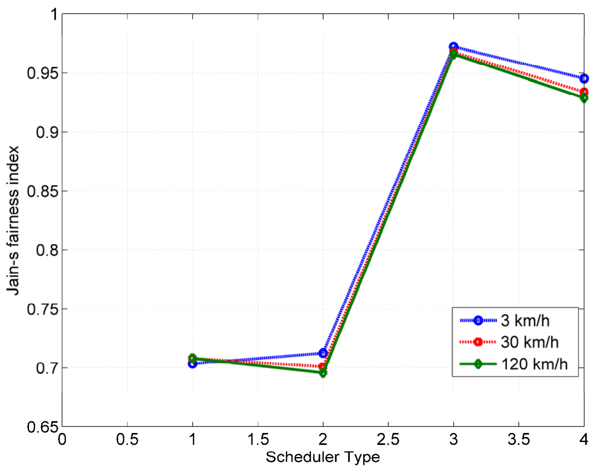


Figure 30: Jain's fairness index per scheduling scheme.

Table 6: Obtained performance statistics compared to RR scheduler for different UE velocities.

Coverage Gain [%] / Throughput Gain[%]	Performance gains		
	3 km/h	*30 km/h*	*120 km/h*
PF	30 / 38	6 / 18	1 / 11
MMPF-M1	86 / 17	46 / 2	34 / 4
MMPF-M2	73 / 22	39 / 22	24 / 8

5.6 QoS-Aware Scheduling and Video Streaming

Here we summarize the results from QoS provisioning framework based on proposed QoS-aware PS schemes and metrics, which are described earlier in Chapter 3.7 and in original publications [P7]-[P8]. The main target is to estimate the number of UEs or streaming services operating at a given fixed bit rate that can be supported by the system based on deployed scenarios in a Micro cell environment. The obtained results are based on 3GPP evaluation metrology, where a user is considered to be in outage if more than 2% of the video packets for the user are erroneous or discarded due to exceeding delay limit when monitored over the whole video session duration. Similarly, video capacity is defined as the number of supported users per cell without exceeding the system saturation point. Here, the system saturation point is set at 5% of the cell outage level, i.e., to the point where 95% of the users in the cell are satisfied (having a maximum of 2% packet loss rate, as described above). Thus in this sense, the performance evaluation framework is quite different compared to earlier best-effort full-buffer scenarios.

The performance studies and corresponding results are based on two different video streaming services with 128 kbps (Scenario 1) and 256 kbps (Scenario 2) constant bit rates (CBR) source video data. The corresponding mean packet size with truncated Pareto distribution and mean inter-arrival packet time are 100 bytes - 6ms for Scenario 1, and 200 bytes - 4ms for Scenario 2 [P7]. Packet delay budget, as well as discard timer threshold is set to 20 ms, following from [40]-[42].

The simulation results in Figure 31 indicate that employing the proposed QoS-aware PS algorithm provides video capacity gain in the order of 13% and an additional 8% throughput increase for Scenario 1 when compared to reference PF scheme achieving 94 UEs support. Similarly for Scenario 2 illustrated in Figure 32, where the obtained gains are in the same range, but the absolute number of supported users or streams is naturally reduced due to increased video packet size and source bit rate. At the air interface level, this is reflected as fulfilling the QoS requirements with more PRBs scheduled per user in general. The main observation here is that incorporating the QoS requirements into the scheduling metrics, as in (3.18), yields considerable gain in the system performance.

Further illustrations on the obtainable system performance are illustrated in Figure 33 with CDF of the number of scheduled users per TTI. As the figure clearly shows, the proposed scheduling scheme allows more users to be scheduled in each TTI and therefore the achieved capacity gains are due to the increased and more efficient usage of PRB resources.

The analysis is also validated in Figure 34 showing the CDF of video packet delay for different schedulers. The CQI reporting process has a direct impact on the delay performance and therefore a combined approach of HOL and CQI criteria in priority metric calculation is beneficial in such scenarios. The delay performance is strictly within the bounds of 20 ms for the simulated video traffic scenarios. More detailed results can be found in [P8] in which the RRM framework and performance analysis also entail reduced feedback schemes.

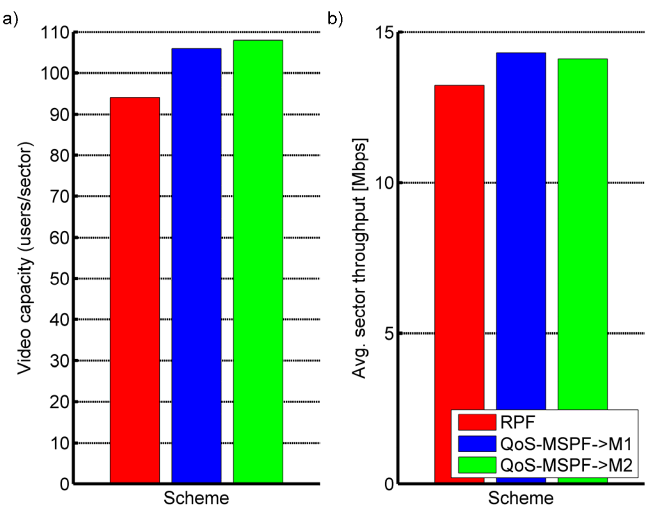


Figure 31: Video capacity (a) and average sector throughput (b) for different scheduling schemes in Scenario 1.

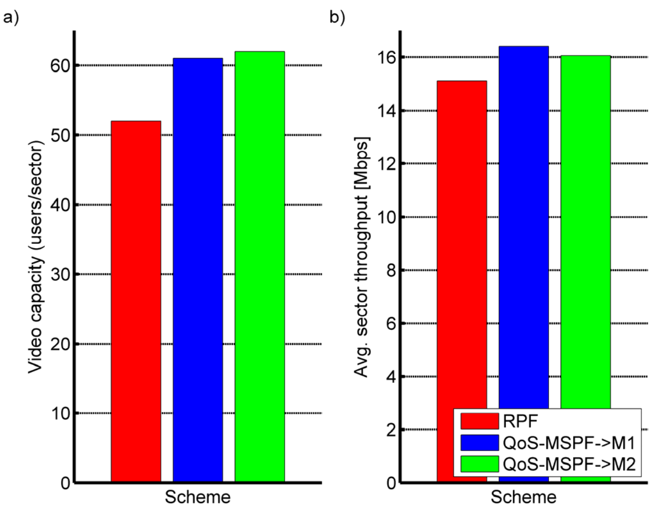


Figure 32: Video capacity (a) and average sector throughput (b) for different scheduling schemes in Scenario 2.

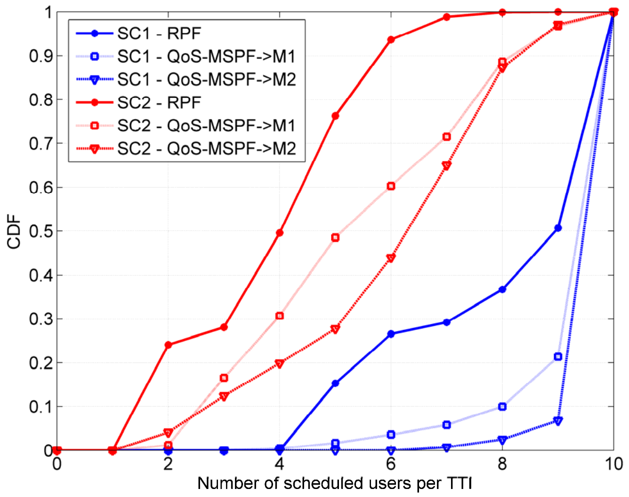


Figure 33: CDF of scheduled users per TTI for simulated scenarios.

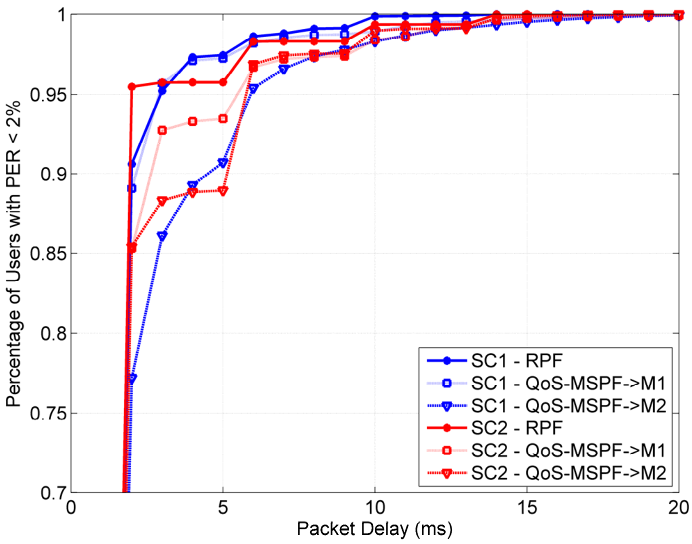


Figure 34: CDF of packet delay for simulated scenarios.

Conclusions

The constantly increasing data content and number of wireless applications for mobile users is one of the main reasons for mobile cellular radio network evolution. The number of people using mobile devices for content production and consumption, in addition to classical voice calls and text messaging, has grown tremendously. The need for content sharing, video calls and Internet connectivity requires improved levels of performance from mobile networks and their operators.

At the radio interface level, most emerging wireless connectivity solutions build on OFDMA-type solutions. Another central ingredient is use of multiantenna or MIMO techniques. A good example of commercial deployment of these techniques is 3GPP mobile cellular radio evolution and LTE/LTE-Advanced. LTE technology, including LTE-Advanced, is currently seen as the primary 4G technology globally.

In addition to radio interface and physical layer advancements, the deployment of advanced radio resource management techniques is one of the most effective performance gaining mechanisms in cellular radio evolution. This dissertation work focused on the application of advanced RRM methods, such as novel packet scheduling, link adaptation, admission control and retransmission mechanisms, as an efficient way of enhancing the system-level performance of future multiuser packet radio systems utilizing OFDMA and multiantenna techniques at the physical layer. The specific emphasis in this dissertation work was on increased user fairness through appropriate scheduling metrics, implying improved cell-edge performance and coverage which is critical in the overall user experience in highly-loaded mobile networks. Several new practically-oriented scheduling metrics were proposed and evaluated in an LTE-type OFDMA mobile network context using comprehensive quasi-static cellular network system simulations. The impacts of practical channel quality feedback mechanisms, soft frequency reuse schemes, user velocities and micro and macro cell scenarios were also studied. Overall, the results demonstrate that highly improved fairness and thus improved cell-edge performance and coverage can be obtained using the proposed scheduling metrics with built-in fairness control.

While the main emphasis in this work has been in best-effort full-buffer data scenarios, another main line of study in the dissertation work has been constant bit rate or video streaming services, and associated advanced scheduling methods. Novel QoS-aware scheduling metrics were proposed with built-in bit rate and delay requirements for streaming services. Clear performance improvements were demonstrated over existing reference schedulers, resulting in a higher number of simultaneously supported video streams in the cell.

Future work should focus on extending the performance studies and RRM mechanisms to more advanced MIMO schemes, including closed-loop precoded single-user and multiuser MIMO and collaborative multicell techniques.

Bibliography

[1] ”3GPP,” http://www.3gpp.org

[2] H. Holma and A. Toskala, WCDMA for UMTS: HSPA Evolution and LTE, John Wiley & Sons Ltd., Fourth Edition, 2007, 539 pages.

[3] 3GPP Technical Report TR 25.814, ver. 7.1.0 “Physical Layer Aspects for Evolved UTRA”, Oct. 2006.

[4] 3GPP Technical Specification 25.211, ver. 5.8.0, “Physical channels and mapping of transport channels onto physical channels (FDD)”(release 5), December 2005

[5] H. Holma and A. Toskala, HSDPA/HSUPA for UMTS- High Speed Radio Access for Mobile Communications, John Wiley & Sons Ltd,2006

[6] E. Dahlman et al., 3G Evolution: HSPA and LTE for Mobile Broadband, Academic Press, Second edition, 2008.

[7] 3GPP TSG-RAN, RP040461, Proposed Study Item on Evolved UTRA and UTRAN, December 2004.

[8] ITU global standard for international mobile telecommunications “IMT-Advanced”, IMT-ADV/1-E,Mar. 2008.

[9] 3GPP Technical Report TR 36.912, Feasibility study for Further Advancements to E-UTRA (LTE-Advanced)

[10] ITU-R, Report m.2134: Requirements related to technical performance for IMT-Advanced radio interface(s), 2008.

[11] N. D. Tripathi et al., Radio Resource management in Cellular Systems, Springer, 2001.

[12] T. Kashima, M. Rinne, I. Kovacs, M. Kuusela, P. Michaelsen, K. Pedersen, C. Rosa, and E. Tuomaala, “Performance evaluation of Evolved-UTRA data rates and throughput," in Technical Report of Japan Society of Simulation Technology, JSST- MM2007-21, Tokio, Japan, Oct. 2007, pp. 80{85.

[13] J. Puttonen, N. Kolehmainen, T. Henttonen, M. Moisio, and M. Rinne, ”Mixed traffic packet scheduling in UTRAN long term evolution downlink," in Proc. IEEE Nineteenth International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2008, Cannes, France, Sep. 2008

[14] 3GPP Technical Specification Group Radio Access Network Technical Specification 36.300 v8.0.0: Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (EUTRAN); Overall description, march 2007

[15] G. J. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. Bell Labs Technical Journal, 1996.

[16] I. E. Telatar. Capacity of multi-antenna Gaussian channels. Technical Memorandum BL011217-950615-07TM, AT&T Bell Labs, June 1995.

[17] T. Buot and M. Rinne, "Performance of WCDMA with radio resource management," in Twelfth Annual International Conference on Wireless Communications, Wireless 2000, Calgary, Alberta, Canada, Jul. 2000, pp. 203{209.

[18] R. Kwan and M. Rinne, \Performance analysis of the downlink shared channel in a WCDMA network," in Proc. IEEE International Conference on Telecommunications, ICT 2001, Bucharest, Romania, Jun. 2001, vol. 3, pp. 379-384.

[19] H. Bölcskei, D. Gesbert, and A. J. Paulraj. On the capacity of ofdm-based spatial multiplexing systems. IEEE Transactions on Communications, 50(2):225–, Feb 2002.

[20] Ericsson. System-level evaluation of OFDM - further considerations. Technical Report R1-03-1303, Technical Specification Group Radio Access Network, 2004.

[21] Ericsson. E-UTRA Performance Checkpoint: Downlink. TSGR1 AH R1-071956, 3GPP, Apr. 2007.

[22] Huawei. LTE Downlink Performance Evaluation Results. TSGR1 AH R1-071961, 3GPP, Apr. 2007.

[23] P. E.Mogensen, N.Wei, A. Pokhariyal, I. Kovacs, F. Frederiksen, K. Pedersen, K. Hugl,T. E. Kolding, and M. Kuusela. LTE capacity versus Shannon. In IEEE 65th Vehicular Technology Conference, Apr. 2007.

[24] Motorola. LTE Downlink System Performance Verification Results. TSGR1 AH R1- 071976, 3GPP, Apr. 2007.

[25] NEC. LTE Downlink Performance. TSGR1 AH R1-071978, 3GPP, Apr. 2007.

[26] Nokia-Siemens-Networks. LTE Performance Benchmarking. TSGR1 AH R1-071960, 3GPP, Apr. 2007.

[27] T. E. Kolding, "Link and system performance aspects of proportional fair scheduling in WCDMA/HSDPA", in Proc. IEEE Vehicular Technology Conference (VTC-2003), Orlando, FL, USA, Oct. 2003, pp. 1717-1723.

[28] A. Pokhariyal, T. E. Kolding, and P. E. Mogensen, "Performance of downlink frequency domain packet scheduling for the UTRAN long term evolution," in Proc. IEEE Personal, Indoor and Mobile Radio Communications Conference (PIMRC-2006), Helsinki, Finland, Sept. 2006.

[29] A. Pokhariyal, K. I. Pedersen, G. Monghal, I. Z. Kovacs, C. Rosa, T. E.Kolding, and P. E. Mogensen, “HARQ aware frequency domain packet scheduler with different degrees of fairness for the UTRAN long term evolution,” in Proc. IEEE Vehicular Technology Conference (VTC-2007), Dublin, Ireland, April 2007, pp. 2761-2765.

[30] N. Wei, et al., “Performance of MIMO with frequency domain packet scheduling”, in Proc. IEEE Vehicular Technology Conf. (VTC-2007 Spring), Dublin, Ireland, May 2007.

[31] Motorola. LTE Downlink System Performance Verification Results. TSGR1 AH R1-071976, 3GPP, Apr. 2007.

[32] Samsung. E-UTRA Performance Verification: DL Throughput. TSGR1AH R1-071969, 3GPP, Apr. 2007.

[33] 3GPP Technical Specification 36.321, Release 8, version 8.3.0, Evolved Universal Terrestrial Radio Access (E-UTRA) Medium Access Control (MAC) protocol specification, September 2008.

[34] P. Svedman, S.K. Wilson, L.J. Cimini, B. Ottersten, "A Simplified Opportunistic Feedback and Scheduling Scheme for OFDMA", in IEEE Proc. VTC. pp. 1878-1882, May 2004.

[35] S. Sanayei, A. Nosratinia, N. Aldhahir, "Opportunistic Dynamic SubChannel Allocation in Multiuser OFDM Networks with Limited Feedback", in IEEE Proc lnform. Theory Workshop, pp. 182-186, October2004.

[36] T.E. Kolding, F. Frederiksen, A. Pokhariyal, "Low-Bandwidth Channel Quality lndication for OFDMA Frequency Domain Packet Scheduling", Proc. ISWCS'06, Spain, September 2006.

[37] Y. Sun, et.al., "Multi-user Scheduling for OFDMA Downlink with Limited Feedback for Evolved UTRA", in IEEE Proc. VTC-06-fall, September 2006.

[38] S. Yoon, C. Suh, Y. Cho, D. Park, "Orthogonal Frequency Division Multiple Access with an Aggregated Sub-Channel Structure and Statistical Channel Quality Measurements", in IEEE Proc. VTC-2004, September 2004.

[39] Nokia. Reduced CQI design for DL SU-MIMO. Technical Report R1-07-0388, Technical Specification Group Radio Access Network, 2007.

[40] K.I. Pedersen, G. Monghal, I.Z. Kovacs, T.E. Kolding, A. Pokhariyal, F. Frederiksen, P. Mogensen, “Frequency domain scheduling for OFDMA with limited and noisy channel feedback,” in Proc. IEEE Vehicular Technology Conf. (VTC-2007 Fall), Baltimore, MD, Sept. 2007.

[41] T.-D. Nguyen, Y. Han, A Proportional Fairness Algorithm with QoS Provision in Downlink OFDMA Systems, IEEE Communication Letters, vol. 10, No. 11, November 2006.

[42] T.E. Kolding, QoS-Aware Proportional Fair Packet Scheduling with Required Activity Detection, VTC2006-Fall, September 2006.

[43] G. Monghal, et al., “QoS Oriented Time and Frequency Domain Packet Schedulers for The UTRAN Long Term Evolution”, in Proc. VTC2008-Spring, vol.3, pp. 2532-2536, May 2008.

[44] Wei-Kuan Shih; J. Liu, "Algorithms for scheduling imprecise computations with timing constraints to minimize maximum error," , IEEE Transactions on Computers, vol.44, no.3, pp.466-471, Mar 1995.

[45] M. Rahnema, "Smart trunk scheduling strategies for future integrated services packet-switched networks," IEEE Communications Magazine, vol.26, no.2, pp.43-50, Feb. 1988.

[46] J. Holtzman, "Asymptotic analysis of proportional fair algorithm," 12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2001, vol.2, no., pp.F-33-F-37 vol.2, Sep/Oct 2001.

[47] “Universal Mobile Telecommunications System (UMTS); requirements for the UMTS terrestrial radio access system (UTRA),” ETSI TR 101 111 v 3.0.1 (1997-11), 1997.

[48] R. Castello, et al., “ IMT – 2000 standards: radio aspects, “ Proceedings of IEEE Personal Communications, vol. 4, issue 4, pp. 30-40.August 1997.

[49] 3GPP Technical Specification 21.101, version 5.14.0, “Technical Specifications and Technical reports for UTRAN-based 3GPP system (Release 5), “ December 2009.

[50] 3GPP Technical Specification 21.101, version 6.10.0, “Technical Specifications and Technical reports for UTRAN-based 3GPP system (Release 6), “ December 2009.

[51] 3GPP Technical Report 25.814, Release 7, version 7.1.0, “Physical layer aspects for evolved universal terrestrial radio access (UTRA), “ September 2006.

[52] 3GPP Technical Specification 36.300, release 8, version 8.5.0, Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description ; Stage 2,” May 2008.

[53] 3GPP Technical Specification 36.912, Version 9.0.0, “ Feasibility Study for Further Advancements for E-UTRA (LTE-Advanced), “ September 2009.

[54] R. Nee and R. Prasad, OFDM for wireless Multimedia Communications, Artech house, 2000.

[55] I. Koffman and V. Roman “ Broadband wireless access solutions based on OFDM access in IEEE 802.16, ” IEEE Communications Magazine, vol. 40, issue 4, pp. 96-103, April 2002.

[56] Y. Yaghoobi, “ Scalable OFDMA physical layer in IEEE 802.16 wireless man, “ Intel Technology Journal, vol. 8, issue 3, pp. 201-212, August 2004.

[57] 3GPP Technical Report 25.913, version 7.3.0, Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN), March 2006.

[58] D. Astely, et al.," A future radio-access framework," IEEE Journal on Selected Areas in Communications, vol 24, issue 3, pp. 693-706, March 2006.

[59] S. M. Alamouti, "A simple transmit diversity technique for wireless communications ," IEEE Journal on Selected Areas in Communications, vol.16, no.8, pp.1451-1458, Oct 1998.

[60] E. L. Hahne and R. Gallager, "Round robin scheduling for fair flow control in data communication networks," in Proceedings of IEEE International Conference in Communications, June 1986, pp. 103-107.

[61] D. Stiliadis, A. Varma, "Efficient fair queuing algorithms for packet-switched networks," IEEE/ACM Transactions on Networking, , vol.6, no.2, pp.175-185, Apr 1998.

[62] A. Gyasi-Agyei, S.L. Kim, "Cross-Layer Multiservice Opportunistic Scheduling for Wireless Networks", IEEE Communications Magazine, June 2006.

[63] P. Svedman, D. Hammarwall, B. Ottersten, "Sub-carrier SNR estimation at the transmitter for reduced feedback OFDMA," in Proceedings of European Signal Processing Conference, Florence, Italy, September 2006.

[64] A. Jalali, R. Podovani, and R. Pankaj, " Data Throughput of CDMA-HDR High Efficiency-High Data Rate Personal Communication Wireless System, " Proceedings of the IEEE Vehicular Technology Conference (VTC), vol. 3, pp. 1854-1858, Tokyo, Japan, May 2000.

[65] J. M. Holtzman. " Asymptotic Analysis of proportional Fair Algorithm," in IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), vol. 2, pp. F33-F37, San Diego, California, USA, September 2001.

[66] R. C. Elliot and W. A. Krzymien, " Scheduling Algorithms for the cdma2000 Packet Data Evolution, " Proceedings of the IEEE Vehicular Technology Conference (VTC), vol. 1, pp. 304-310, Vancouver, Canada, September 2002.

[67] F. D. Calabrese, P. H. Michaelsen, et al., " Search-Tree based Uplink Channel Aware Packet Scheduling for UTRAN LTE, " in Proceedings of the 67th IEEE Vehicular Conference, pp. 1949-1953, Singapore, May 2008.

[68] F. D. Calabrese, C. Rosa, et al., " Adaptive Transmission Bandwidth Based Packet Scheduling for LTE Uplink, " in Proceedings of the 68th IEEE Vehicular Conference, pp 1-5, Calgary, Canada, September 2008.

[69] A. J. Goldsmith and S. –G. Chua, " Adaptive Coded Modulation for Fading Channels, " IEEE Transactions on Communications, vol. 46, no. 5, pp. 595-602, May 1998.

[70] K. B. Letaief and Y. J. Zhang, " Dynamic Multiuser Resource Allocation and Adaptation for Wireless Systems, " IEEE Wireless Communications, vol. 13, no. 4, pp. 38-47, August 2006.

[71] M. Nakamura, Y. Awad and S. Vadgama, “Adaptive control of link adaptation for high speed downlink packet access (HSDPA) in W-CDMA,” in Proceedings of The 5th Wireless Personal Multimedia Communications Conference, vol. 2, pp. 382-386, October 2002.

[72] R. Cam and C. Leung, “ Multiplexed ARQ for time-varying channels,” IEEE Transactions on Communications, vol. 46, issue 1, pp. 41-51, January 1998.

[73] S. Hara, A. Ogino, M. Araki, M. Okada and N. Morinaga, ” Throughput performance of SAW-ARQ protocol with adaptive packet length in mobile packet data transmission,“ IEEE Transactions on Vehicular Technology, vol. 45, issue 3, pp. 561-569, August 1996.

[74] S. Falahti and A. Swenson, “ Hybrid type-II ARQ schemes with adaptive modulation systems for wireless channels, “ in Proceedings of IEEE Fiftieth Vehicular Technology Conference VTC 1999-Fall, vol. 5, pp. 2691-2695, Amsterdam, Holland, September 1999.

[75] E. Malkamäki, D. Mathew and S. Hämälainen, “ Performance of hybrid ARQ techniques for WCDMA high data rates,” in Proceeding of IEEE Fifty-Third Vehicular Technology Conference VTC 2001 Spring, vol. 4, pp. 2720-2724, Rhodes, Greece, May 2001.

[76] A. Das, F. Khan, A. Sampath and S. Hsuan-Jung, “ Performance of hybrid ARQ for high speed downlink packet access in UMTS,” in Proceeding of IEEE Eighteenth Vehicular Technology Conference VTC 2001 Fall, vol. 4, pp. 2133-2237,Atlantic City, USA, October 2001.

[77] M. Wrulich, W. Weiler and M. Rupp, “ HSDPA performance in a mixed traffic network,” in Proceeding of IEEE Vehicular Technology Conference VTC 2008 Spring, pp. 2056-2060, May 2008.

[78] I. Z. Kovacs, K. I. Pedersen, J. Wigard, F. Frediriksen, and T. E. Kolding, “ Hsdpa performance in mixed outdoor-indoor micro cell scenarios, “in Proceeding of IEEE Personal Indoor and Mobile Radio Communications (PIMRC) , September 2006.

[79] A. Sampath, P. S. Kumar, J. M. Holtzman, “ On Setting Reverse Link Target SIR in a CDMA System, “ in Proceeding of IEEE Vehicular Technology Conference VTC 1997, pp. 929-933, May 1997.

[80] F. Frediriksen and T. E. Kolding, “ Performance and modeling of WCDMA/HSDPA transmission/H-ARQ schemes, “in Proceeding of IEEE Fifty-Six Vehicular Technology Conference VTC 2002 Fall, vol. 1, pp. 472-476, Vancouver (BC), Canada, September 2002.

[81] S. Hämällinen, H. Holma and Kari Sipilä, “ Advanced WCDMA network simulator,” in Proceedings of IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC ’99), pp. 951-955, Osaka, Japan, September 1999.

[82] Ericsson. Effective SNR mapping for modeling frame error rates in multiple-state channels. Technical report C30-20030429-010, 3GPP2, 2003.

[83] K. Brueninghaus, D. Astely, T. Salzer, et. al., "Link performance models for system level simulations on broadband radio access systems, " in Proceedings of the 16th IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), vol. 4, pp. 2306-2311, Berlin, Germany, September 2005.

[84] D. Chui and R. Jain,"Analysis of the increase and decrease algorithms for congestion avoidance in computer networks," Computer Networks and ISDN Systems, 1989.

Publications

Publication 1

S. Nonchev, J. Venäläinen and M. Valkama, “New frequency domain packet scheduling schemes for UTRAN LTE downlink,” in *Proceedings of ICT-MobileSummit 2008 Conference* (ICT-2008), Stockholm, Sweden, June 2008, 8 pages.

ICT-MobileSummit 2008 Conference Proceedings
Paul Cunningham and Miriam Cunningham (Eds)
IIMC International Information Management Corporation, 2008
ISBN: 978-1-905824-08-3

New Frequency Domain Packet Scheduling Schemes for UTRAN LTE Downlink

Stanislav NONCHEV, Juha VENÄLÄINEN, Mikko VALKAMA
Tampere University of Technology, Department of Communications Engineering, P.O.Box 553, Tampere, FIN-33101, Finland
Email: stanislav.nonchev@tut.fi, juha.venalainen@tut.fi, mikko.e.valkama@tut.fi

Abstract: In this paper, we study the performance of different packet scheduling schemes in 3GPP UTRAN long term evolution (LTE) downlink context. The downlink is based on orthogonal frequency division multiple access (OFDMA) air interface and relocates the packet scheduler from RNC to the base station. This gives the possibility to implement more intelligent scheduling schemes that are aware of the instantaneous state of the radio channel utilizing both time and frequency domain resources in an efficient manner. Stemming from the earlier Proportional Fair (PF) scheduler studies, we continue to investigate its efficiency in increasing the system capacity in terms of throughput and coverage by simulating practical OFDMA system environment with frequency reuse of 1. We show that by using new modified PF schemes, significant coverage improvements in the order of 70% at the expense of only 10% throughput loss can be obtained. Furthermore, we investigate the behaviour of the PF scheduling using single antenna as well as multi-antenna diversity receivers.

Keywords: radio resource management, packet scheduling, proportional-fair scheduling

1. Introduction

Development of new radio interface technologies for beyond 3G cellular systems (long term evolution of 3G, LTE) with support to high-data-rates, low-latency and packet-optimised radio-access, has led to the use of OFDM/OFDMA [1]. Performance improvements are basically obtained through proper deployment of fast link adaptation and new packet scheduling algorithms, together with exploiting the available multi-user diversity in both time and frequency domains. The common feature for advanced schedulers is their ability to adapt to the fast channel quality variations. Such fast adaptation rate results in a throughput gain compared to the scheduling algorithms that do not take advantage of channel variations. In this paper, we introduce a new scheduling scheme based on proportional fair (PF) principle and investigate its performance using different receiver schemes in LTE context.

The research of different packet scheduling algorithms has been very active based on existing publications on different approaches, and their comparisons in different simulator environments [2]-[5]. Using [2]-[4] as starting points for LTE, it has been reported that frequency domain packet scheduling (FDPS) algorithms are always a compromise between the overall cell throughput and resource fairness among users. We focus on the Proportional Fair (PF) algorithm, which in general offers an attractive balance between cell throughput and user fairness. One of the drawbacks of FDPS, in turn, is relatively higher scheduling complexity and increased signalling overhead. Here, we extend our studies on algorithm development by deploying both time domain (TD) and frequency domain (FD) PF scheduling schemes that can efficiently utilise the provided feedback information from all the user equipments (UEs), typically reported in terms of channel quality indicators (CQI).

A new scheduler is proposed, called modified PF, which is shown to improve the system performance considerably, depending somewhat on the receiver scheme used. The performance evaluations of the scheduling methods presented in this paper are based on the system model according to the 3GPP evaluation criteria [1] and measured in terms of average cell throughput and coverage.

The rest of the paper is organised as follows: Section 2 describes the proposed modified PF scheduling scheme. Section 3 presents the overall system model and simulation assumptions. The simulation results and analysis are presented in Section 4, while the conclusions are drawn in Section 5.

2. Scheduling Process

In general, the task of a Packet Scheduler (PS) is to select the most suitable user to access the channel in order to optimize throughput, fairness, and delay performances. Fast packet scheduler is the mechanism determining which user to transmit to in a given transmission time interval (TTI).

To efficiently utilize the limited radio resources, the scheduler should consider the state of the channel when selecting the user to be scheduled based on ACK/NACK signalling information and CQI's [6]. This information is also used by link adaptation mechanism to ensure that the supported throughput estimation for UEs is above certain block error rate (BLER) target for the first transmission. HARQ management is then typically used to provide the necessary buffer information and transmission format of pending retransmissions [3]. Maximizing the system throughput can be regarded as the most important requirement of a packet scheduling algorithm [4]. In LTE, the system bandwidth of 10 MHz is divided into 50 physical resource blocks (PRB's, or k sub-bands), each consisting of 12 subcarriers. The FD multiplexing (scheduling) is done per one PRB at minimum.

We use the winning two step scheduling algorithm as proposed and described in [4], and illustrated in Figure 1. First step is TD scheduling followed by FD scheduling in the next step. The basic constraint for this algorithm is simple and can be formulated as the following requirement: The number of users scheduled per TTI (D_{buff}) should always be smaller than the number of users (D) in the cell. On the other hand, the number of UEs for frequency multiplexing should also be less than the number of PRBs in the system, since we are allocating one RB for each user. By the use of TD scheduler we have a simple controlling mechanism based on QoS of the users which is one of the crucial requirements for advanced packet scheduling [6]. Spectral efficiency, in turn, is obtained by the consequent FD scheduler. One of the main issues is that TD scheduling takes into account the instantaneous CQI over the whole bandwidth, instead of instantaneous supported throughput, and estimation of the average delivered throughput to a specific UE over average delivered throughput to all UEs. The scheduler picks users that have highest value of the priority metric at the scheduling time interval. The HARQ aware FD scheduler [4] then allocates the available sub-bands to the prioritized UEs from TD scheduler based on CQI feedback for different PRBs.

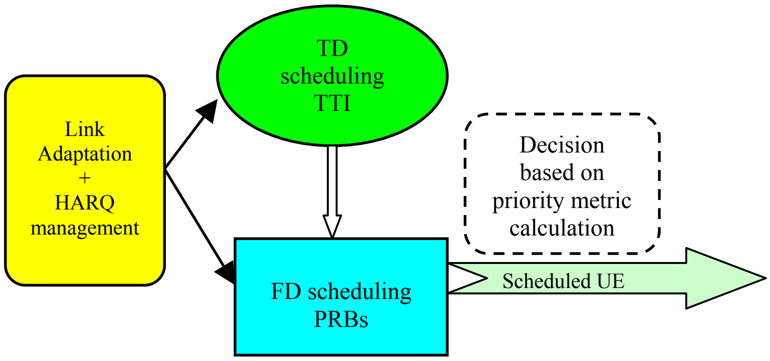


Figure 1: Joint time and frequency-domain scheduling process.

2.1 Proportional Fair (PF)

For the PF scheduler, scheduling decision per TTI is based on the following priority metric

$$RCQI_{i,k}(n) = P_{i.k}(n) = \frac{R_{i,k}(n)}{T_i(n)} \tag{1}$$

in which $R_{i,k}(n)$ is the instantaneous throughput of user i at sub-band k for the time instant n (TTI), $T_i(n)$ is the average delivered throughput to the UE i during the recent past and is calculated by

$$T_i(n) = \left(1 - \frac{1}{t_c}\right) T_i(n-1) + \frac{1}{t_c} R_{i,k}(n). \tag{2}$$

Here above, t_c is a parameter expressing the discrete time in TTI's over which the average delivered throughput is calculated. In general, $1/t_c$ is called the forgetting factor and has an initial value of 0.0025. $T_i(n-1)$ is the old value of T_i for the previous TTI, and $R_{i,k}(n)$ is the instantaneously delivered throughput to the UE i and sub-band k.

The basic algorithm can be finally summarized as follows:

1. For the D_{buff} selected users (available for scheduling) by the TD scheduler, the link adaptation unit is used to compute the instantaneously supported data rate for each sub-band. These users are sorted in descending order.
2. The scheduler selects the final target user from the above D_{buff} and a sub-band is assigned to the user with the highest priority metric.

2.2 PF- Modified

The proposed modified PF scheduler uses the following priority metric shown in (3):

$$RCQI_{i,k}(n) = P_{i.k}(n) = \left(\frac{CQI^i}{CQI^i_{avg}}\right)^{s1} \left(\frac{T_i(n)}{T_D(n)}\right)^{-s2} \tag{3}$$

Here, $s1$ and $s2$ are power coefficients from 0 to infinity, CQI^i is the CQI of user i and CQI^i_{avg} is the average CQI of user i calculated using

$$CQI^i_{avg} = \left(1 - \frac{1}{t_c}\right) CQI^{old,i}_{avg} + \left(\frac{1}{t_c}\right) CQI^i \tag{4}$$

Above, $T_D(n)$ is averaged delivered throughput (during the recent past) to all users (D) served by the BS and is calculated as

$$T_D(n) = \left(1 - \frac{1}{t_c}\right) T_D(n-1) + \frac{1}{t_c} \sum_{i=1}^{D} \frac{T_i(n)}{D} \quad (5)$$

As seen from above, the priority metric has two different terms. The first one aims to increase the total throughput and/or coverage, while the second one will improve only the coverage. Therefore the desired system performance can be tuned in a flexible manner by the use of the power coefficients *s1* and *s2*. The modified proportional fair algorithm uses the same overall structure as PF but the difference is in the priority metric calculations used in step 2.

3. System Simulation Model

Quasi-static system simulator for downlink has been created and used for several studies based on the OFDMA air interface according to [1]. 10MHz system bandwidth case is assumed and frequency reuse of 1 is used. The main simulation parameters and assumptions are summarized in the Table 1 for macro case 1 and follow the LTE working proposals. 18 UEs are uniformly distributed within each cell and experience inter-cell interferences from the surrounding cells. Every UE has an individual HARQ entry, which operates the physical layer retransmission functionalities. It is based on Stop-And-Wait (SAW) protocol and the number of entries is fixed to six. Modulation schemes supported are QPSK, 16QAM and 64QAM with variable rates for the encoder. Distance dependant path loss and shadow fading for a sector of cell site are taken into account for each PRB. Typical Urban channel is entirely described by a Power Delay Profile (PDP) and the fast fading characteristics updated once per TTI are based on it. Infinite buffer traffic model is applied in the simulations, i.e. every user has data to transmit for the entire duration of a simulation cycle. Exponential effective SINR metric (EESM) model is used for link-to-system level mapping, as described in [1].

Table 1: Default simulation parameters

Parameter	**Assumption**
Cellular Layout	Hexagonal grid, 19 cell sites, 3 sectors per site
Inter-site distance	500 m
Carrier Frequency / Bandwidth	2000MHz / 10 MHz
Channel estimation	Ideal
PDP	ITU Typical Urban 20 paths
Minimum distance between UE and cell	>= 35 meters
Average number of UEs per cell	18
Max. number of frequency multiplexed UEs	10
UE receiver	2-Rx MRC, IRC, SISO
Shadowing standard deviation	8 dB
UE speed	3km/h
Total BS TX power (*Ptotal*)	46dBm - 10MHz carrier
Traffic model	Full Buffer
Fast Fading Model	Jakes Spectrum
CQI reporting time	5 TTI
CQI delay	2 TTIs
MCS rates	QPSK (1/3, 1/2, 2/3),16QAM(1/2, 2/3, 4/5),64QAM(1/2, 2/3, 4/5)
ACK/NACK delay	2ms
Number of SAW channels	6
Maximum number of retransmissions	3
HARQ model	Ideal CC
1st transmission BLER target	20%
Forgetting factor	0.002
Scheduling schemes	PF, PF modified

4. Results

In this section we present the results obtained from the system simulations using the PS algorithms described in the paper. The performance is evaluated in terms of:

- Throughput - the total number of successfully delivered bits. Typically measured either in kbps or Mbps.
- CDF - Cumulative Distribution Function of UEs throughput. The slope of the CDF defines the fairness of the algorithms.
- Coverage – the experienced data rate per UE at the 95% coverage probability.
- EESM distribution of UEs

We present the behaviour of the modified PF scheduler with the use of different power coefficients *s1* and *s2* as shown in Table 2. As it was mentioned earlier, the main issue in the priority metric calculation is the first term from the equation (3). Therefore, we use large coefficient scale for *s1*, while the second power coefficient *s2* is here fixed to 1. The increase of *s2* does not lead to increasing the system performance, but only the relative values of *s1* and *s2* are important. More specifically, with large s2 values, the effect of second term in priority metric calculation based on throughput estimation would be emphasized and the scheduling algorithm would behave like equal resource scheduler. Furthermore, we also take into account the CQI imperfections as given in Table 1 when calculating the feedback, which is an important practical issue. Number of users used for frequency multiplexing is set to 10.

Table 2: Different combinations of power-coefficients

Coefficient	Value					
s1	1	1	2	6	10	20
s2	0	1	1	1	1	1

Figure 2 (left column) illustrates the average user throughput and coverage for the different schedulers and different UE receiver cases – single receiver antenna based on single input single output (SISO) principle, dual antenna interference rejection combining (IRC) and dual antenna maximal ratio combining (MRC) receivers. The power coefficient values from Table 2 are presented as index M, where M1 represents the first couple, etc. The used reference scheduler is TD-FD PF. By using the first term (M1) of the new metric calculation for modified PF scheduler we achieve both throughput and coverage gains of 5% and 15% correspondingly for SISO scenario as shown in (a). In the case of IRC and MRC presented in (c) and (e) we have coverage increases by 11% and 6% with no average throughput loss.

Continuing on the evaluation of modified PF scheduler, we clearly see a trade-off between average cell throughput and coverage for different power coefficient cases. The SISO scenario in (b) shows that with a loss of 15% in the cell throughput the coverage has been improved by 75% for power coefficient values (1,1). The same situation appears in the case of IRC (d) and MRC (f) where for the loss of 12% in cell throughput we gain a 65% increase in the user coverage. The remaining power coefficient values shown in Table 2 are used for tuning the overall scheduling performance. For the cases where *s1* varies between 2, 6 and 10 (M3, M4 and M5) the cell throughput loss is decreased to around 3%, which on the other hand corresponds to 7% -21% losses in the coverage (compared to earlier cases). Thus this establishes the principal behaviour of the new scheduling algorithm. The cell throughput being determined by the users experiencing the less favourable conditions is increasing and therefore the averaged system throughput is steadily growing by using coefficient values like M3, M4 and M5. In the same manner the user coverage is observed to have a steady monotonic decrease. Thus trade-off between average cell throughput and coverage is clearly seen and in order to preserve the average sector throughput and still to gain from the coverage increase coefficient values should be higher than M5. Consequently, for M6, the same average cell throughput is achieved as in case M2, but the coverage is increased by 10% corresponding to 20% gain for reference scheduler performance (case M1).

Figure 2 (right column) illustrates the CDF's of UE data rates for the different simulation scenarios and scheduling combinations. The slope of the CDF defines the fairness of the algorithms. Therefore we aim to achieve steeper slope corresponding to algorithm fairness. Similar slope change behaviour can clearly be established for each simulation scenario. Clearly, at for example 5% point of CDF curve, corresponding to users situated at the cell edge, we observe significant data rate increases indicated by shift to the right. This indicates improved overall cell coverage at the expense of slight total throughput loss.

Figure 3 illustrates the behaviour of the scheduling schemes for the simulated receiver types – (a) SISO, (b) IRC and (c) MRC in terms of the experienced SINR at UE based on the scheduler (left) and the usage of different modulation and coding scheme rates (right).The performance varies within the range of (-1,1) dB for the cases M1-M6. The modified PF scheme utilizes better higher order MCS, which is clearly visible in (b) and (c) cases. This further indicates the improved system performance achieved by employing the modified PF schemes.

5. Conclusions

In this work, we studied the potential of modified packet scheduling in time and frequency domains within the framework of UTRAN long term evolution (LTE). A new efficient and flexible packet scheduling algorithm based on modified PF principle was proposed. To evaluate the overall system performance and compare the schemes to each other, the throughput and coverage gains are assessed against more traditional scheduling. The efficient use of measured CQI reports with 1dB std. error provided from each UE is the key issue in the used PS schemes. In the case of fixed coverage requirements, the new metric calculation offers better control over the ratio between throughput loss and coverage increase. As practical examples, both throughput and coverage can be increased by several percentages, or by allowing a small decrease (in the order of only 5%) in the cell throughput, more than 40% increases in cell coverage are then available, due to the new priority metric calculation used in the proposed modified PF scheme.

References

[1] "Physical Layer Aspects for Evolved UTRA", 3GPP Technical Report TR 25.814, ver. 7.1.0, Oct. 2006.

[2] C. Wengerter, J. Ohlhorst, A.G.E Von Elbwert, "Fairness and Throughput Analysis for Generalized Proportional Fair Frequency Scheduling in OFDMA", IEEE Proc. VTC 2005-Spring, Stockholm, Sweden, May-2005.

[3] T. E. Kolding, "Link and system performance aspects of proportional fair scheduling in WCDMA/HSDPA", in Proc. IEEE Vehicular Technology Conference (VTC-2003), Orlando, FL, USA, Oct. 2003, pp. 1717-1723.

[4] A. Pokhariyal, K. I. Pedersen, G. Monghal, I. Z. Kovacs, C. Rosa, T. E. Kolding, and P. E. Mogensen, "HARQ aware frequency domain packet scheduler with different degrees of fairness for the UTRAN long term evolution," in Proc. IEEE Vehicular Technology Conference (VTC-2007), Dublin, Ireland, April 2007, pp. 2761-2765.

[5] A. Pokhariyal, T. E. Kolding, and P. E. Mogensen, "Performance of downlink frequency domain packet scheduling for the UTRAN long term evolution," in Proc. IEEE Personal, Indoor and Mobile Radio Communications Conference (PIMRC-2006), Helsinki, Finland, Sept. 2006.

[6] H. Holma and A. Toskala, Eds., HSDPA/HSUPA for UMTS – High Speed Radio Access for Mobile Communications. Wiley, 2006.

[7] T.E. Kolding, "QoS-aware proportional fair scheduling with required activity detection," in Proc. IEEE Vehicular Technology Conference (VTC-2006), Montreal, Canada, Sept. 2006.

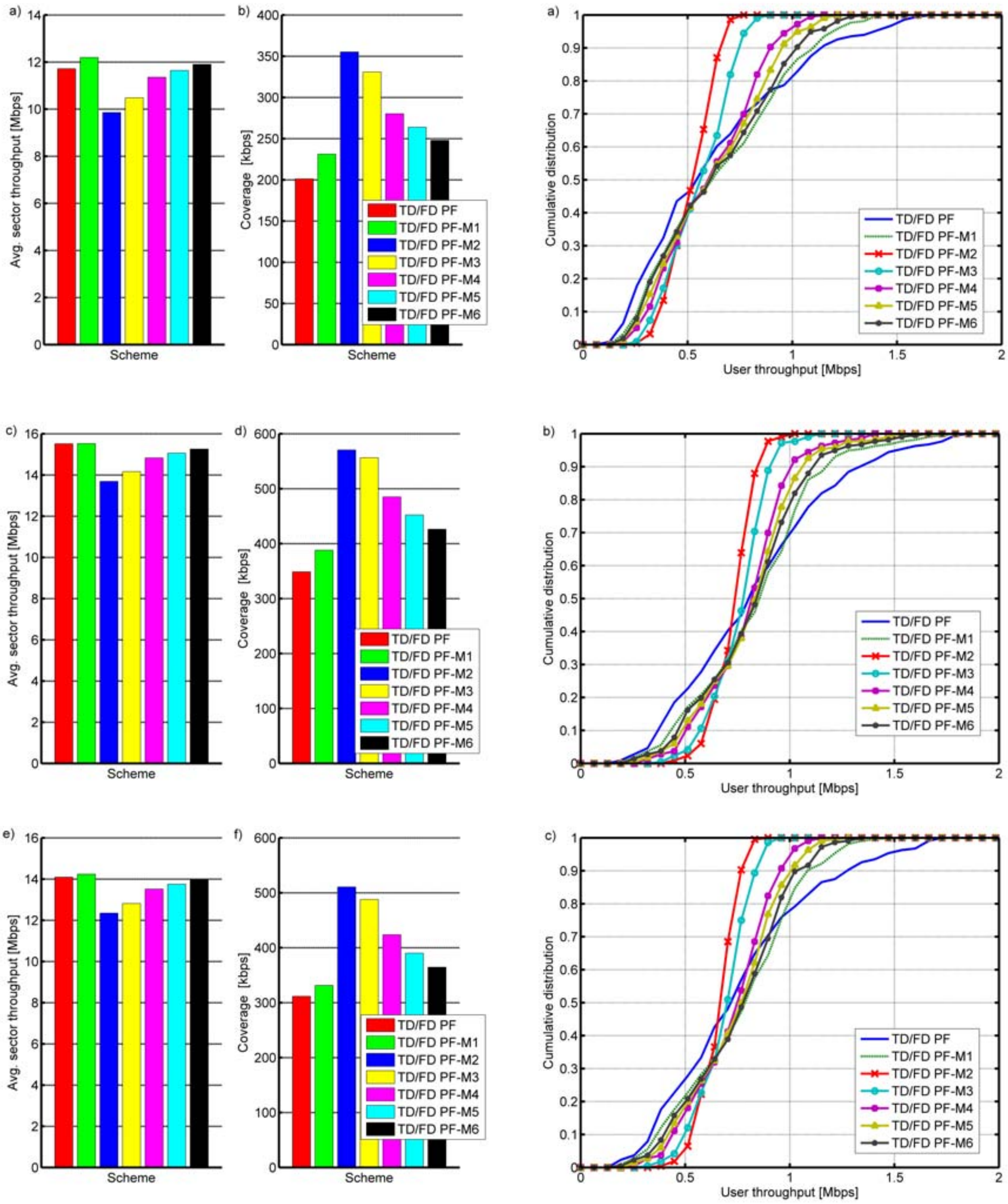


Figure 2: (left column) Average cell throughput and coverage gain over the reference PF scheduling scheme for the different receiver cases - SISO (a, b), IRC (c, d) and MRC (e, f) scenarios. The schemes M1-M6 refer to the new modified PF scheduler with power coefficient values as given in Table 1 (M1: s1=1, s2=0, etc.).

(right column): CDF's of User Throughput for three different receiver scenarios: (a) SISO, (b) IRC, and (c) MRC receiver cases.

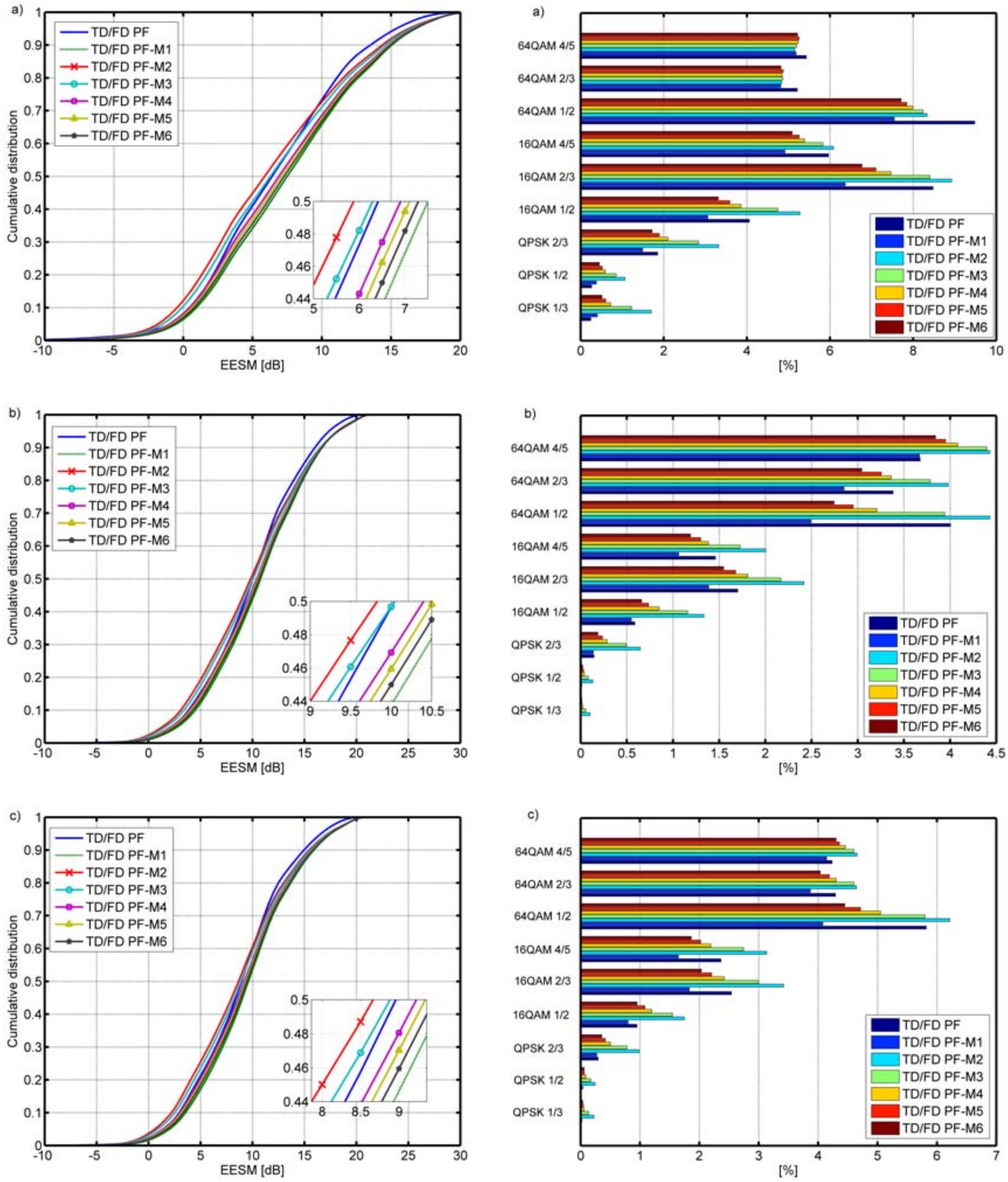


Figure 3: (left column) Experienced SINR at UE [dB] for the different receiver cases - SISO (a), IRC (b) and MRC (c) scenarios.

(right column) MCS distribution [%]: (a) SISO, (b) IRC, and (c) MRC receiver cases.

Publication 2

S. Nonchev and M. Valkama, “Efficient packet scheduling schemes for multiantenna packet radio downlink”, in *Proceedings of Fifth Advanced Int. Conf. Telecommunications* (AICT-2009),pp. 404-409, Venice, Italy, May 2009.

Efficient Packet Scheduling Schemes for Multiantenna Packet Radio Downlink

Stanislav Nonchev and Mikko Valkama
Tampere University of Technology,
Department of Communications Engineering,
Email: stanislav.nonchev@tut.fi, mikko.e.valkama@tut.fi

Abstract

In this paper, we propose fairness-oriented dual-stream multiple-input multiple-output (MIMO) packet scheduling schemes for future multi-antenna packet radio systems. In general, multi-user transmit-receive schemes allow users to be scheduled on different parallel streams on the same time-frequency resource. Based on that, implementations of more intelligent scheduling schemes that are aware of the instantaneous state of the radio channel require utilization of time, frequency and spatial domain resources in an efficient manner. Stemming from the earlier advanced Proportional Fair (PF) scheduler studies, we extend the developments to dual-stream MIMO packet radios. Furthermore, we investigate the resulting fairness distribution among users together with efficiency in increasing the system capacity in terms of throughput and coverage by simulating practical orthogonal frequency division multiple access (OFDMA) system environment with MIMO functionality in Micro and Macro cell scenarios. As a concrete example, we demonstrate that by using new multi-user scheduling schemes, significant coverage improvements in the order of 30% can be obtained at the expense of only 14% throughput loss. Furthermore, the user fairness is also greatly increased, by more than 18%, when measured using Jain's fairness index.

1. Introduction

The growth of new radio technologies for beyond 3G cellular systems continues. This includes e.g. 3GPP (3rd Generation Partnership Project) LTE (Long Term Evolution) [1], WiMAX [2] and the work in various research projects, like WINNER [3]. Some common elements in most of these developments are e.g. OFDMA based air interface, operating bandwidths of at least 10-20 MHz, and the exploitation of MIMO techniques and frequency domain channel-aware packet scheduling (FDPS) principles [1]. MIMO in terms of *Spatial Division Multiplexing* (SDM), combined by pre-coding, is considered as a promising technology due to increased spectral efficiency. In addition, it also provides scheduler with extra degree of freedom by offering a possibility to multiplex data of one or more users on the same physical resource. Thus the two principal concepts widely analyzed in literature [4]-[5] are single-user (SU) and multi-user (MU) MIMO. The SU-MIMO allow only one UE scheduled at the same time-frequency resource block (RB), while MU-MIMO provides flexibility so that multiple users can be scheduled on separate parallel streams over the same time-frequency RB. Considering relatively accurate channel state feedback in terms of channel quality indicator (CQI) reports from mobile station (MS) to base station (BS), together with fast link adaptation mechanisms, the new packet scheduling schemes have major impact on the performance optimization in terms of throughput and coverage. Another important feature of scheduling is fairness, implying that also users with less favorable channel conditions should anyway give some reasonable access to the radio spectrum. This is especially important in serving users at, e.g., cell edges in cellular networks. In this paper, we introduce a new scheduling scheme based on proportional fair (PF) principle and investigate its performance using multi-user receive-transmit scheme in 3GPP long term evolution (LTE) system context and in Micro and Macro Cell scenarios.

Recently, multi-antenna oriented packet scheduling principles have been started to be investigated in the literature, see e.g. [6]-[11]. New scheduling algorithms are proposed and their performance is evaluated in different simulator environments [8]-[10]. In this paper, we concentrate on the Proportional Fair (PF) principle, which in general offers an attractive balance between cell throughput and user fairness, and extend it with spatial domain functionality for the needs of MU-MIMO operation. More specifically, we extend our earlier studies [10] on algorithm development by deploying SDM functionality to frequency domain

(FD) PF scheduling schemes that can efficiently utilise the provided feedback information from all the user equipments (UEs), in terms of CQI. A new scheduler is proposed, called MIMO modified PF (MMPF), which is shown to improve the system performance considerably, in terms of coverage and overall user fairness. The performance evaluations of the scheduling methods presented in this paper are based on the system model according to the 3GPP evaluation criteria [1] and measured in terms of average cell throughput, cell coverage and Jain's fairness index [12].

In general, while the increased flexibility of channel-aware scheduling can offer performance enhancements, compared to fixed resource allocation, it also has some practical disadvantages. This includes e.g. relatively higher scheduling complexity, in terms of scheduling metric calculations and increased signaling overhead. Keeping these at reasonable levels requires thus some constraints on the scheduling algorithm, so for simplicity we assume here that only one MIMO mode (SU or MU) and fixed modulation and coding scheme (MCS) is allowed per user within one scheduling element.

The rest of the paper is organised as follows: Section 2 describes the MIMO channel-aware scheduling principles and proposed scheduling scheme. Section 3 gives an overview of the overall system model and simulation assumptions. The simulation results and analysis are presented in Section 4, while the conclusions are drawn in Section 5.

2. Channel-Aware MIMO Scheduling

In general, Packet Scheduler (PS) functionality is to select the most suitable users to access the radio resources based on some selected priority metric calculations. In order to do so, the scheduler interacts with other radio resource management (RRM) units such as *link adaptation* (LA) and *Hybrid ARQ* (HARQ) manager as shown in Figure 1. The scheduling decision is based on received users' signaling information in terms of acknowledgements and channel state information (CQI reports) per given transmission time interval (TTI) [9] and per frequency domain physical resource block (PRB). More specifically, we assume that both single-stream and dual-stream CQIs are reported to BS by each UE. Based on this information, the BS then decides, among other things, whether the particular time-frequency resource is used for (i) transmitting only one stream to a specific UE, (ii) two streams for a specific UE (SU-MIMO) or (iii) two streams to two different UEs (MU-MIMO).

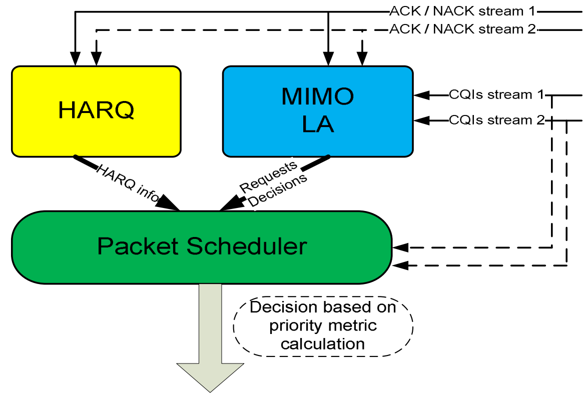


Figure 1. Joint time- and frequency-domain scheduling process.

2.1. Proposed Scheduler at Principal Level

In general, we use the well-known two-step PF approach with extended functionality in extra spatial dimension to enable MIMO operation [4],[10]. The first step is time-domain (TD) scheduling in which the scheduler selects the number of users in each TTI based on the full bandwidth channel state information. More specifically, the TD step selects those I_{BUFF} UE's (out of the total number of UE's, say I_{TOT}) whose total instantaneous throughput, per TTI, calculated over the full bandwidth is highest [11]. In this stage, we also take the different spatial multiplexing possibilities (one-stream, dual-stream SU, dual-stream MU) into account, in calculating all the possible reference throughputs.

The second step is then frequency-domain (FD) / spatial-domain (SD) scheduling in which the scheduler first reserves the needed PRB's for pending re-transmissions (on one stream-basis only) and the rest available PRB's are allocated to the selected UE's from the TD. The actual metric in FD/SD allocation is based on the PRB-level and stream-wise channel state information, and the corresponding throughput calculations.

2.2. Exact Scheduling Metrics

Here we describe the actual scheduling metrics used in ranking users in the TD scheduling step as well as mapping the users to FD/SD resources in the second step. First a multistream extension of "ordinary" PF is described in sub-section 2.2.1, used as a reference in the performance simulations, and then the actual proposed modified metric is described sub-section 2.2.2.

2.2.1 Multistream Proportional Fair: For the PF scheduler, scheduling decision per TTI is based on the following priority metric

$$\gamma_{i,k,s} = \underset{i}{arg\ max}\left\{\frac{R_{i,k,s}(n)}{T_i(n)}\right\} \tag{1}$$

in which $R_{i,k,s}(n)$ is the estimated instantaneous throughput of user i at sub-band k on stream s for the time instant (TTI) n (calculated based on the CQI reports through e.g. EESM mapping [1]). $T_i(n)$, in turn, is the average delivered throughput to the UE i during the recent past and is calculated by

$$T_i(n) = \left(1-\frac{1}{t_c}\right)T_i(n-1)+\frac{1}{t_c}R_i\ (n-1) \tag{2}$$

Here t_c controls averaging window length over which the average delivered throughput is calculated [7] and $R_i\ (n\text{-}1)$ is the actually delivered throughput to user i at previous TTI n-1, calculated over all sub-bands k and streams s. In general, $1/\ t_c$ is also called the forgetting factor.

Considering the previous TD and FD/SD steps described in Section 2.1, the above metrics are used as follows:

- TD: Metric (1) is evaluated over the full bandwidth and for different stream options to rank the I_{TOT} UE's. Out of these, $I_{\text{BUFF}} < I_{\text{TOT}}$ UE's with highest metric are picked to the following FD/SD stage. In the following, this subset is called scheduling candidate set (SCS), and is denoted by $\Omega(n)$.

- FD/SD: The access to individual PRB and stream(s) is granted for the user belonging to the above SCS with the highest metric (1) evaluated for the particular PRB and stream at hand.

2.2.2 Modified Multistream Proportional Fair (MMPF): Stemming from the earlier work in [10], the following modified multistream PF metric is proposed:

$$\overline{\gamma}_{i,k,s} = \underset{i}{arg\ max}\left\{\left(\frac{CQI_{i,k,s}(n)}{CQI_i^{avg}(n)}\right)^{\alpha 1}\left(\frac{T_i(n)}{T_{tot}(n)}\right)^{-\alpha 2}\right\} \tag{3}$$

Here, $\alpha 1$ and $\alpha 2$ are scheduler optimization parameters ranging basically from 0 to infinity, $CQI_{i,k,s}$ is the CQI of user i at PRB k and stream s, and $CQI^{avg}{}_i$ is the average CQI of user i calculated using

$$\begin{aligned} CQI_i^{avg}(n) &= \left(1-\frac{1}{t_c}\right)CQI_i^{avg}(n-1) \\ &+\frac{1}{t_c}\frac{1}{K_{TOT}}\frac{1}{S_{TOT}}\sum_{s=1}^{S_{TOT}}\sum_{k=1}^{K_{TOT}}CQI_{i,k,s}(n) \end{aligned} \tag{4}$$

In above, K_{TOT} is the total number of available PRB's while S_{TOT} denotes the number of streams which is here always two (max two streams). In (3), $T_{tot}(n)$ is the average delivered throughput (during the recent past) to all users I_{BUFF} served by the BS and is calculated as

$$T_{tot}(n) = \left(1-\frac{1}{t_c}\right)T_{tot}(n-1)+\frac{1}{t_c}\frac{1}{I_{BUFF}}\sum_{i\in\Omega(n-1)}T_i(n-1) \tag{5}$$

Intuitively, the proposed scheduling metric in (3) is composed of two elements affecting the overall scheduling decisions. The first dimension measures in a stream-wise manner the relative instantaneous quality of the individual user's radio channels against their own average channel qualities while the second dimension is related to measuring the achievable throughput of individual UE's against the corresponding average throughput of scheduled users. In this way, and by understanding the power coefficients $\alpha 1$ and $\alpha 2$ as additional adjustable parameters, the exact scheduler statistics can be tuned and controlled to obtain a desired balance between the throughput and fairness. This will be demonstrated later using computer simulations. Considering finally the actual TD and FD/SD steps described at general level in Section 2.1, the same approach as in sub-section 2.2.1 is deployed but the metrics (1)-(2) are of course here replaced by the metrics in (3)-(5).

3. System Simulator

System-level performance of the proposed scheduling scheme is evaluated based on developed quasi-static system simulator for LTE downlink providing traffic modeling, multiuser packet scheduling and link adaptation [1]. As a practical example, the 10 MHz system bandwidth case of LTE is assumed, meaning that there are 50 physical resource blocks each consisting of 12 sub-carriers with sub-carrier spacing of 15 kHz. This sets also the basic resolution in FD/SD UE multiplexing (scheduling), i.e., the allocated individual UE bandwidths are multiples of the PRB bandwidth. The actual reported CQI's are based on received signal-to-interference-and-noise ratios (SINR), calculated by the UE's for each PRB. Here the UE's are assumed to use linear MMSE (LMMSE) receiver principle, and utilize the in SINR calculations the actual radio channel response, the received noise level, and the structure of the detector. Furthermore, like mentioned already earlier, the UE's always report single-stream SINR as well as both SU and MU dual-stream SINR's (at the corresponding detector output).

In a single simulation run, mobile stations are randomly dropped or positioned over each sector and cell. Then based on the individual distances between the mobile and the serving base station, the path losses for individual links are directly determined, while the actual fading characteristics of the radio channels depend on the assumed mobility and power delay profile. In updating the fading statistics, the time resolution in our simulator is set to one TTI (1ms). In general, a standard hexagonal cellular layout is utilized with altogether 19 cell sites each having 3 sectors in Macro case and 1 sector in Micro case. In the performance evaluations, statistics are collected only from the central cell site while the others simply act as sources of inter-cell interference.

The main simulation parameters and assumptions are generally summarized in the Table 1 for both Macro and Micro cell scenarios, following again the LTE working assumptions. The used MIMO scheme is *per- antenna rate control* (PARC) with two transmit antennas at a BS and two receive antennas at the UE and the receivers are equipped with LMMSE detectors. As illustrated in Figure 1, the RRM functionalities are controlled by the packet scheduler together with link adaptation and HARQ mechanisms. Notice that the maximum number of simultaneously multiplexed users (I_{BUFF}) is set to 6 here while I_{TOT} is 10. In general, we assume that the BS transmission power is equally distributed among all PRB's. In the basic simulations, 10 UE's are uniformly dropped within each cell and experience inter-cell interferences from the surrounding cells, in addition to path loss and fading. The UE velocity equals 3km/h, and the typical urban (TU) channel model standardized by ITU is assumed in modeling the power-delay spread of the radio channels. Infinite buffer traffic model is applied in the simulations, i.e. every user has data to transmit for the entire duration of a simulation cycle. Exponential effective SINR mapping (EESM) is used for link-to-system level mapping (throughput calculations), as described in [1]. The length of a single simulation run is set to 5 seconds which is then repeated for 10 times to collect reliable statistics.

Considering MIMO functionality, every UE has an individual HARQ entry per stream, which operates the physical layer re-transmission functionalities. It is based on the stop-and-wait (SAW) protocol and for simplicity, the number of entries per UE is fixed to six. HARQ retransmissions are always transmitted with the same MCS and on the same PRB's (if scheduled) as the first transmissions in a single-stream mode. The supported modulation schemes are QPSK, 16QAM and 64QAM with variable rates for the encoder as shown in Table 1.

Link adaptation handles the received UE reports containing the channel quality information for each PRB based on single and dual-stream MIMO modes. The implemented link adaptation mechanism consists of two separate elements – the inner loop (ILLA) and outer loop (OLLA) LA's – and are used for removing CQI imperfections and estimating supported data rates and MCS. It is assumed that the CQI reporting errors are log-normal distributed with 1dB standard deviation.

Table 1. Basic simulation parameters.

Parameter	Assumption
Cellular Layout	Hexagonal grid, 19 cell sites, 3 sectors per site for Macro / 1 sector per site for Micro
Inter-site distance	500 m - Macro / 1732 m - Micro
Carrier Frequency / Bandwidth	2000MHz / 10 MHz
Number of active sub-carriers	600
Sub-carrier spacing	15kHz
Sub-frame duration	0.5 ms
Channel estimation	Ideal
PDP	ITU Typical Urban 20 paths
Minimum distance between UE and cell	>= 35 meters - Macro >= 10 meters - Micro
Average number of UE's per sector	20
Max. number of frequency multiplexed UEs (I_{BUFF})	10
UE receiver type	LMMSE
Shadowing standard deviation	8 dB
UE speed	3km/h
Total BS TX power (Ptotal)	46dBm
Traffic model	Full Buffer
Fast Fading Model	Jakes Spectrum
CQI reporting schemes	Full CQI
CQI log-normal error std.	1 dB
CQI reporting time	5 TTI
CQI delay	2 TTIs
CQI quantization	1 dB
CQI std error	1 dB
MCS rates	QPSK (1/3, 1/2, 2/3), 16QAM (1/2, 2/3, 4/5), 64QAM (1/2, 2/3, 4/5)
ACK/NACK delay	2ms
Number of SAW channels	6
Maximum number of retransmissions	3
HARQ model	Ideal chase combining (CC)
1st transmission BLER target	20%
Scheduler forgetting factor	0.0025
Scheduling schemes used	Ordinary PF Modified PF (proposed)
Simulation duration (one drop)	5 seconds
Number of drops	10

4. Numerical Results

In this section, we present the results obtained from the system simulations using the PS algorithms described in the paper. The system-level performance is generally measured and evaluated in terms of:

- Throughput - the total number of successfully delivered bits per unit time. Usually measured either in kbps or Mbps.
- Coverage – the experienced data rate per UE at the 95% coverage probability.
- Fairness per scheduling scheme measured using Jain's fairness index

We illustrate the behavior of the MMPF scheduler with using different power coefficients *α1* and *α2* as shown in Table 2. To focus mostly on the role of the channel quality reporting in the priority metric calculation in equation (3), *α2* is fixed here to 1 and different values are then demonstrated for *α1*. More specifically, with large *α2* values, the effect of second term in priority metric calculation based on throughput estimation would be emphasized and the scheduling algorithm would behave like maximum throughput scheduler, which implies reduced fairness distribution.

Figure 2 illustrates the average user throughput and coverage for the different schedulers and different simulation cases – Macro cell and Micro cell. The power coefficient values from Table 2 are presented as index M, where M1 represents the first couple, etc. The obtained results are compared with the reference PF scheduler described in Section 2. By using the first term (M1) of the new metric calculation for modified PF scheduler, we achieve coverage gains in the order of 75% at the expense of 23% throughput loss for Macro case scenario presented in (a) and (b). In Micro case scenario illustrated on (c) and (d) we obtain even 11% more coverage gain for the same throughput loss as in Micro case scenario.

Continuing the evaluation of the proposed scheduler, we clearly see a trade-off between average cell throughput and coverage for different power coefficient cases. Furthermore, the remaining power coefficient values are used for tuning the overall scheduling performance. For the rest of the cases, the cell throughput loss is decreased stepwise with around 3% per index M, which on the other hand corresponds to coverage gains reducing from 47% to 11%. Consequently, an obvious trade-off between average cell throughput and coverage is clearly seen. In order to preserve the average sector throughput and still to gain from the coverage increase, coefficient values should thus be properly chosen. Similarly, in Micro case scenario we obtain coverage gains between 26% to 67% corresponding to throughput losses from 11% to 20% as illustrated by the performance statistics.

Table 2. Different power coefficient combinations

Coefficient	Value				
α1	1	2	4	6	8
α2	1	1	1	1	1

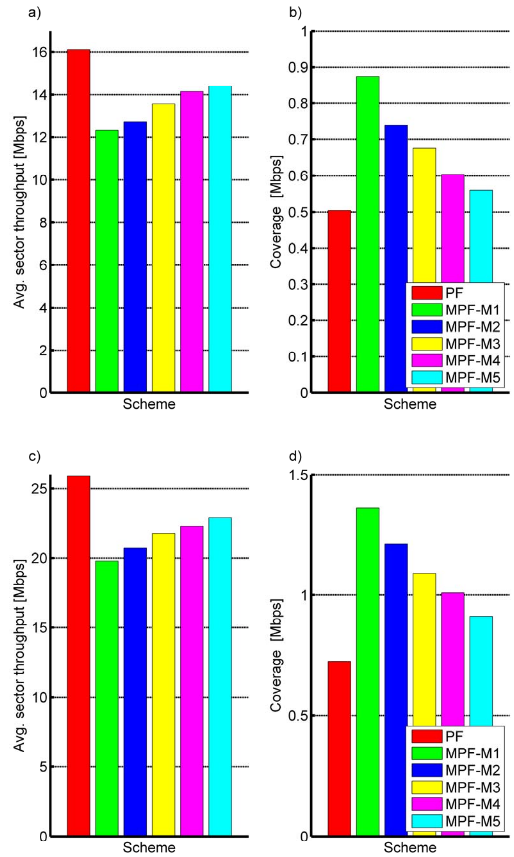


Figure 2. Average cell throughput and coverage gain over the reference PF scheduling scheme for the different simulation scenarios – Macro cell (a, b) and Micro cell (c, d). The schemes M1-M5 refer to the new MMPF scheduler with power coefficient values as given in Table 2.

Figure 3 illustrates the Jain's fairness index per scheduling scheme for Micro and Macro cell scenarios, calculated over all the I_{TOT} = 10 UE's. The value on the x axis corresponds to used scheduler type, where 1 refers to the reference PF scheduler, 2 refers to MMPF with index M1, etc. The value of Jain's fairness index is generally in the range of [0,1], where value of 1 corresponds to all users having the same amount of resources. Clearly, the fairness distribution with MMPF outperforms the used reference PF scheduler for both cases. The received fairness gains are in range of 18%-32% in the Macro case and 25%-34% in the Micro case.

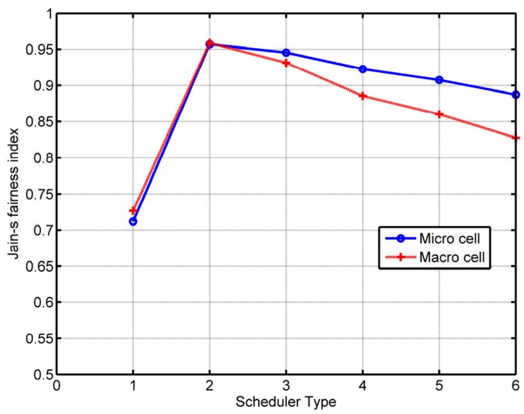


Figure 3. Jain's fairness index per scheduling scheme for Macro and Micro cases. Scheduler type 1 means ordinary PF, while 2-6 means proposed PF with power coefficients as described in Table 2.

5. Conclusions

In this paper, we have studied the potential of advanced multi-user packet scheduling algorithms in OFDMA type radio system context, using UTRAN long term evolution (LTE) downlink in Macro and Micro cell environment as practical example cases. New multi-stream proportional fair scheduler metric covering time-, frequency- and spatial domains was proposed that takes into account both the instantaneous channel qualities (CQI's) as well as resource allocation fairness. Overall, the achieved throughput performance together with coverage and fairness statistics were assessed, by using extensive system simulations, and compared against more traditional proportional fair scheduling with SDM functionality. In the case of fixed coverage requirements, the proposed scheduling metric calculations based on UE channel feedback offers better control over the ratio between the achievable cell/UE throughput and coverage increase, as well as increased UE fairness. As a practical example, the fairness in resource allocation together with cell coverage can be increased significantly (more than 30%) by allowing a small decrease (in the order of only 10-15%) in the cell throughput.

6. References

[1] "Physical Layer Aspects for Evolved UTRA", *3GPP Technical Report TR 25.814*, ver. 7.1.0, Oct. 2006.

[2] WiMAX Forum™, http://www.wimaxforum.org, Aug. 2006.

[3] "Wireless World Initiative New Radio (WINNER)," IST-507581, European Information Society Technologies, Technical Report, Dec. 2006.

[4] N. Wei, B. Talha, T. B. Sørensen, T. E. Kolding, and P. E. Mogensen, "Spectral efficiency of closed-loop transmit diversity with limited feedback for UTRA long term evolution", in *The 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Helsinki, Sep. 2006.

[5] R. W. Heath, M. Airy, and A. Paulraj, "Multiuser diversity for mimo wireless systems with linear receivers", in *Signals, Systems and Computers,2001. Conference Record of the Thirty-Ffth Asilomar Conference on*, vol. 2, 2001, pp. 1194-1199, vol. 2.

[6] C. Wengerter, J. Ohlhorst, and A.G.E Von Elbwert, "Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA", in *Proc. IEEE Vehicular Technology Conference* (VTC 2005-Spring), Stockholm, Sweden, May-2005.

[7] T. E. Kolding, "Link and system performance aspects of proportional fair scheduling in WCDMA/HSDPA", in *Proc. IEEE Vehicular Technology Conference (VTC-2003)*, Orlando, FL, USA, Oct. 2003, pp. 1717-1723.

[8] P. Svedman, S.K.Wilson, L.J. Cimini, and B. Ottersten, "A simplified opportunistic feedback and scheduling scheme for OFDMA", in *Proc. IEEE Vehicular Technology Conference IEEE*, pp. 1878-1882, May 2004.

[9] Y. Sun, et al., "Multi-user scheduling for OFDMA downlink with limited feedback for evolved UTRA", in *Proc. IEEE Vehicular Technology Conference* (VTC-2006 Fall), September 2006

[10] S. Nonchev, J. Venäläinen and M. Valkama, "New frequency domain packet scheduling schemes for UTRAN LTE downlink", in *Proc. ICT Mobile Summit*, Stockholm, Sweden, June 2008

[11] N. Wei, et al., "Performance of MIMO with frequency domain packet scheduling", in *Proc. IEEE Vehicular Technology Conference*, May 2007.

[12] D. Chui and R. Jain,"Analysis of the increase and decrease algorithms for congestion avoidance in computer networks", *Computer Networks and ISDN Systems*, 1989.

Publication 3

S. Nonchev and M. Valkama, “Efficient power-aware packet scheduling for multiantenna packet radio systems,” in *Proceedings of First International Conference on Networks and Communications* (NETCOM-2009), pp. 344-348, Chennai (Madras), India, December 2009.

Efficient Power-Aware Packet Scheduling for Multiantenna Packet Radio Systems

Stanislav Nonchev and Mikko Valkama
Department of Communications Engineering
Tampere University of Technology
P.O.Box 553, FI-33101, Tampere, FINLAND
Email: stanislav.nonchev@tut.fi, mikko.e.valkama@tut.fi

In this paper, we propose fairness-oriented packet scheduling (PS) schemes with power-efficient control mechanism for future packet radio systems. In general, multi-antenna transmit-receive schemes provide additional flexibility to packet scheduler functionality. Stemming from the earlier enhanced proportional fair scheduler studies for single-input multiple-output (SIMO) and multiple-input multiple-output (MIMO) systems, we extend the development of efficient packet scheduling algorithms by adding transmit power considerations in the overall priority metrics calculations and scheduling decisions. Furthermore, we evaluate the proposed scheduling schemes by simulating practical orthogonal frequency division multiple access (OFDMA) based packet radio system in terms of throughput, coverage and fairness distribution among users. As a concrete example, under reduced overall transmit power constraint and unequal power distribution for different sub-bands, we demonstrate that by using the proposed power-aware multi-user scheduling schemes, significant coverage and fairness improvements in the order of 70% and 20% can be obtained, at the expense of average throughput loss of only 15%.

I. INTRODUCTION

Development of new advanced packet radio systems continues progressively. Relevant work in this direction includes, e.g., 3GPP long term evolution (LTE) [1], WiMAX [2], IMT-Advanced [3] and the related work in various research projects, like WINNER [4]. These developments rely on OFDMA air interface, scalable bandwidth operation, exploitation of MIMO technologies and advanced convergence techniques. The major driving forces behind these developments are the increased radio system performance, in terms of average and peak cell throughputs, low latency and reduced operating expenditures. While the average and peak throughputs are typically emphasized, also the fairness and cell-edge coverage are equally important quality measures of cellular radio systems [5], [6]. Yet another key measure becoming all the time more and more important is the energy consumption of the radio access network.

In general, all user equipments (UEs) within base station (BS) coverage are sharing the available radio resource (radio spectrum). This is controlled by the packet scheduler (PS) utilizing selected scheduling metrics. In both uplink (UL) and downlink (DL), the packet scheduler works in a centralized manner being located at the BS. It is generally fairly well understood that the overall radio system performance, in terms of throughput and fairness, depends heavily on the PS functionality. The PS operation can build on instantaneous radio channel conditions, quality of service (QoS) requirements and traffic situations of the served UEs [5], [6]. In the emerging radio systems, multi-antenna MIMO technologies increase spectral efficiency and also provide the PS with extra degree of freedom by offering a possibility to multiplex data of one or more users on the same physical resource (spatial domain multiplexing, SDM) [7], [8]. Such functionality requires additional feedback information provided by mobile stations (MS) to the base station, typically in terms of channel quality information (CQI) reports. The main disadvantage of MIMO-SDM technique is increased signaling overhead and scheduling complexity. Finally, an additional important aspect in scheduling functionality is the interaction with other radio resource management (RRM) entities, namely fast link adaptation (LA) and reliable re-transmission mechanisms.

Packet scheduling principles are generally rather widely investigated in the literature and many new scheduling algorithms have been proposed, see, e.g., [9]-[13]. Most of them are considering equal BS transmission power distribution among all physical resource blocks (PRBs), which is not necessarily a practical or optimum case. In practice, unequal power allocation between different PRB's can be used, e.g., to control the interference between neighboring cells or sectors in frequency re-use 1 radio systems. Stemming now from our previous work in advanced PS algorithms reported in [12]-[13], we extend the studies here to incorporate energy efficiency considerations in the scheduling decisions. Our starting point is the advanced modified proportional fair scheduler for MIMO reported in [13], which was shown to improve cell-edge coverage and user fairness compared to state-of-the-art. In this paper, we introduce a new energy- or power-aware scheduling scheme taking into account the applied power pattern within the overall radio spectrum. This is called MIMO power-aware modified proportional fair (MPMPF) scheduler in the continuation. The performance of the proposed scheduler is investigated using both single-user MIMO (spatial multiplexing for individual UE streams) and multi-user MIMO (spatial multiplexing also for streams of different UEs) transmit-receive schemes in 3GPP LTE system context. The used system-level figures of merit are cell throughput distribution, cell-edge coverage and Jain's fairness index [14]. For simplicity and illustration purposes, 1x2 (SIMO) and 2x2 (MIMO) multi-antenna scenarios are assumed in the continuation.

The rest of the paper is organized as follows: Section II describes the MIMO scheduling principles and proposed power-aware scheduling scheme. The overall radio system simulation methodology and simulation assumptions are described in Section III. The simulation results and analysis are presented in Section IV, while the conclusions are summarized in Section V.

II. POWER-AWARE MIMO SCHEDULING

In general, packet scheduler is part of the overall radio resource management mechanism at BS. Therefore, PS functionality depends heavily on the collaboration with other RRM units such as *link adaptation* (LA) and *Hybrid ARQ* (HARQ) manager, as depicted in Figure 1. The scheduling decisions are based on the selected scheduling metric, utilizing typically the CQI reports and acknowledgements from UEs, per given transmission time interval (TTI) in time domain and per physical resource block (PRB) in frequency domain [5]. Moreover, we assume here that the LA unit can feed the PS with power allocation information on a PRB-per-PRB basis. MIMO functionality, in turn, requires both single-stream and dual-stream CQI feedback by each UE. Consequently, the BS decides whether the particular time-frequency resource is used for transmitting (*i*) only one stream to a specific UE, (*ii*) two streams to a specific UE (SU-MIMO) or (*iii*) 1+1 streams to two different UEs (MU-MIMO) [6]. BS also handles proper transmit power allocation in all the cases (*i*)-(*iii*), respectively, such that target packet error rate is reached with selected modulation and coding scheme (MCS).

A. Proposed Scheduler at Principal Level

The proposed scheduler developments are stemming from the widely used two-stage (see e.g. [15], [16]) multi-stream PF approach. In the first stage, within each TTI, UEs are ranked based on the full bandwidth channel state information and throughput calculations. This is called time-domain (TD) scheduling step. For MIMO case, different spatial multiplexing possibilities (one-stream, dual-stream SU, dual-stream MU) are taken into account, in calculating all the possible reference throughputs. In the second stage, the scheduling functionality is expanded in frequency-domain (FD) and spatial-domain (SD) where the actual PRB allocation takes place. First the needed PRB's for pending re-transmissions (on one stream-basis only) are reserved and the rest available PRB's are allocated to the selected UE's from the first stage. The actual priority metric in FD/SD stage is evaluated at PRB-level taking into account the available stream-wise channel state information, the transmit power allocations and the corresponding throughput calculations. The exact scheduling metrics are described below.

B. Scheduling Metrics

Here we describe the actual scheduling metrics used in ranking users in the TD scheduling stage as well as mapping the users to FD/SD resources in the second stage. First a

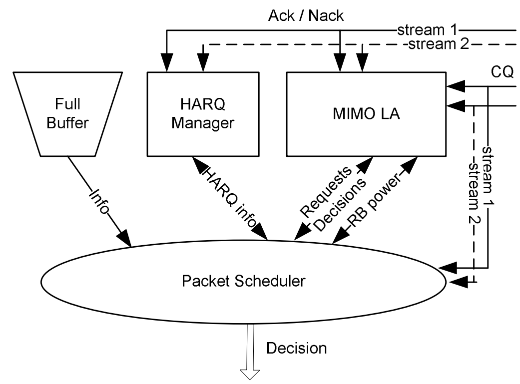


Figure 1. Principal RRM functionalities.

power-aware extension of ordinary multi-stream PF scheduler is described in sub-section B1, used as a reference in the performance simulations. Then the actual proposed modified power-aware metric is described in sub-section B2.

1) Power-Aware Multistream Proportional Fair:

Stemming from ordinary multi-stream PF principle, the following power-aware PF (PPF) scheduling metric evaluated at each TTI is deployed

$$\gamma_{i,k,s} = \underset{i}{arg\ max}\left\{\frac{P_{k,s}}{P_{\max}}\frac{R_{i,k,s}(n)}{T_i(n)}\right\} \qquad (1)$$

Here $R_{i,k,s}(n)$ is the estimated instantaneous throughput of user i at sub-band k on stream s at TTI n. $T_i(n)$ is the corresponding average delivered throughput to the UE i during the recent past [12],[13]. $P_{k,s}$, in turn, is the transmission power for sub-band k on stream s and P_{max} is the maximum transmission power for any sub-band. Notice that the power ratio $P_{k,s} / P_{max}$ has a direct impact on the ranking of the users, since the deployed sub-band power levels affect the corresponding estimated and delivered throughput quantities. The scheduler in (1) is only used as a reference in evaluating the performance of the actual proposed scheduler to be described below.

2) Modified Power-Aware Multi-stream Proportional Fair (MPMPF):

Starting from the earlier work in [13], the following modified power-aware multi-stream PF metric is proposed in this paper:

$$\overline{\gamma}_{i,k,s} = \underset{i}{arg\ max}\left\{\frac{P_{k,s}}{P_{\max}}\left(\frac{CQI_{i,k,s}(n)}{CQI_i^{avg}(n)}\right)^{\alpha_1}\left(\frac{T_i(n)}{T_{tot}(n)}\right)^{-\alpha_2}\right\} \qquad (2)$$

In the above metric, α_1 and α_2 are scheduler optimization parameters ranging basically from 0 to infinity. $CQI_{i,k,s}$ is the CQI of user i at PRB k and stream s, and CQI^{avg}_i is the average CQI of user i. $T_{tot}(n)$ is the average delivered throughput (during the recent past) to all users ranked in TD stage served by the BS.

The proposed scheduling metric in (2) is essentially composed of three elements affecting the overall scheduling decisions. The first ratio takes into account the transmit power fluctuations in BS for each PRB. The second ratio is the relative instantaneous quality of the individual user's radio channels over their own average channel qualities in a stream-wise manner. The third ratio is related to measuring the achievable throughput of individual UE's against the corresponding average throughput of scheduled users. The power coefficients α_1 and α_2 are additional adjustable parameters that can be tuned and controlled to obtain a desired balance between throughput and fairness. This will be illustrated in Section IV.

The basic idea of incorporating the sub-band power ratio $P_{k,s}/P_{max}$ into the scheduling metric (2) is that the sub-band power levels obviously affect the link adaptation and thereon the estimated supportable throughput as well as the actual delivered throughput. Thus by taking the power fluctuations into account, we seek higher fairness in the scheduling between truly realized UE throughputs, and thereon better cell-edge coverage. This will be demonstrated in Section IV. Also since the power is understood in a stream-wise manner, this gives somewhat higher priority to single-stream transmissions which also helps in increasing the coverage.

III. System Simulation Methodology

Extensive quasi-static system simulator for LTE downlink providing traffic modeling, multiuser packet scheduling and link adaptation including HARQ is used for evaluating the system level performance of the proposed packet scheduling scheme, following 3GPP evaluation criteria [1]. The 10 MHz system bandwidth case is assumed, being composed of 1024 sub-carriers (out of which 600 are active) and divided into 50 physical resource blocks (PRB) each consisting of 12 sub-carriers with sub-carrier spacing of 15 kHz. Pilot signals are sent from base station to mobile station to determine the instantaneous channel condition. The mobile stations measure the actual channel states and the information is reported to the BS. The actual reported CQI's are based on received signal-to-interference-and-noise ratios (SINR), calculated by the UE's for each PRB. Here the UE's are assumed to use two different receiver principles – maximal ratio combining (MRC) or linear MMSE (LMMSE), depending whether 1x2 or 2x2 system is simulated. Additionally, the UEs always report single-stream SINR as well as both single user (SU) and multi-user (MU) dual-stream SINR's (in the 2x2 case) at the corresponding detector output. As a concrete example of unequal PRB power allocation for inter-cell or inter-sector interference management, the case of reducing the transmit power of every third RB by 1dB and the neighboring ones yet by another 1dB (i.e. the relative power pattern is 0dB, -1dB, -2dB, 0dB, -1dB, -2dB, …) compared to

TABLE I. Basic radio system simulation parameters.

Simulation Parameters	
Parameter	***Assumption***
Carrier Frequency / Bandwidth	2000MHz / 10 MHz
Number of active sub-carriers	600
Sub-carrier spacing	15kHz
Sub-frame duration	0.5 ms
Simulation scenario	Macro cell hexagonal 19-site layout
Channel estimation	Ideal
PDP	ITU Typical Urban 20 paths
Number of UE's per sector	10
Max. number of frequency multiplexed UEs (I_{BUFF})	6
UE receiver type	2-Rx MRC and LMMSE
UE speed	3km/h
Total BS TX power (P_{total})	46dBm
Traffic model	Full Buffer
Fast Fading Model	Jakes Spectrum
MCS rates	QPSK, 16QAM ,64QAM
ACK/NACK delay	2ms
Number of SAW channels	6
Maximum number of retransmisions	3
HARQ model	Ideal chase combining (CC)
1st transmission BLER target	20%
Scheduler forgetting factor	0.002
Scheduling schemes used	Ordinary PF, modified PF [13], power-aware PF (1), and modified power-aware PF (2)
Simulation duration (one drop)	5 seconds
Number of drops	10

maximum transmit power is deployed. Notice that this is just one concrete example selected for evaluating and comparing the performance of different schedulers.

In a single simulation run, mobile stations are randomly distributed over standard hexagonal cellular layout with altogether 19 cells each having 3 sectors. As a concrete example, the number of active users in the cell is set to 10 and the UE velocities equal 3km/h. The path losses for individual links are directly determined based on the individual distances between the mobile and the serving base station. On the other hand, the actual fading characteristics of the radio channels are collected each TTI (1ms) and depend on the assumed mobility and power delay profile. Due to the centralized approach statistics are collected only from the central cell site while the others simply act as sources of inter-cell interference.

Two different multi-antenna scenarios are used for performance evaluation purposes – SIMO and MIMO. The first scheme consist of one transmit antenna at the BS and two receive antennas at each UE and the receivers are equipped with MRC detectors. The used MIMO scheme, in turn, is per-antenna rate control (PARC) with two transmit antennas at the BS and two receive antennas at each UE and the receivers are equipped with LMMSE detectors. The main simulation parameters and assumptions are summarized in Table I.

The RRM functionalities are controlled by the packet scheduler together with link adaptation and HARQ entities. Link adaptation consists of two separate elements – the inner loop LA (ILLA) and the outer loop LA (OLLA). These are used for removing CQI imperfections, estimating supported data rates and MCS's, and stabilizing the 1st transmission Block Error Probability (BLEP) to the target range (typically 10-20%). Simple admission control scheme is used for keeping the number of UEs per cell constant. HARQ is based on SAW protocol and a maximum number of three re-transmissions is allowed. MIMO functionality requires individual HARQ entry per stream which is also implemented. Link-to-system level mapping is based on the effective SINR mapping (EESM) principle [1].

IV. Obtained Results

This section presents the obtained results from the radio system simulations using different PS algorithms. The system-level performance evaluation is based on the following statistics:

- Cell throughput (kbps or Mbps) distribution
- Cell-edge coverage
- Fairness distribution
- CDF of the number of users scheduled per TTI

Throughput is defined as the number of successfully delivered user bits per unit time. Coverage, in turn, corresponds to 5% probability point in the throughput CDF. Finally fairness is measured using the Jain's fairness index [14].

We demonstrate the behavior of the proposed MPMPF scheduler by using different power coefficients α_1 and α_2 and comparing it against other PF scheduling algorithms. To emphasize the role of power-aware and CQI based priority metric calculation in (2), we fix the value of α_2 to 1 and change the values of α_1 as $\alpha_1 = \{1,2,4\}$ [10], [13].

Figure 2 illustrates the average cell throughput and cell-edge coverage for the different schedulers in SIMO (sub-figures (a) and (b)) and MIMO (sub-figures (c) and (d)) system simulation cases. The power coefficient values are presented as index M, where M1 represents the first couple, i.e., $\alpha_1=1$, $\alpha_2=1$, M2: $\alpha_1=2$, $\alpha_2=1$ and M3: $\alpha_1=4$, $\alpha_2=1$. The obtained results with the proposed scheduler are compared with the reference PF schedulers – ordinary PF, power-aware PF described in (1) and MMPF scheduler from [13]. For the cases M1-M3, based on Figure 2, the new MPMPF scheduler achieves coverage gains in the order of 63-91% at the expense of only 15-22% throughput loss compared to ordinary PF in the SIMO scenario (sub-figures (a) and (b)). In the corresponding MIMO (sub-figures (c) and (d)) scenario, we obtain 71% constant coverage gain for the same throughput loss as in previous SIMO case. Compared to the MMPF scheduling principle [13] or to the power-aware PF in (1), similar coverage gains are obtained, as can be read from the figure.

Figure 3 illustrates the Jain's fairness index [14] per scheduling scheme for SIMO and MIMO scenarios, calculated using the truly realized throughputs at each TTI for all 10 UE's and over all the simulation runs. The value on the x axis corresponds to the used scheduler type (1 refers to ordinary PF scheduler, 2 refers to PPF, etc.). Clearly, the proposed MPMPF scheduler outperforms all other scheduling algorithms and the received fairness gains are in range of 17%-20% in the SIMO case and 35%-37% in the MIMO case.

Further illustrations on increased fairness are presented in Figure 4 (a) and (b) in terms of cumulative distribution of scheduled UEs per TTI for the simulated cases. Clearly, the power-aware term in MPMPF scheduling decision has a major impact on the increases fairness distribution seen as reaching the maximum number of UEs multiplexed per TTI.

V. Conclusion

In this paper, a new power-aware multi-stream proportional fair scheduling metric covering time-, frequency- and spatial domains was proposed. The proposed metric takes into account the transmit power allocation on RB basis, the instantaneous channel qualities (CQI's) as well as resource allocation fairness. The potential of advanced power-aware multi-user packet scheduling algorithms was evaluated in OFDMA type radio system context, using UTRAN LTE downlink as a practical example. The achievable throughput performance together with coverage and fairness distributions were analyzed and compared against the corresponding statistics of more traditional earlier-reported proportional fair scheduling techniques with SDM functionality in SIMO and MIMO cases. The proposed scheduling scheme offers better control over the ratio between the achievable cell/UE throughput and coverage increase, as well as increased UE fairness taking into account irregular BS transmission power.

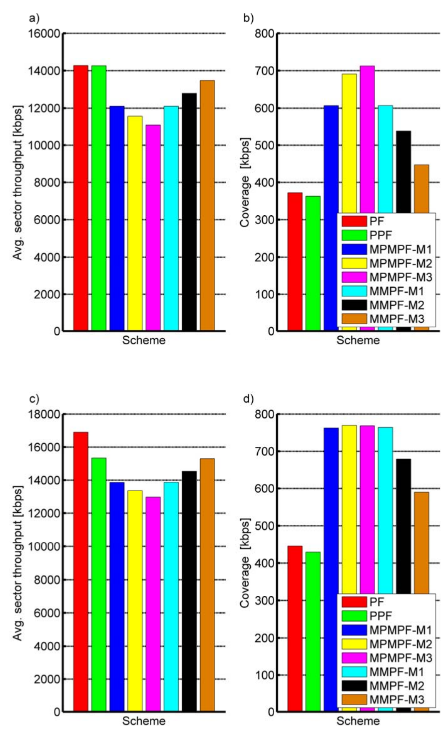


Figure 2. Average cell throughput and coverage gain over the reference PF scheduling scheme for the different simulation cases – SIMO (a, b) and MIMO (c, d).

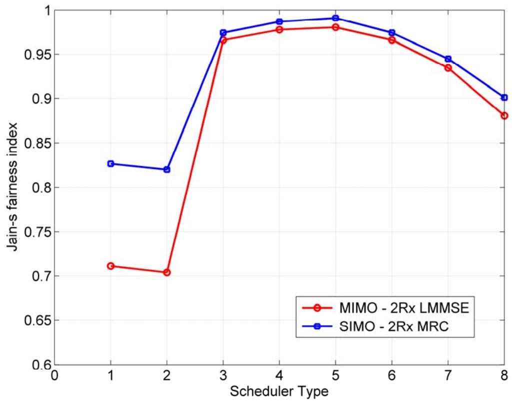


Figure 3. Jain's fairness index per scheduling scheme for SIMO and MIMO cases. Scheduler type 1 means ordinary PF, 2 means power-aware PF from (1), 3-5 mean proposed modified power-aware PF from (2) and 6-8 mean modified PF from [13].

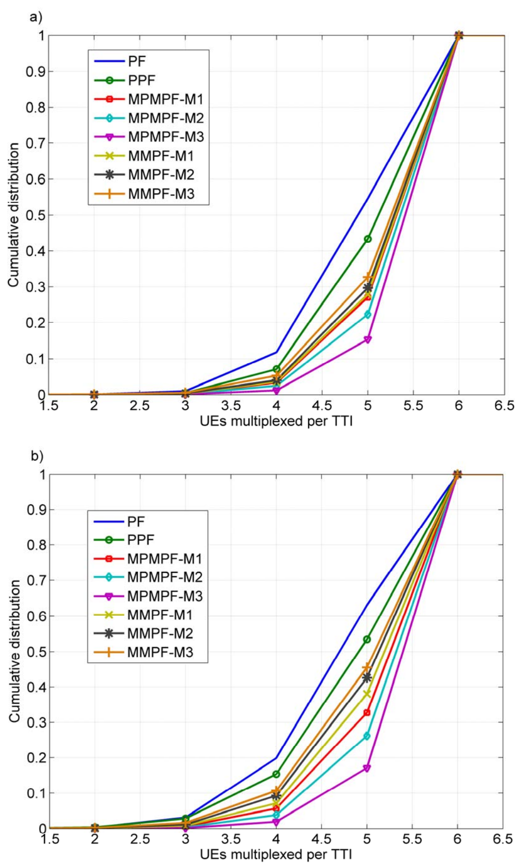


Figure 4. CDF of the number of UEs multiplexed in frequency domain per TTI for (a) SIMO and (b) MIMO cases.

References

[1] "Physical Layer Aspects for Evolved UTRA", *3GPP Technical Report TR 25.814*, ver. 7.1.0, Oct. 2006.

[2] WiMAX Forum™, http://www.wimaxforum.org, Aug. 2006.

[3] ITU global standard for international mobile telecommunications "IMT-Advanced", IMT-ADV/1-E,Mar. 2008.

[4] "Wireless World Initiative New Radio (WINNER)," IST-507581, European Information Society Technologies, Technical Report, Dec. 2006.

[5] N. D. Tripathi et al., *Radio Resource Management in Cellular Systems*, Springer, 2001.

[6] E. Dahlman et al., *3G Evolution: HSPA and LTE for Mobile Broadband*, Academic Press, 2007.

[7] Y. Sun et al., "Multi-user scheduling for OFDMA downlink with limited feedback for evolved UTRA", in *Proc. IEEE VTC-2006,* Montreal, Canada, September 2006.

[8] N. Wei, et al., "Performance of MIMO with frequency domain packet scheduling", in *Proc. IEEE VTC-2007*, Dublin, Ireland, May 2007.

[9] C. Wengerter, J. Ohlhorst, and A.G.E Von Elbwert, "Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA", in *Proc. IEEE VTC-2005,* Stockholm, Sweden, May 2005.

[10] P. Svedman, S.K.Wilson, L.J. Cimini, and B. Ottersten, "A simplified opportunistic feedback and scheduling scheme for OFDMA", in *Proc. IEEE VTC-2004*, Los Angeles, USA, pp. 1878-1882, May 2004.

[11] S. Y. Kyoung et al., "Power efficient frequency domain packet scheduling for OFDMA systems", in *Proc. Int. Conf. Convergence Information Technology, 2007,* Dhaka, Bangladesh, pp. 1906-1911.

[12] S. Nonchev, J. Venäläinen and M. Valkama, "New frequency domain packet scheduling schemes for UTRAN LTE downlink", in *Proc. ICT Mobile Summit*, Stockholm, Sweden, June 2008.

[13] S. Nonchev and M. Valkama, "Efficient packet scheduling schemes for multiantenna packet radio downlink", in *Proc. Fifth Advanced Int. Conf. Telecommunications (AICT-2009)*, Venice, Italy, May 2009.

[14] D. Chui and R. Jain, "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks", *Computer Networks and ISDN Systems*, 1989.

[15] K.I. Pedersen et al., "Frequency domain scheduling for OFDMA with limited and noisy channel feed-back," in *Proc. IEEE VTC-2007*, Baltimore, MD, Sept. 2007, pp. 1792-1796.

[16] A. Pokhariyal, T. E. Kolding, and P. E. Mogensen, "Performance of downlink frequency domain packet scheduling for the UTRAN long term evolution," in *Proc. IEEE PIMRC-2006*, Helsinki, Finland, Sept. 2006.

Publication 4

S. Nonchev and M. Valkama, “Advanced packet scheduling in soft frequency reuse scenarios for multiantenna packet radio systems,” in *Proceedings of Future Network and MobileSummit 2010,* pp. 1-8, Florence, Italy, June 2010.

Advanced Packet Scheduling in Soft Frequency Reuse Scenarios for Multiantenna Packet Radio Systems

Stanislav NONCHEV, Mikko VALKAMA
Tampere University of Technology, Department of Communications Engineering,
P.O.Box 553, Tampere, FIN-33101, Finland
Email: stanislav.nonchev@tut.fi, mikko.e.valkama@tut.fi

Abstract: The radio resource management functionality plays an important role in new OFDMA based networks. The control of the network resource division among the users is performed by packet scheduling functionality based on maximizing cell coverage and capacity while satisfying certain quality of service constrains. In order to mitigate inter-cell and co-channel interference problems in OFDMA cellular networks, soft frequency reuse with different power mask patterns is used. Stemming from our previous studies in scheduling algorithm developments, we extend the usability of packet scheduling techniques with built in transmit power considerations in the overall priority metrics calculations and scheduling decisions. In this paper, we investigate the system performance of combined soft frequency reuse schemes with advanced power-aware packet scheduling algorithms for further optimization between user fairness and cell throughput. As a concrete example, we demonstrate that by using such technique significant coverage improvements in the order of 30% can be obtained at the expense of only 14% throughput loss. Furthermore, the user fairness is also greatly increased, by more than 18%, when measured using Jain's fairness index.

Keywords: radio resource management, packet scheduling, soft frequency reuse, proportional-fair, power masks, channel quality feedback, fairness

1. Introduction

The new radio technologies for beyond 3G cellular systems such as 3GPP (3rd Generation Partnership Project) LTE (Long Term Evolution) [1] and WiMAX [2] are based on OFDMA air interface. The operating bandwidths of at least 10-20 MHz are divided into large number of orthogonal subcarriers for efficient reduction of the effects of inter-symbol interference (ISI) and inter-carrier interference (ICI). Consequently OFDMA is providing high spectral efficiency, flexible multi-user access and dynamic resource allocation.

On the other hand, important issue in those deployments is co-channel interference (CCI) that users at the cell borders are experiencing based on low signal to noise ratio (SNR) and high power emits from neighboring cells' base station in their communication channel. Solution for this problem in OFDMA cellular networks is utilization of frequency reuse schemes. Recent literature studies indicate that soft frequency reuse scheme has a capacity gain over the other hard frequency reuse schemes [3]. The overall functionality of frequency reuse schemes is based on applying specific power masks (fraction of the maximum transmission power level) over the whole system bandwidth. In soft frequency reuse (SFR) [4],[5] the frequency band, shared by all base stations (BS) (reuse factor is equal to 1) is divided into sub-bands with predefined power levels as illustrated in Figure1.

Consequently, the near and far from the BS users are allocated with different powers by limiting the impact of neighboring cells.

In general, all user equipments (UEs) within base station coverage are sharing the available radio resource (radio spectrum). This is controlled by the packet scheduler (PS) utilizing selected scheduling metrics. In both uplink (UL) and downlink (DL), the packet scheduler works in a centralized manner being located at the BS. It is generally fairly well understood that the overall radio system performance, in terms of throughput and fairness, depends heavily on the PS functionality. The PS operation is built on instantaneous radio channel conditions, quality of service (QoS) requirements and traffic situations of the served UEs [6], [7]. Finally, an additional important aspect in scheduling functionality is the interaction with other RRM entities, namely fast link adaptation (LA) and reliable re-transmission mechanisms.

Recently, multi-antenna oriented packet scheduling principles have been started to be investigated in the literature and their performance is evaluated in different simulator environments [8]-[13]. Stemming now from our previous work in advanced PS algorithms reported in [10]-[12], we extend the studies here to incorporate cell power bounds considerations in the scheduling decisions based on different SFR power mask configurations. Specifically in [12], the proposed scheduling schemes are investigated under the assumptions of PRB based power fluctuations due to instability of the network's link budget. In this paper, we introduce a combined method of SFR with advanced scheduling scheme based on proportional fair (PF) principle and investigate its performance using multi-user multi-antenna receive-transmit scheme in 3GPP long term evolution (LTE) context.

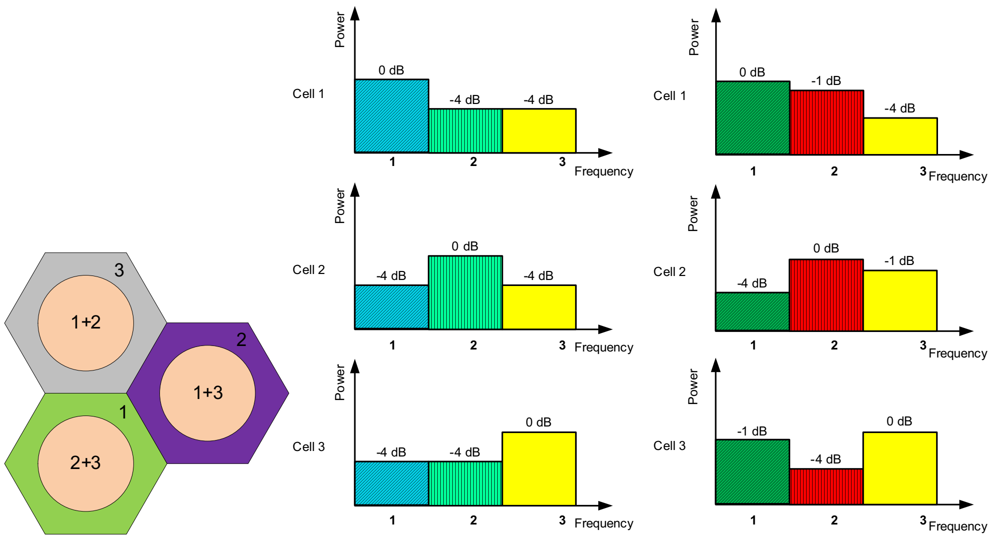


Figure 1: Example Soft Frequency reuse scheme with 3 sub-bands division.

The scheduling decisions are based on received users' signaling information in terms of acknowledgements and channel state information (CQI reports) per given transmission time interval (TTI) and per frequency domain physical resource block (PRB) [5]. In generall, in addition to throughput and capacity, another important feature of scheduling is fairness, implying that also users with less favorable channel conditions should anyway give some reasonable access to the radio spectrum. This is especially important in serving users at, e.g., cell edges in cellular networks.

The rest of the paper is organized as follows: Section 2 describes the SFR scheme and power mask configurations, while Section 3 is dedicated to power aware scheduling principles. Section 4 gives an overview of the overall system model and simulation assumptions. The simulation results and analysis are presented in Section 5, while the conclusions are drawn in Section 6.

2. Soft frequency reuse scheme

In general, SFR scheme reserves part of the frequency band for the cell-edge users and uses the power bound specified for it by the power mask. The rest of the unallocated sub-bands are dedicated to the near to BS users. One practical solution for sub-band division reported in the literature studies is 3 and therefore we are considering it in our case. Generally there is no restriction on the soft reuse factor [3],[4]. The same apply for the choice of the power mask. We have chosen for our evaluation the following power mask configurations: PM1 (0dB, -4dB, -4dB) and PM2 (0dB, -1dB, -4dB). The values in the brackets represent normalized transmission power values in dB as shown in Figure 1. If the sub-band allocated to the cell-edge UEs is not fully occupied, it can be still used by the other UEs.

3. Packet Scheduler

Packet scheduler is located in the BS and as a part of the whole RRM process it has an important role. Together with other RRM units such as *link adaptation* (LA) and *Hybrid ARQ* (HARQ) manager PS selects the most suitable users to access the radio resources based on some selected priority metric calculations. Moreover, we assume here that the LA unit can feed the PS with power allocation information on a PRB-per-PRB basis. MIMO functionality, in turn, requires both single-stream and dual-stream CQI feedback by each UE.

The advanced scheduler developments are stemming from the widely used two-stage (see e.g. [10], [11]) multi-stream PF approach. In the first stage time-domain (TD), within each TTI, UEs are ranked based on the full bandwidth channel state information and throughput calculations. Considering MIMO case, different spatial multiplexing possibilities (one-stream, dual-stream SU, dual-stream MU) are taken into account, in calculating all the possible reference throughputs. In the second stage, the scheduling functionality is expanded in frequency-domain (FD) and spatial-domain (SD) where the actual PRB allocation takes place. Initially, the needed PRB's for pending re-transmissions (on one stream-basis only) are reserved and the rest available PRB's are allocated to the selected UE's from the first stage. The actual priority metric in FD/SD stage is evaluated at PRB-level taking into account the available stream-wise channel state information, the transmit power allocations and the corresponding throughput calculations. The exact scheduling metrics are described below.

The modified power-aware multi-stream PF (MPMPF) metric proposed in [12] is calculated as:

$$\overline{\gamma}_{i,k,s} = arg\,\underset{i}{max}\left\{\frac{P_{k,s}}{P_{\max}}\left(\frac{CQI_{i,k,s}(n)}{CQI_i^{avg}(n)}\right)^{\alpha_1}\left(\frac{T_i(n)}{T_{tot}(n)}\right)^{-\alpha_2}\right\} \tag{1}$$

In the above metric, α_1 and α_2 are scheduler optimization parameters ranging basically from 0 to infinity. In (1), $P_{k,s}$, in turn, is the transmission power for sub-band k on stream s and P_{max} is the maximum transmission power for any sub-band. $T_{tot}(n)$ is the average delivered throughput (during the recent past) to all users ranked in TD stage served by the BS. $T_i(n)$ is the corresponding average delivered throughput to the UE i during the recent past. $CQI_{i,k,s}$ is the CQI of user i at PRB k and stream s, and $CQI^{avg}{}_i$ is the average CQI of user i calculated as:

$$CQI_i^{avg}(n)=\left(1-\frac{1}{t_c}\right)CQI_i^{avg}(n-1)+\frac{1}{t_c}\frac{1}{K_{TOT}}\frac{1}{S_{TOT}}\sum_{s=1}^{S_{TOT}}\sum_{k=1}^{K_{TOT}}CQI_{i,k,s}(n) \tag{2}$$

In above, K_{TOT} is the total number of available PRB's while S_{TOT} denotes the maximum number of streams which is here always two (max two streams).

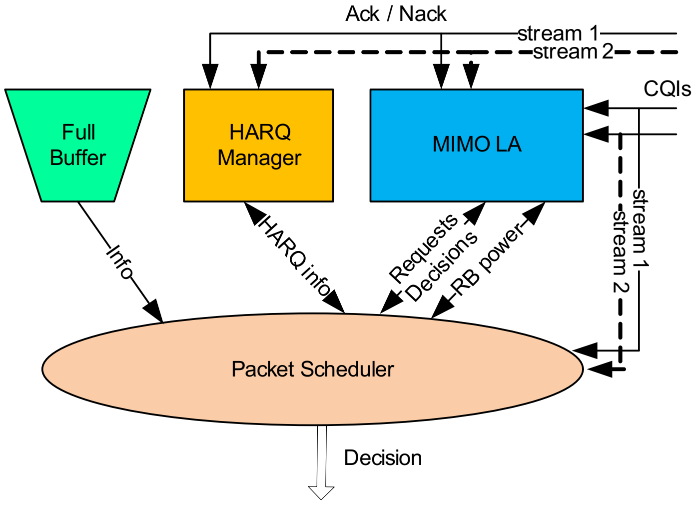


Figure 2: RRM functionality-scheduling process.

The scheduling metric in (1) is essentially composed of three parts affecting the overall scheduling decisions. The first ratio takes into account the transmit power levels in BS for each PRB due to SFR scheme. The second ratio is the relative instantaneous quality of the individual user's radio channels over their own average channel qualities in a stream-wise manner. The third ratio is related to measuring the achievable throughput of individual UE's against the corresponding average throughput of scheduled users. The power coefficients $\alpha 1$ and $\alpha 2$ are additional adjustable parameters that can be tuned and controlled to obtain a desired balance between throughput and fairness. This will be illustrated in Section 5.

4. System Simulation Methodology

Extensive quasi-static system simulator for LTE downlink providing traffic modelling, multiuser packet scheduling and link adaptation including HARQ is used for evaluating the system level performance of the proposed packet scheduling scheme, following 3GPP evaluation criteria [1]. The 10 MHz system bandwidth case is assumed, being composed of 1024 sub-carriers (out of which 600 are active) and divided into 50 physical resource blocks (PRB) each consisting of 12 sub-carriers with sub-carrier spacing of 15 kHz. Pilot signals are sent from base station to mobile station to determine the instantaneous channel condition. The mobile stations measure the actual channel states and the information is reported to the BS. The actual reported CQI's are based on received signal-to-interference-and-noise ratios (SINR), calculated by the UE's for each PRB. Here the UE's are assumed to use linear MMSE (LMMSE) receivers for MIMO 2x2 system simulation. Additionally, the UEs always report single-stream SINR as well as both single user (SU) and multi-user (MU) dual-stream SINR's at the corresponding detector output.

In a single simulation run, mobile stations are randomly distributed over standard hexagonal cellular layout with altogether 19 cells each having 3 sectors. As a concrete example, the number of active users in the cell is set to 15 and the UE velocities equal 3km/h. The path losses for individual links are directly determined based on the individual distances between the mobile and the serving base station. On the other hand, the actual fading characteristics of the radio channels are collected each TTI (1ms) and depend on the assumed mobility and power delay profile. Due to the centralized approach statistics are collected only from the central cell site while the others simply act as sources of inter-cell interference.

Table 1: Default simulation parameters

Parameter	Assumption
Cellular Layout	Hexagonal grid, 19 cell sites, 3 sectors per site
Inter-site distance	500 m
Carrier Frequency / Bandwidth	2000MHz / 10 MHz
Channel estimation	Ideal
PDP	ITU Typical Urban 20 paths
Minimum distance between UE and cell	>= 35 meters
Average number of UEs per cell	15
Max. number of frequency multiplexed UEs	10
UE receiver	LMMSE
Shadowing standard deviation	8 dB
UE speed	3km/h
Total BS TX power (*Ptotal*)	46dBm - 10MHz carrier
Traffic model	Full Buffer
Fast Fading Model	Jakes Spectrum
CQI reporting time	5 TTI
CQI delay	2 TTIs
MCS rates	QPSK (1/3, 1/2, 2/3),16QAM(1/2, 2/3, 4/5),64QAM(1/2, 2/3, 4/5)
ACK/NACK delay	2ms
Number of SAW channels	6
Maximum number of retransmissions	3
HARQ model	Ideal CC
1st transmission BLER target	20%
Forgetting factor	0.002
Scheduling schemes	PF, MPMPF

The used MIMO scheme for performance evaluation purposes is per-antenna rate control (PARC) with two transmit antennas at the BS and two receive antennas at each UE The main simulation parameters and assumptions are summarized in Table 1.

The RRM functionalities are controlled by the packet scheduler together with link adaptation and HARQ entities. Link adaptation consists of two separate elements – the inner loop LA (ILLA) and the outer loop LA (OLLA). These are used for removing CQI imperfections, estimating supported data rates and MCS’s, and stabilizing the 1st transmission Block Error Probability (BLEP) to the target range (typically 10-20%). Simple admission control scheme is used for keeping the number of UEs per cell constant. HARQ is based on SAW protocol and a maximum number of three re-transmissions is allowed. MIMO functionality requires individual HARQ entry per stream which is also implemented. Link-to-system level mapping is based on the effective SINR mapping (EESM) principle [1].

SFR scheme benefits from full bandwidth utilization for each BS and sub-band division helps in mitigating the CCI. The reuse factor used in the simulation scenarios is 3, which corresponds to sub-band division of (17, 17, 16) RB. Two different power patterns i.e. SFR power masks (section 2) define the relative power levels used for each RB group. In the first case the corresponding power pattern is (0dB, -4dB, -4dB) and in second case (0dB, -1dB, -4dB).

5. Numerical Results

In this section, we present the results obtained from the system simulations using the combined approach of SFR and PS algorithms described in the paper. The system-level performance is generally measured and evaluated in terms of:

- Throughput - the total number of successfully delivered bits per unit time. Usually measured either in kbps or Mbps.
- Coverage – the experienced data rate per UE at the 95% coverage probability.
- Fairness per scheduling scheme measured using Jain's fairness index [14]

We illustrate the potential of such technique using different SFR power masks and tuning the MPMPF scheduler with power coefficients α1 and α2. To emphasize the role of power-aware and CQI based priority metric calculation in (2), we fix the value of α2 to 1 and change the values of $\alpha 1$ as $\alpha 1 = \{1,2,4\}$ [11], [12].

Figure 3 illustrates the average user throughput and coverage for the different scheduling approaches and different power mask cases. The power coefficient values are presented as index M, where M1 represents the first couple, i.e., α_1=1, α_2=1, M2: α_1=2, α_2=1 and M3: α_1=4, α_2=1. By combining SFR and PF scheduling we achieve small throughput gain, while coverage is increased with 10%. Changing the scheduling method to MPMPF and coefficient values (M1), we achieve coverage gains in the order of 58% at the expense of 20% throughput loss for PM1case scenario presented in (a) and (b). In PM2 case scenario illustrated on (c) and (d) we obtain nearly the same coverage gain (60%) for small throughput loss (13%).

Continuing the evaluation of the proposed method – SFR combined with MPMPF, we clearly see a trade-off between average cell throughput and coverage for different power coefficient values in PM1 and PM2 cases. Furthermore, the remaining power coefficient values are used for tuning the overall scheduling performance. For the rest of the cases, the cell throughput loss is decreased stepwise with around 4% per index M, which on the other hand corresponds to coverage gains increase from 74% to 80%. Consequently, an obvious trade-off between average cell throughput and coverage is clearly seen. Similarly, in PM2 case scenario we obtain coverage gains between 75% to 82% corresponding to throughput losses from 10% to 8% as illustrated by the performance statistics. Clearly, combining SFR with MPMPF scheduler reduces the overall throughput loss by 5-10% yet still achieving the same coverage gains compared to [12].

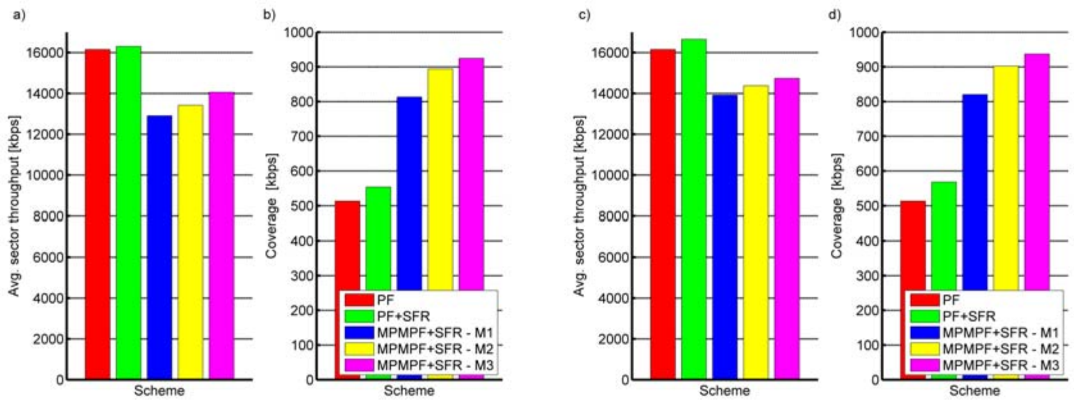


Figure 3: Average cell throughput and coverage gain over the reference PF scheduling scheme and PF +SFR scheme for the different power masks scenarios – PM1 (a, b) and PM2 (c, d). The schemes M1-M3 refer to the MPMPF scheduler with power different coefficient values.

Figure 4 illustrates the Jain's fairness index per scheduling scheme for different SFR power masks scenarios, calculated over all the I_{TOT} = 15 UE's. The value on the x axis corresponds to used scheduler type, where 1 refers to the reference PF scheduler, 2 refers to PF+SFR, 3 refers to the MPMPF+SFR with index M1, etc. The value of Jain's fairness index is generally in the range of [0,1], where value of 1 corresponds to all users having the same amount of resources. Clearly, the fairness distribution with MPMPF+SFR outperforms the used reference plain PF scheduler and PF+SFR for both analyzed cases. Compared to our earlier work in [12] the gain over MPMPF is in the range of 1-3%. The received fairness gains are in range of 17%-31% in the PM1 case and 18%-32% in the PM2 case. Compared to the PF+SFR case the corresponding gains are in the range of 15-17% for both cases.

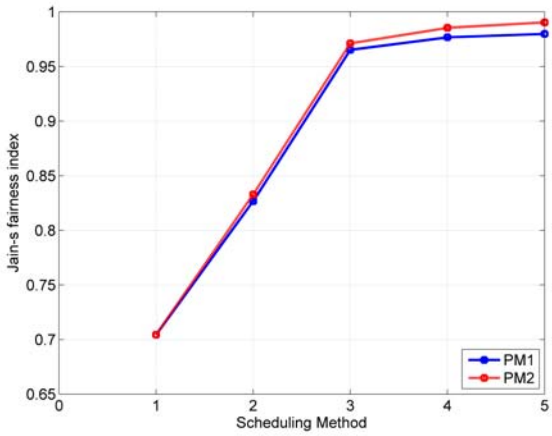


Figure 4: Jain's fairness index per scheduling scheme. Scheduler type 1 means ordinary PF, while 2 means PF +SFR, 3-5 means MPMPF +SFR with power coefficients M1,M2 and M3 correspondingly.

6. Conclusions

In this paper, we have studied the potential of combining soft frequency reuse schemes with advanced power-aware multi-user packet scheduling algorithms in OFDMA type radio system context, using UTRAN long term evolution (LTE) downlink in Macro cell environment. The MPMPF scheduler metric covering time-, frequency- and spatial domains was used that takes into account the corresponding power levels on the RB basis defined by SFR power masks, the instantaneous channel qualities (CQI's) as well as resource allocation fairness. Overall, the achieved throughput performance together with coverage and fairness statistics were assessed, by using extensive system simulations, and compared against plain proportional fair scheduling and combined with SFR functionality. In the case of fixed coverage requirements, the proposed scheduling metric calculations based on UE channel feedback and SFR power mask offers better control over the ratio between the achievable cell/UE throughput and coverage increase, as well as increased UE fairness. As a practical example, the fairness in resource allocation together with cell coverage can be increased significantly (more than 50%) by allowing a small decrease (in the order of only 10-15%) in the cell throughput for plain PF scheduling case and more than 15% increase over the traditional PF combined with SFR. Compared to MPMPF scheme the obtained performance results are in the same range with additional decrease of 5-7% in total throughput loss.

References

[1] "Physical Layer Aspects for Evolved UTRA", 3GPP Technical Report TR 25.814, ver. 7.1.0, Oct. 2006.

[2] WiMAX Forum™, http://www.wimaxforum.org, Aug. 2006.

[3] Bohge Matias, et. al., "Optimal Power Masking in Soft Frequency Reuse based OFDMA Networks," in Proc. of the European Wireless Conference 2009 (EW'09), Aalborg, Denmark,May 2009

[4] 3GPP; Huawei, "Soft frequency reuse scheme for UTRAN LTE," R1- 050507, May 2005.

[5] K. Doppler, X. He, C. Witjing, and A. Sorri, "Adaptive soft reuse for relay enhanced cells," in Proc. of the 65th IEEE Vehicular Technology Conference (VTC-Spring 07), Apr. 2007, pp. 758–762.

[6] N. D. Tripathi et al., Radio Resource Management in Cellular Systems, Springer, 2001.

[7] E. Dahlman et al., 3G Evolution: HSPA and LTE for Mobile Broadband, Academic Press, 2007.

[8] N. Wei, et al., "Performance of MIMO with frequency domain packet scheduling", in Proc. IEEE VTC-2007, Dublin, Ireland, May 2007.

[9] Q. Wang, J. Xu, and Z. Bu, "Proportional-fair bit and power adaptation in multi-user OFDM systems," in Proc. of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Helsinki, Finland, Sep. 2006, pp. 1–4.

[10] S. Nonchev and M. Valkama, "Efficient packet scheduling schemes for multiantenna packet radio downlink", in Proc. Fifth Advanced Int. Conf. Telecommunications (AICT-2009), Venice, Italy, May 2009.

[11] S. Nonchev and M. Valkama, "A new fairness-oriented packet scheduling scheme with reduced channel feedback for OFDMA packet radio systems", Int. Journal of Communications, Network and System Sciences (IJCNS), 2009.

[12] S. Nonchev and M. Valkama, "Efficient power-aware packet scheduling schemes for multiantenna packet radio systems", in The First International Workshop on Wireless & Mobile Networks (WiMoNe - 2009), December 2009, Chennai (Madras), India

[13] S. Y. Kyoung et al., "Power efficient frequency domain packet scheduling for OFDMA systems", in Proc. Int. Conf. Convergence Information Technology, 2007, Dhaka, Bangladesh, pp. 1906-1911

[14] D. Chui and R. Jain, "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks", Computer Networks and ISDN Systems, 1989.

Publication 5

S. Nonchev, M. Valkama, R. Hamila and M. Hasna "On the performance of advanced QoS-aware packet scheduling for multiantenna packet radio systems" in *Proceedings of 3rd IEEE International Conference on Broadband Network and Multimedia Technology 2010* (IC-BNMT-2010), pp. 209-304, Beijing, China, October 2010.

ON THE PERFORMANCE OF ADVANCED QOS-AWARE PACKET SCHEDULING FOR MULTIANTENNA PACKET RADIO SYSTEMS

Stanislav Nonchev[1], Mikko Valkama[1], Ridha Hamila[2], Mazen Hasna[2]

1 Department of Communications Engineering, Tampere University of Technology, Tampere, Finland
2 College of Engineering, Qatar University, Doha, Qatar
stanislav.nonchev@tut.fi, mikko.e.valkama@tut.fi, hamila@qu.edu.qa, hasna@qu.edu.qa

Abstract

In this paper, we study and analyze the performance of advanced OFDM packet scheduling techniques with given QoS requirements under the 3GPP UTRAN Long Term Evolution (LTE) framework. Stemming from our previous studies in scheduling algorithm developments, we show that the main radio resource management (RRM) performance measures such as fairness, throughput and coverage are efficiently controlled by proper scheduling priority metric calculations utilizing time-, frequency- and spatial domains. The proposed scheduling algorithms are able to meet the higher layer quality of service (QoS) targets, taking into account Constant Bit Rate (CBR) traffic scenarios with Guaranteed Bit Rate (GBR) constrains as well as instantaneous user channel conditions. Moreover, we demonstrate that by using such algorithms, significant coverage improvements in the order of 40% can be obtained at the expense of 8% total throughput loss compared to more traditional scheduling approaches. Furthermore, the user fairness is also greatly increased, by more than 15% when measured using Jain's fairness index.

Keywords: cellular system performance; radio resource management; packet scheduling; QoS; proportional-fair; channel quality feedback; fairness

1 Introduction

Following the latest trends in beyond 3G cellular network developments, the OFDMA air interface is the practical choice for achieving increased data rates, improved mobility and flexible spectrum use. Relevant work in this direction includes, e.g., latest releases of 3GPP Long Term Evolution (LTE) [1, 2], IMT-Advanced [3], WiMAX [4], etc. Reaching the performance requirements set in these developments call for sophisticated optimization of, e.g., base stations' (BS) radio resource management (RRM) functionalities. Moreover, scalable bandwidth operation and MIMO technologies are the major driving forces behind these developments to achieve these primary goals. MIMO in terms of *Spatial Multiplexing* (SM), possibly combined with pre-coding, is considered as one core physical layer technology towards increased link spectral efficiencies compared to existing radio systems. In addition, it also provides the packet scheduler (PS) with an extra degree of freedom (spatial domain), by offering a possibility to multiplex multiple data streams of one or more users on the same physical time-frequency resource. Thus the two principal concepts widely analyzed in literature [5] are single-user (SU) and multi-user (MU) MIMO. SU-MIMO allows single UE scheduling at the same time-frequency resource block (RB), while MU-MIMO provides flexibility so that multiple users can be scheduled on separate parallel streams over the same time-frequency RB.

In general, QoS means traffic differentiation and configuring multiple bearers and priorities to ensure satisfactory service quality for each user. Several LTE QoS parameters are defined in order to provide effective solution to differentiating packet services. Here we mainly focus on GBR, which is defined as the expected rate to be provided by the access network.

In cellular performance optimization, link- and cell-level performance improvements are basically obtained through proper deployment of fast link adaptation and advanced packet scheduling algorithms, combined with efficient retransmission schemes, exploiting the available multi-user diversity in both time and frequency as well as spatial domains [6, 7]. These are typically enabled by having some form of channel state information on the transmitter side, e.g. in terms of codeword containing both MIMO ranking information and channel quality indicator (CQI) measurement from each mobile station within the serving cell. In this paper, we study the performance of advanced QoS aware proportional fair (PF) scheduler schemes in combination with realistic traffic scenario - namely

Constant Bit Rate (CBR) traffic with Guaranteed Bit Rate (GBR) constrains in LTE context.

It is generally fairly well understood that the overall radio system performance in terms of throughput, coverage and fairness, depends heavily on PS functionality, being the key ingredient in the radio resource management (RRM) process. Recent literature studies [8, 9, 10] on different multi-user packet scheduling principles reveal the potential of OFDMA based systems. Moreover, most of them are based on well-known proportional fair (PF) scheduling principle and do not consider extensively the user's QoS requirements, practical system constrains and limitations. On the other hand, the QoS aware packet scheduling strategies investigated in the literature consider GBR scenarios with certain allowance of packet delays [11], while many of the advanced techniques are taking into account the varying channel state information [12, 13]. Stemming now from our previous work in advanced PS scheduling developments reported in [5, 6, 14], we extend our studies here to incorporate QoS requirements into scheduling decisions by effectively controlling user fairness and BS's RRM process. Furthermore, we apply different simulation cases with CBR traffic investigating the limits of achieved gains from time-frequency-spatial domain packet scheduling. The system model used for the performance evaluations of the proposed scheduling methods presented in this paper is according to the 3GPP evaluation criteria [2]. The overall outcome is measured in terms of average *cell throughput* and *coverage,* and *fairness* distribution.

The rest of the paper is organized as follows: Section 2 gives an overview of the packet scheduling process and describes the proposed QoS aware multi-stream PF (QoS-MSPF) scheduling scheme. Section 3 presents the overall system model and simulation assumptions. The simulation results and analysis are presented in Section 4, while the conclusions are drawn in Section 5.

2 Packet scheduling process

Packet scheduler is located in the BS and is a part of the overall RRM framework [4] together with e.g. LA and HARQ manager. Packet scheduler distributes the available cellular resources to different UEs based on selected priority or scheduling metric calculations. In more details, the scheduling decisions are typically based on received users' signaling information in terms of acknowledgements (ACK/NACK) and channel state information (CQI reports) per given transmission time interval (TTI) and per frequency domain physical resource block (PRB). Spatial domain functionality requires additional UEs' ranking information for possible MIMO modes extracted also from the reported codeword. Based on this information, and assuming two transmit antenna downlink case with dual-antenna UE receivers for practicality, the BS then decides whether the particular time-frequency resource is used for (i) transmitting only one stream to a specific UE, (ii) two streams for a specific UE (SU-MIMO) or (iii) two streams (one plus one) to two different UEs (MU-MIMO).

The proposed QoS-aware multi-stream PF (QoS-MSPF) scheduler is presented at conceptual level in Figure 1. It is based on widely used two-stage multi-stream PF approach with additional QoS enabled services and MIMO functionality [6, 11, 14]. In the first stage, which is the time-domain (TD) scheduling step executed within each TTI, UEs are ranked based on the full bandwidth channel state information, QoS parameters and the corresponding throughput calculations. Considering MIMO case, different spatial multiplexing possibilities (one-stream, dual-stream SU, dual-stream MU) are taken into account, in calculating all the possible reference throughputs. In the second stage, the scheduling functionality is expanded in frequency-domain (FD) and spatial-domain (SD) where the actual PRB allocation takes place. Initially, the schedulable UEs are determined through available buffer information. Consequently, the needed PRB's for pending re-transmissions (on one stream-basis only) signaled through HARQ channels are reserved and the rest available PRB's are allocated to the selected UE's from the first stage. The actual priority metric in FD/SD stage is evaluated at PRB-level taking into account the available stream-wise channel state information, QoS parameters and the corresponding throughput calculations.

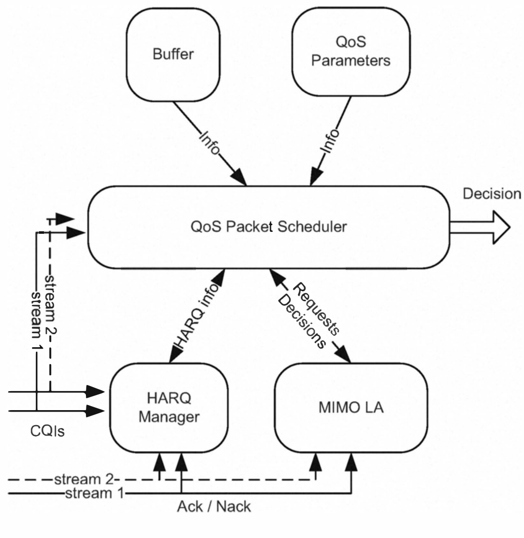


Figure 1. RRM entities and packet scheduling process.

In terms of the actual metric calculations, the proposed QoS-aware multi–stream PF (QoS-MSPF) scheduler uses the following metric

$$\overline{\gamma}_{i,k,s} = arg\max_{i}\left\{\eta_i(n)\left(\frac{CQI_{i,k,s}(n)}{CQI_i^{avg}(n)}\right)^{\alpha_1}\times\right.$$
$$\left.\left(\frac{T_i^{se}(n)}{T_{i,k,s}(n)}\frac{T_i(n)}{T_{tot}(n)}\right)^{-\alpha_2}\right\} \quad (1)$$

which can be understood as an extension of the authors' earlier work in [7]. In the above metric, α_1 and α_2 are scheduler optimization parameters ranging basically from 0 to infinity. In expression (1), $T_i^{se}(n)$ is an estimate of the user throughput if user i is scheduled on sub-frame basis according to [12] and $T_{i,k,s}(n)$ is the estimated achievable throughput of user i at PRB k and stream s. $T_i(n)$ corresponds to average delivered throughput to the UE over the past and T_{tot} is the average delivered throughput (during the recent past) to all users ranked in TD stage served by the BS. $CQI_{i,k,s}$ is the CQI of user i at PRB k and stream s, and CQI^{avg}_i is the average CQI of user i calculated as:

$$CQI_i^{avg}(n) = \left(1-\frac{1}{t_c}\right)CQI_i^{avg}(n-1) + $$
$$+\frac{1}{t_c}\frac{1}{K_{TOT}}\frac{1}{S_{TOT}}\sum_{s=1}^{S_{TOT}}\sum_{k=1}^{K_{TOT}}CQI_{i,k,s}(n) \quad (2)$$

In above, K_{TOT} is the total number of available PRB's while S_{TOT} denotes the number of streams which is here always two (max two streams). The $\eta_i(n)$ is, in turn, a QoS specific factor defined as [12]

$$\eta_i(n) = 1 + \xi e^{[-\psi(T_i(n)-GBR_i)]} \quad (3)$$

The parameters ξ and ψ are used to control user traffic requirements and example values are proposed in [15]. GBR_i corresponds to the desired user throughput.

The scheduling metric in (1) is essentially composed of three parts affecting the overall scheduling decisions. The first term takes into account the QoS requirements. The second ratio is the relative instantaneous quality of the individual user's radio channels over their own average channel qualities in a stream-wise manner. The third ratio is divided into two parts. The first one takes into account the estimated throughputs of individual UE's and the second one the achievable over total throughputs. The power coefficients α_1 and α_2 are additional adjustable parameters that can be tuned and controlled to obtain a desired balance between throughput and fairness. This will be illustrated in Section 5.

In general, while the increased flexibility of combined QoS- and channel-aware scheduling can offer performance enhancements, compared to fixed resource allocation, it also has some practical disadvantages. This includes e.g. relatively higher scheduling complexity, in terms of scheduling metric calculations and increased signaling overhead. Keeping these at reasonable levels requires thus some constraints on the scheduling algorithm, so for simplicity we assume here that only one MIMO mode (SU or MU) and fixed modulation and coding scheme (MCS) is allowed per user within one scheduling element. Moreover, we limit the number of users for multiplexing in TD stage to further reduce the signaling overhead and complexity of FD/SD scheduling [11].

3 System simulation model

The performance of the proposed scheduling scheme is evaluated in extensive quasi-static system simulator for LTE downlink providing traffic modeling, multiuser packet scheduling and link adaptation including HARQ, following the 3GPP evaluation criteria [2]. The 10 MHz system bandwidth is divided into 50 physical resource blocks (PRB) each consisting of 12 sub-carriers with sub-carrier spacing of 15 kHz. The mobile stations measure the actual channel states based on pilot signals and report them to the BS together with MIMO ranking possibilities. The actual reported CQI's are based on received signal-to-interference-and-noise ratios (SINR) after receiver spatial signal processing, calculated by the UE's for each PRB. Here the UE's are assumed to use dual-antenna linear MMSE (LMMSE) receivers, and the reporting includes single-stream SINR as well as both single user (SU) and multi-user (MU) dual-stream SINR's.

In a single simulation run (Macro case 1 scenario [2]), mobile stations are randomly distributed over standard hexagonal cellular layout with altogether 19 cells each having 3 sectors. The number of active users per sector is set to 20 and the UE velocities equal 3km/h. The used MIMO scheme for performance evaluation purposes is per-antenna rate control (PARC) with two transmit antennas at the BS and two receive antennas at each UE. The main simulation parameters and assumptions are summarized in Table 1.

The RRM functionalities are controlled by the packet scheduler together with link adaptation and HARQ entities. Moreover, link adaptation functionality consist of removing CQI imperfections, estimating supported data rates and MCS's, and stabilizing the 1st transmission Block Error Probability (BLEP) to the target range (typically 10-20%). Poisson call arrival model and simple admission control scheme for keeping the number of UEs per cell constant are used in the performance evaluation simulations. HARQ is based on SAW protocol and a maximum of three re-transmissions is allowed. MIMO functionality requires individual HARQ entry per stream which is also implemented. Link-to-system level mapping is based on the effective SINR mapping (EESM) principle [2].

The CBR traffic model implies fixing the amount of individual UE data packets arriving at the BS with constant size and inter-arrival time (Table 1). The UE GBR requirement is equal to their constant source bit rate.

Table 1 Default simulation parameters

Parameter	Assumption
Carrier Frequency / Bandwidth	2000MHz / 10 MHz
Number of active sub-carriers	600
Sub-carrier spacing	15kHz
Sub-frame duration	0.5 ms
Channel estimation	Ideal
PDP	ITU Typical Urban 20 paths
Minimum distance between UE and cell	>= 35 meters – Macro
UE's per sector	20
Max. number of frequency multiplexed UEs	10
UE receiver type	LMMSE
Shadowing standard deviation	8 dB
UE speed	3km/h
Total BS TX power (Ptotal)	46dBm
Fast Fading Model	Jakes Spectrum
CQI reporting schemes	Full CQI
CQI log-normal error std.	1 dB
CQI reporting time	5 TTI
CQI delay CQI quantization CQI std error	2 TTIs 1 dB 1 dB
MCS rates	QPSK (1/3, 1/2, 2/3), 16QAM (1/2, 2/3, 4/5), 64QAM (1/2, 2/3, 4/5)
ACK/NACK delay	2ms
Number of SAW channels	6
Maximum number of retransmissions	3
HARQ model	Ideal chase combining (CC)
1st transmission BLER target	20%
Scheduler forgetting factor	0.0025
Scheduling schemes used	QoS- PF, QoS-PFsch [11], QoS -MSPF (proposed)
Traffic type	Constant Bit Rate (256 kbps)
Packet size	2048 bits
Inter packet arrival time	8ms
Packet transfer delay budget	80ms
Simulation duration (one drop)	5 seconds
Number of drops	10

4 Simulation results

In this section, we present the results obtained from the system simulations using the PS algorithms described in the paper. The system-level performance is generally measured and evaluated in terms of:

- Cell throughput distribution (Mbps)
- Cell-edge coverage (kpbs or Mbps)
- Fairness distribution
- CDF of the number of users scheduled per TTI
- Scheduling distribution over G-factor

In general, cell throughput is defined as the number of successfully delivered user bits per unit time. Coverage, in turn, corresponds to 5% probability point in the throughput CDF. Fairness is measured using the Jain's fairness index [16]. G-factor is defined as the ratio of total received wideband BS power and total interference power (noise and other cell interferences) at UE averaged over short term fading (excluding shadowing).

The performance of the proposed QoS-MSPF scheduler is evaluated using different power coefficients α_1 and α_2 and CBR traffic configurations. The role of the channel quality reporting in the QoS-aware priority metric calculation in equation (1), α_2 is fixed here to 1 and different values are then demonstrated for α_1 [14]. At the other extreme, large α_2 values increase the effect of the second term in priority metric calculation based on throughput estimation and the scheduling algorithm would behave like maximum throughput scheduler, which implies reduced fairness distribution. The used values for α_1 coefficient are defined as $\alpha_1 = \{1,2,4\}$. Moreover, the power coefficient values are presented as index M, where M1 represents the first couple, i.e., $\alpha_1=1$, $\alpha_2=1$, M2: $\alpha_1=2$, $\alpha_2=1$ and M3: $\alpha_1=4$, $\alpha_2=1$.

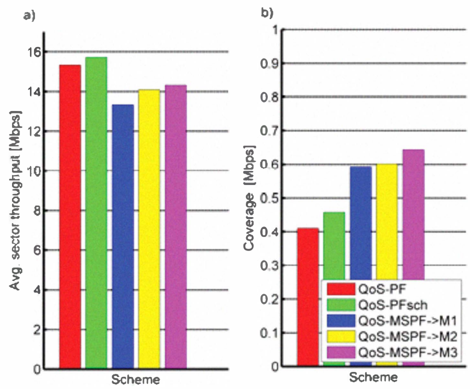


Figure 2. Average cell throughput and coverage gain over the reference PF scheduling schemes

Figure 2 illustrates the average cell throughput and cell-edge coverage for the different QoS-aware schedulers in MIMO (sub-figures (a) and (b)) system simulation case. The obtained results with the proposed scheduler are compared with the reference PF schedulers – ordinary QoS-PF and QoS-PFsch [11]. We observe that the overall system performance is reduced between 10-15% compared to our previous studies [6, 7] due to the applied CBR traffic model and the constant bit-rates demands from all the UEs in the serving BS. For the cases M1-M3, based on Figure 2, the new QoS-MMPF scheduler achieves coverage gains in the order of 44-57% at the expense of only 7-13% throughput loss compared to ordinary QoS-PF. Similarly, the corresponding coverage gains are 30-41% for the 9-15% throughput loss when compared to QoS-PFsch case. Compared to the MMPF scheduling principle [6], lower coverage gains are obtained, as can be expected for CBR traffic scenarios.

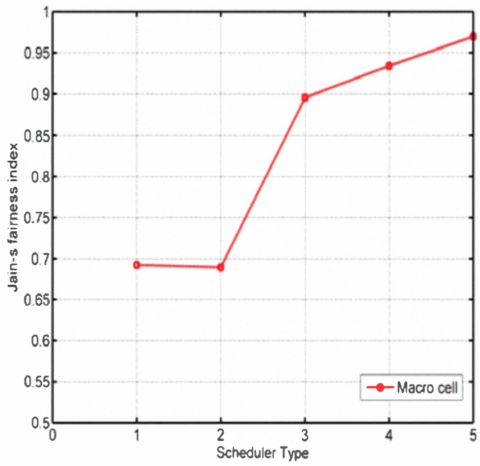


Figure 3 Jain's fairness index per scheduling scheme.

Figure 3 illustrates the Jain's fairness index [16] per scheduling scheme calculated using the truly realized throughputs at each TTI for all 20 UE's and over all the simulation runs. The value on the x-axis corresponds to the used scheduler type (1 refers to ordinary QoS-aware PF scheduler, 2 refers to QoS-PFsch, etc.). We observe that both reference scheduling schemes ordinary QoS-PF and the QoS-PFsch have identical fairness distributions and the proposed QoS-MMPF scheduler clearly outperforms them in terms of fairness. Having CBR traffic model with GBR requirements, the received fairness gains are in range of 29%-40% when compared to the reference scheduling schemes.

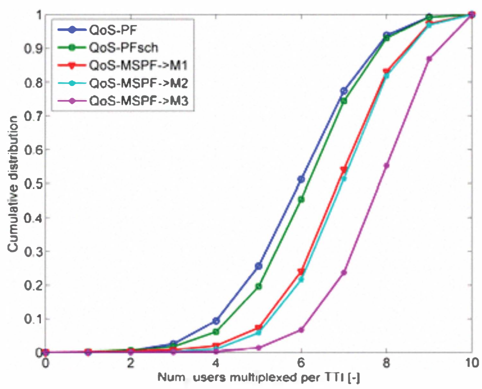


Figure 4. CDF of the number of UEs multiplexed in frequency domain per TTI

Further illustrations on increased fairness are presented in Figure 4 in terms of cumulative distribution of scheduled UEs per TTI for the simulated cases. Clearly, thanks to the metric decoupling as well as efficient usage of the provided feedback have a major impact on the increases fairness distribution seen as reaching the maximum number of UEs multiplexed per TTI.

Figure 5 illustrates the behavior of the scheduling schemes in terms of resource distribution over the G-factor. Clearly, the proposed QoS – MSPF scheduling scheme tends to allocate equal resources to all users. The reference QoS-PF and QoS-PFsch schemes are favoring users with better channel condition and this affect their fairness distribution as show in Figure 3. On the other hand, changing the power coefficient values of the QoS-MMPF scheduling metric increases the resources allocated to lower channel quality users and guarantees the GBR requirements.

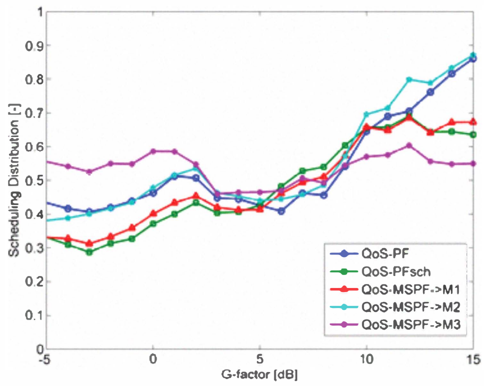


Figure 5. Scheduling distribution over G-factor.

5 Conclusions

In this paper, we have studied the potential of advanced QoS-aware multi-stream packet scheduling algorithms in OFDMA type radio system context, using UTRAN long term evolution (LTE) downlink in Macro cell environment as practical example case. We demonstrated the benefits of multi-domain (time-, frequency- and spatial) scheduling metric in achieving increased user fairness and coverage taking into account certain QoS requirements, instantaneous channel qualities (CQI's) as well as resource allocation fairness. Overall, the achieved throughput performance together with coverage and fairness statistics were assessed, by using extensive system simulations, and compared against more traditional scheduling schemes. In the case of specific traffic requirements, the proposed scheduling metric calculations based on UE channel feedback and QoS metric offers better control over the ratio between the achievable cell/UE throughput and coverage increase, as well as increased UE fairness. As a practical example, the fairness in resource allocation together with cell coverage can be increased significantly (more than 30%) by allowing a small decrease (in the order of only 7-15%) in the average cell throughput.

References

[1] 3GPP Technical Specification Group Radio Access Network Technical Specification 36.300 v8.0.0: Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (EUTRAN); Overall description, march 2007

[2] Physical Layer Aspects for Evolved UTRA", 3GPP Technical Report TR 25.814, ver. 7.1.0, Oct. 2006.

[3] ITU global standard for international mobile telecommunications "IMT-Advanced", IMT-ADV/1-E,Mar. 2008.

[4] WiMAX Forum™, http://www.wimaxforum.org, Aug. 2006.

[5] E. Dahlman et al., 3G Evolution: HSPA and LTE for Mobile Broadband, Academic Press, 2007.

[6] S. Nonchev and M. Valkama, "Efficient packet scheduling schemes for multiantenna packet radio downlink", in *Proc. Fifth Advanced Int. Conf. Telecommunications (AICT-2009)*, Venice, Italy, May 2009.

[7] S. Nonchev and M. Valkama, "Advanced radio resource management for multiantenna packet radio systems", *International Journal of Wireless & Mobile Networks (IJWMN),* 2010, vol. 2, No. 2, pp.1-14

[8] Y. Sun et al., "Multi-user scheduling for OFDMA downlink with limited feedback for evolved UTRA", *in Proc. IEEE VTC-2006*, Montreal, Canada, September 2006.

[9] N. Wei, et al., "Performance of MIMO with frequency domain packet scheduling", *in Proc. IEEE VTC-2007*, Dublin, Ireland, May 2007.

[10] C. Wengerter, J. Ohlhorst, and A.G.E Von Elbwert, "Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA", *in Proc. IEEE VTC-2005*, Stockholm, Sweden, May 2005.

[11] G. Monghal, et al., QoS Oriented Time and Frequency Domain Packet Schedulers for The UTRAN Long Term Evolution, *in Proc. VTC2008-Spring*, vol.3, pp. 2532-2536, May 2008.

[12] T.E. Kolding, QoS-Aware Proportional Fair Packet Scheduling with Required Activity Detection, *VTC2006-Fall*, September 2006

[13] T.-D. Nguyen, Y. Han, A Proportional Fairness Algorithm with QoS Provision in Downlink OFDMA Systems, *IEEE Communication Letters*, vol. 10, No. 11, November 2006.

[14] S. Nonchev and M. Valkama, "A new fairness-oriented packet scheduling scheme with reduced channel feedback for OFDMA packet radio systems", *Int. Journal of Communications, Network and System Sciences (IJCNS),* 2009, vol. 7, pp. 608-618.

[15] M. Lundevall, B. Olin, J. Olsson, N. Wiberg, S. Wanstedt, J. Eriksson, and F. Eng, "Streaming Applications over HSDPA in Mixed Service Scenarios," *in Proceedings of Vehicular Technology Conference (VTC)*, vol. 2, Los Angeles, California, USA, September 2004, pp. 841–845.

[16] D. Chui and R. Jain, "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks", *Computer Networks and ISDN Systems*, 1989.

Publication 6

S. Nonchev, M. Valkama and R. Hamila, “Effect of high-velocity scenarios on the performance of MIMO LTE packet scheduling,” in *Proc. 8th International Multi-Conference on Systems, Signals and Devices* (SSD-2011), pp. 1-6, Sousse, Tunisia, March 2011.

Effect of High-Velocity Scenarios on the Performance of MIMO LTE Packet Scheduling

Stanislav Nonchev[1], Mikko Valkama[1] and Ridha Hamila[2]
[1]Tampere University of Technology, Tampere, Finland
[2]College of Engineering, Qatar University ,Doha, Qatar
stanislav.nonchev@tut.fi, mikko.e.valkama@tut.fi, hamila@qu.edu.qa

Abstract—**In this paper we analyze the performance of advanced OFDM packet scheduling techniques based on 3GPP UTRAN Long Term Evolution (LTE) framework for high velocity scenarios. Based on our previous investigations in system performance analysis, we demonstrate that packet scheduler being part of the whole radio resource management (RRM) process has a major impact on it. On the other hand high user velocities further imply additional system performance limitations due to the introduction of reporting overhead. Therefore, we investigate the resulting fairness distribution among users together with the achievable radio system performance in terms of throughput, coverage and fairness distribution, by simulating practical OFDMA cellular system environment with MIMO functionality in Macro cell Best effort traffic scenario. As a concrete example, we demonstrate that by using the advanced fairness-oriented multi-user scheduling schemes, significant coverage improvements in the order of 30% can be obtained by achieving the same cell throughput. Furthermore, the user fairness is also greatly increased, by more than 25%, when measured using Jain's fairness index.**

Keywords-system performance; radio resource management; packet scheduling; proportional-fair; channel quality feedback; fairness

I. INTRODUCTION

Future network developments are aiming at increased data rates, improved mobility and certain QoS guarantees. Relevant work in this direction includes, e.g., latest 3GPP Long Term Evolution (LTE) [1, 2], IMT –Advanced [3], WiMAX [4] etc. Moreover, utilization of OFDMA air interface combined with MIMO technologies are the major driving forces behind these developments to achieve the primary goals, i.e., increased cell throughput and coverage, user fairness and QoS guarantees [5]. Performance improvements are basically obtained through proper radio resource management (RRM) functionality at the BS side. Thus include new packet scheduling algorithms, fast link adaptation and advanced retransmission schemes. In addition, MIMO operation allows exploiting the available multi-user diversity in both time and frequency as well as spatial domains [6, 7]. Another important aspect for achieving such performance improvements is obtaining accurate channel state information (codeword) containing both MIMO ranking information and channel quality indicator (CQI) measurement from each mobile station within the serving cell. In this paper, we study the performance of advanced proportional fair (PF) scheduler schemes in higher velocity scenario for best effort (BE) traffic in LTE context.

It is generally fairly well understood that the overall radio system performance depends heavily on comprehensive RRM functionality. Moreover, PS is the main RRM entity to determine the crucial system parameters in terms of throughput, coverage and fairness. Recent investigations on different multi-user packet scheduling principles presented in the literature [8, 9, 10] reveal the potential of such techniques in OFDMA based systems. Most of them are based on PF scheduling principle and do not consider high mobility user scenarios. On the other hand, increased user velocities imply severe limitations on the observed SINR reported from UE and hence to overall system throughput. The reported feedback is suffering from errors and is not able to follow the fast fading. Thus, the frequency and spatial packet scheduling gains will be lost. Stemming now from our previous work in advanced PS scheduling development reported in [6, 7, 12], we extend our studies here to find the limitations and throughput losses due to mobility. Furthermore, we apply different simulation cases for 3 km/h, 30 km/h and 120 km/h user velocities and investigate the limits of achieved gains from frequency – spatial domain packet scheduling [13, 14]. The system model used for the performance evaluations of the scheduling methods presented in this paper is according to the 3GPP evaluation criteria [2]. The overall outcome is measured in terms of average cell throughput and coverage, and fairness distribution.

The rest of the paper is organized as follows: Section II overview the packet scheduling process and applied scheduling algorithms in the performance study. Section III presents the overall system model and simulation assumptions. The simulation results and analysis are presented in Section IV, while the conclusions are drawn in Section V.

II. PACKET SCHEDULING PROCESS

First, Packet scheduler is located in the BS considering downlink direction and as a part of RRM framework [4] together with LA and HARQ Manager distributes the available system resources based on specific users' priority metric calculation. The scheduling decision is based on received

This work is supported by the Qatar Telecom (Qtel) Grant No. QUEX-Qtel-09/10-10.

users' signaling information in terms of acknowledgements (ACK/NACK) and channel state information (CQI reports) per given transmission time interval (TTI) and per frequency domain physical resource block (PRB). Spatial domain functionality requires additional UEs' ranking information for possible MIMO modes subtracted also from the reported codeword. Based on this information, the BS then decides, among other RRM issues, whether the particular time-frequency resource is used for (i) transmitting only one stream to a specific UE, (ii) two streams for a specific UE (SU-MIMO) or (iii) two streams to two different UEs (MU-MIMO).

Here we describe the scheduling algorithms used in our performance studies based on priority metric calculation for each user. The advanced scheduling algorithms are based on widely used two-stage multi-stream PF approach with MIMO functionality [6, 11, 12]. In time-domain (TD) stage, within each TTI, UEs are ranked based on the full bandwidth channel state information and corresponding wideband throughput calculations. Considering MIMO case, different spatial multiplexing possibilities (one-stream, dual-stream SU, dual-stream MU) are taken into account, in calculating all the possible reference throughputs. In the second stage, the scheduling functionality is expanded in frequency-domain (FD) and spatial-domain (SD) where the actual PRB allocation takes place. Consequently, the needed PRB's for pending re-transmissions (on one stream-basis only) signaled through HARQ channels are reserved and the rest available PRB's are allocated to the selected UE's from the first stage. The actual priority metric in FD/SD stage is evaluated at PRB-level taking into account the available stream-wise channel state information and the corresponding throughput calculations. The overall scheduling process and RRM functionality are presented in Figure 1.

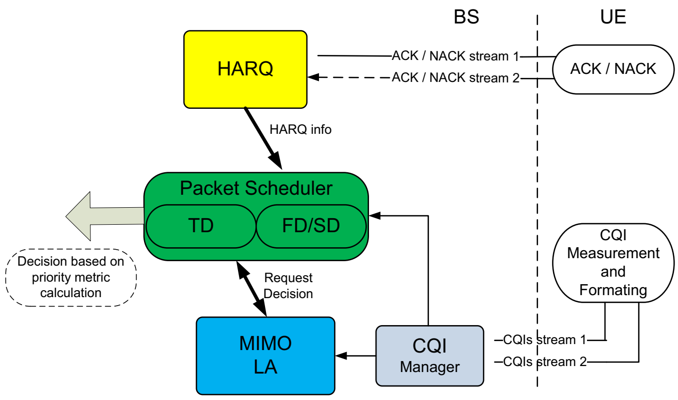


Figure 1. Packet scheduling.

A. Raund Robin

The users are served in a consecutive order without exploiting their feedback information. The priority calculation is based on the queueing time of each UE.

B. MIMO aware Proportinal Fair Scheduler

For the PF scheduler, scheduling decision per TTI is based on the following priority metric

$$\gamma_{i,k,s} = \arg\max_{i} \left\{ \frac{R_{i,k,s}(n)}{T_i(n)} \right\} \tag{1}$$

in which $R_{i,k,s}(n)$ is the estimated instantaneous throughput of user i at sub-band k on stream s for the time instant (TTI) n (calculated based on the CQI reports through e.g. EESM mapping [1]). $T_i(n)$, in turn, is the average delivered throughput to the UE i during the recent past and is calculated by

$$T_i(n) = \left(1 - \frac{1}{t_c}\right) T_i(n-1) + \frac{1}{t_c} R_i(n-1) \tag{2}$$

Here, t_c controls the averaging window length over which the average delivered throughput is calculated [6, 12] and $R_i(n-1)$ is the actually delivered throughput to user i at previous TTI n-1, calculated over all sub-bands k and possible streams s. In general, $1/t_c$ is also called the forgetting factor.

Considering the previous TD and FD/SD stages described earlier, the above metrics are used as follows:

a) TD: Metric (1) is evaluated over the full bandwidth and for different stream options to rank the I_{TOT} UE's. Out of these, $I_{\mathrm{BUFF}} < I_{\mathrm{TOT}}$ UE's with highest metric are picked to the following FD/SD stage. In the following, this subset is called scheduling candidate set (SCS).

b) FD/SD: The access to individual PRB and stream(s) is granted for the user(s) belonging to the above SCS with the highest metric (1) evaluated for the particular PRB and stream at hand.

C. Modified Multistream Proportional Fair (MMPF):

Stemming from the earlier work in [10], the modified multistream PF metric is calculated as:

$$\overline{\gamma}_{i,k,s} = \arg\max_{i} \left\{ \left(\frac{CQI_{i,k,s}(n)}{CQI_i^{avg}(n)} \right)^{\alpha_1} \left(\frac{T_i(n)}{T_{tot}(n)} \right)^{-\alpha_2} \right\} \tag{3}$$

Here, α_1 and α_2 are scheduler optimization parameters ranging basically from 0 to infinity, $CQI_{i,k,s}$ is the CQI of user i at PRB k and stream s, and $CQI^{avg}{}_i$ is the average CQI of user i calculated using

$$\begin{aligned} CQI_i^{avg}(n) &= \left(1 - \frac{1}{t_c}\right) CQI_i^{avg}(n-1) \\ &+ \frac{1}{t_c} \frac{1}{K_{TOT}} \frac{1}{S_{TOT}} \sum_{s=1}^{S_{TOT}} \sum_{k=1}^{K_{TOT}} CQI_{i,k,s}(n) \end{aligned} \tag{4}$$

In above, K_{TOT} is the total number of available PRB's while S_{TOT} denotes the maximum number of streams which is here always two (max two streams). In (3), $T_{tot}(n)$ is the average delivered throughput (during the recent past) to all users I_{BUFF} served by the BS and is calculated as

$$T_{tot}(n) = \left(1-\frac{1}{t_c}\right)T_{tot}(n-1)+\frac{1}{t_c}\frac{1}{I_{BUFF}}\sum_{i\in\Omega(n-1)}T_i(n-1) \qquad (5)$$

In general, the scheduling metric in (3) is composed of two elements affecting the overall scheduling decisions. The first dimension measures in a stream-wise manner the relative instantaneous quality of the individual user's radio channels against their own average channel qualities while the second dimension is related to measuring the achievable throughput of individual UE's against the corresponding average throughput of scheduled users. For high velocity user scenarios both ratios will be affected by the erroneous feedback provided by the UE and thus the metric calculation will impact the user scheduling priority. The power coefficients α_1 and α_2 are additional adjustable parameters for tuning and controlling the exact scheduler statistics. This will be demonstrated later using radio system simulations.

III. System Simulation Model

The performance of the proposed scheduling schemes is evaluated in extensive quasi-static system simulator for LTE downlink providing traffic modeling, multiuser packet scheduling and link adaptation including HARQ following the 3GPP evaluation criteria [2]. The 10 MHz system bandwidth is divided into 50 physical resource blocks (PRB) each consisting of 12 sub-carriers with sub-carrier spacing of 15 kHz. The mobile stations measure the actual channel states based on pilot signals and report it to the BS together with MIMO ranking possibilities - single-stream SINR as well as both single user (SU) and multi-user (MU) dual-stream SINR's at the corresponding detector output. The actual reported CQI's are based on received signal-to-interference-and-noise ratios (SINR), calculated by the UE's for each PRB. Here the UE's are assumed to use linear MMSE (LMMSE) receivers for MIMO 2x2 system simulation.

In a single simulation run (Macro case 1 scenario [2]), mobile stations are randomly distributed over standard hexagonal cellular layout with altogether 19 cells each having 3 sectors. The number of active users per sector is set to 20 and the UE velocities are changed according to the scenario - 3km/h, 30km/h or 120km/h. The used MIMO scheme for performance evaluation purposes is per-antenna rate control (PARC) with two transmit antennas at the BS and two receive antennas at each UE. The main simulation parameters and assumptions are summarized in Table I.

The RRM functionalities are controlled by the packet scheduler together with link adaptation and HARQ entities. Moreover, link adaptation functionality consist of removing CQI imperfections, estimating supported data rates and MCS's, and stabilizing the 1st transmission Block Error Probability (BLEP) to the target range (typically 10-20%). Simple admission control scheme for keeping the number of UEs per cell constant is used for performance evaluation simulations. HARQ is based on SAW protocol and a maximum number of three re-transmissions is allowed. MIMO functionality requires individual HARQ entry per stream which is also implemented. Link-to-system level mapping is based on the effective SINR mapping (EESM) principle [2].

TABLE I. DEFAULT SIMULATION PARAMETERS.

Parameter	Assumption
Carrier Frequency / Bandwidth	2000MHz / 10 MHz
Number of active sub-carriers	600
Sub-carrier spacing	15kHz
Sub-frame duration	0.5 ms
Channel estimation	Ideal
PDP	ITU Typical Urban 20 paths
Minimum distance between UE and cell	>= 35 meters – Macro
Average number of UE's per sector	20
Max. number of frequency multiplexed UEs	10
UE receiver type	LMMSE
Shadowing standard deviation	8 dB
UE speed	3km/h
Total BS TX power (Ptotal)	46dBm
Fast Fading Model	Jakes Spectrum
CQI reporting schemes	Full CQI
CQI log-normal error std.	1 dB
CQI reporting time	5 TTI
CQI delay CQI quantization CQI std error	2 TTIs 1 dB 1 dB
MCS rates	QPSK (1/3, 1/2, 2/3), 16QAM (1/2, 2/3, 4/5), 64QAM (1/2, 2/3, 4/5)
ACK/NACK delay	2ms
Number of SAW channels	6
Maximum number of retransmissions	3
HARQ model	Ideal chase combining (CC)
1st transmission BLER target	20%
Scheduler forgetting factor	0.0025
Scheduling schemes used	RR, PF , MMPF
Traffic type	Best Effort
Simulation duration (one drop)	5 seconds
Number of drops	10

IV. Simulation Results

In this section, we present the results obtained from the system simulations using the PS algorithms described in the paper for different user velocities. The system-level performance is generally measured and evaluated in terms of:

- Cell throughput distribution (Mbps)
- Cell-edge coverage
- Fairness distribution

In general, cell throughput is defined as the number of successfully delivered user bits per unit time. Coverage, in turn, corresponds to 5% probability point in the throughput CDF. Fairness is measured using the Jain's fairness index [15].

The performance of the advanced MMPF scheduler is evaluated using different power coefficients α_1 and α_2. Emphasizing on the importance of the throughput estimate in the priority metric calculation in equation (3), α_2 is fixed here to

1 and different values are then demonstrated for α_1 [12] i.e., the CQI measurement ratio. Consequently, large α_2 values increase the effect of the second term in priority metric calculation based on throughput estimation and the scheduling algorithm would behave like maximum throughput scheduler, which implies reduced fairness distribution. The used values for α_1 coefficient are defined as $\alpha_1 = \{1,2\}$. Moreover, the power coefficient values are presented as index M, where M1 represents the first couple, i.e., α_1=1, α_2=1 and M2: α_1=2, α_2=1.

Figure 2 illustrates the average cell throughput and cell-edge coverage for the different scheduling schemes and velocities. The performance at 3 km/h is illustrated in sub-figures (a) and (b). The performance of PF scheduler in terms of total throughput is significant compared to simple RR scheduler. In this case the achieved gain of 38% is due to frequency- spatial diversities and accurate user's throughput estimates. The throughput losses for MMPF cases M1 and M2 when compared to PF scheduler are 15-11% respectively due to increased cell coverage [12]. Similarly, the corresponding coverage gain for PF over RR scheduler is 23%. The MMPF scheduler clearly outperforms the rest of the scheduling schemes and the achieved gains are 42% for M1 case and 32% for M2 case when compared to PF scheduler, and more than 70% when compared to RR scheduler. Increasing the velocity to 30 km/h (sub- figures (c) and (d)) reduces the overall system throughput for all channel dependent scheduling schemes with around 20%. We still observe small throughput gains compared to blind RR scheduler. Similarly, the coverage has dropped with additional 18% for PF and MMPF schedulers compared to earlier scenario, while no change is observed for RR. The whole system performance changes significantly in 120 km/h case shown on sub-figures (e) and (f). Without any surprise, the PF scheduler has identical performance as RR in this simulation scenario. It can be explained with the reduced channel coherence time and with the obtained erroneous user feedback. The throughput losses for RR and PF schedulers are around 10%, while MMPF scheduler has maintained the nearly same throughput performance. Moreover, increased cell- edge coverage in order of 35% is achieved when compared to PF. Clearly, the role of the power coefficient α_1 in priority metric calculation in MMPF scheduling scheme is demonstrated. The impact of the received over the average user feedback in the scheduling decision is used for achieving additional performance gains in terms of coverage and throughput.

The complete performance statistics for different user velocities are summarized in Table II.

TABLE II. OBTAINED PERFORMANCE STATISTICS COMPARED TO RR SCHEDULER FOR DIFFERENT VELOCITIES.

Coverage Gain [%] / Throughput Gain[%]	Performance gains		
	3 km/h	*30 km/h*	*120 km/h*
PF	30 / 38	6 / 18	1 / 11
MMPF-M1	86 / 17	46 / 2	34 / 4
MMPF-M2	73 / 22	39 / 22	24 / 8

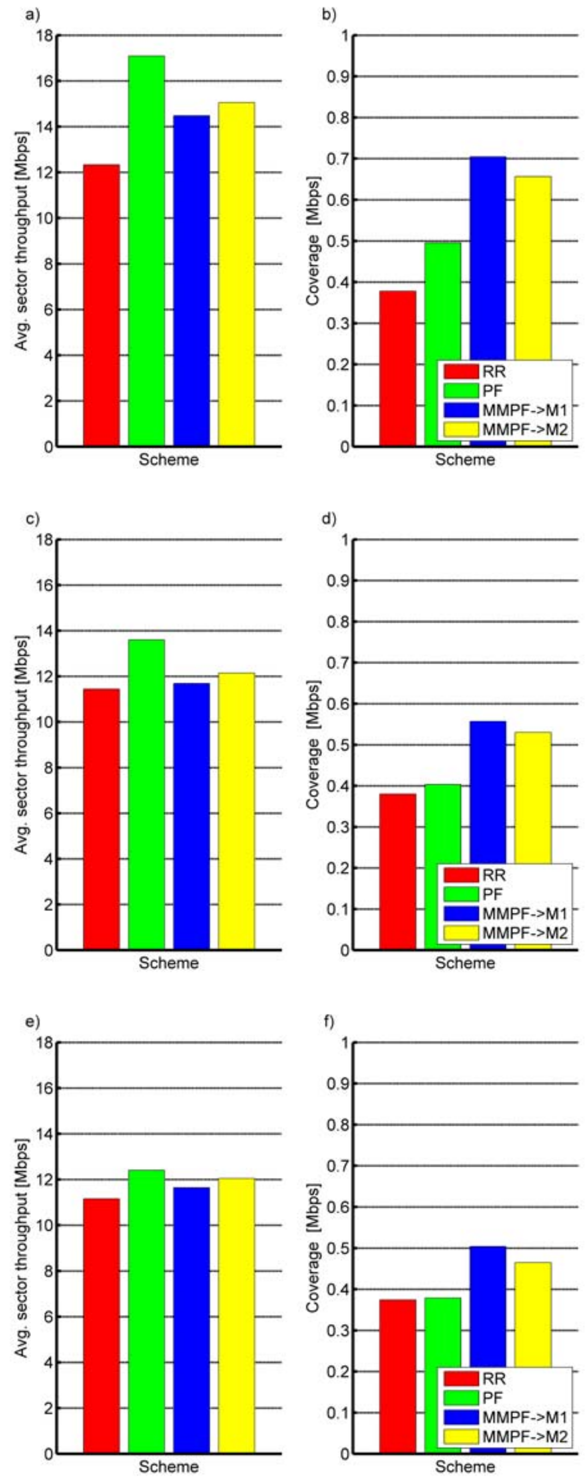


Figure 2. Average sector throughput and coverage for different scheduling schemes and velocity scenarios in 3 km/h (a, b), 30 km/h (c, d) and 120 km/h (e, f).

Figure 3. illustrates the HARQ distributions for the different scheduler schemes and velocity scenarios. Clearly, the 20% BLER target rate is achieved in all simulated cases. Moreover, the MMPF scheduler provides slight increase in probability of successful fist transmission of around 2% for all the cases. On the other hand, increased velocity reflects also on increasing the HARQ retransmissions as can be clearly observed for 30 km/h and 120 km/h simulation cases. The MMPF scheduler tends to show a negligible increase in second retransmission probability compared to normal PF scheduler due to the CQI based priority metric calculation.

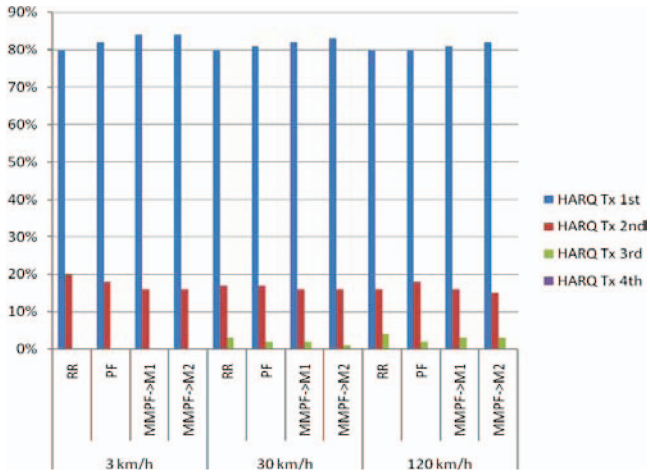


Figure 3. Probability of successful reception for HARQ transmission for different scheduling schemes and velocity scenarios in 3 km/h , 30 km/h and 120 km/h.

Further demonstrations on the effect of the increased user velocity is shown on Figure 4 where the modulation and coding scheme (MCS) distributions for different schedulers and velocity schemes are presented. The decrease in higher order modulation usage (more than 5%) leads to the increase in the lower ones for handling the retransmissions and improving the cell coverage. In all the simulated cases, the MCS distribution behaviour has a relatively similar trend following the PF scheduling principle and the choice of the power coefficients in the MMPF packet scheduling scheme. In general, the use of higher-order modulations corresponds to increased overall system throughput and coverage. Clearly, increased velocity affects on decreasing MCS selection and usage of more robust ones. Even with slightly lower MCS distributions, the MMPF scheduler achieves nearly the same system performance when compared to PF scheduler in terms of throughput and increased cell coverage as shown in Figure 2.

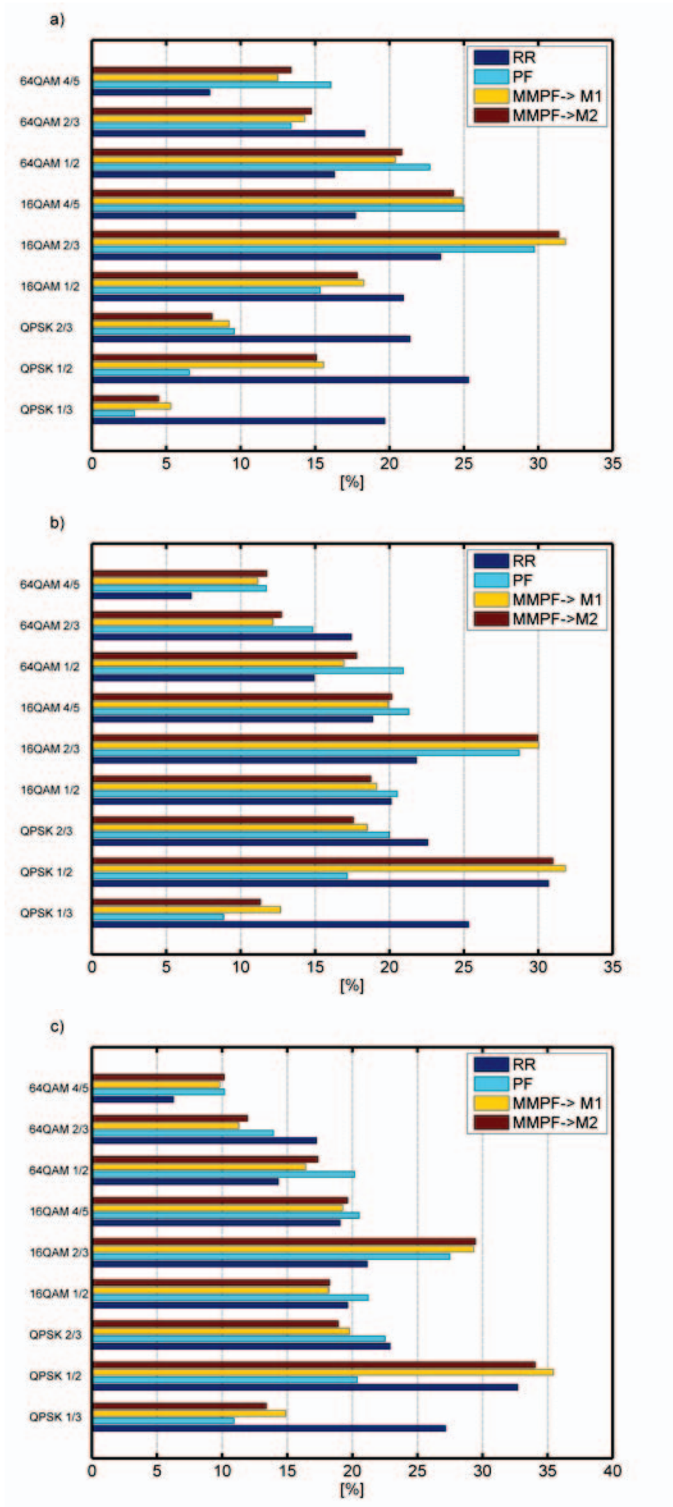


Figure 4. MCS distributions [%] for different scheduling principles and velocity scenarios in 3 km/h (a), 30 km/h (b) and 120 km/h (c).

Figure 5 illustrates the Jain's fairness index [15] per scheduling scheme calculated using the truly realized throughputs at each TTI for all 20 UE's and over all the simulation runs. The value on the x-axis corresponds to the used scheduler type (1 refers to RR scheduler, 2 refers to PF scheduler, etc.). We observe that all scheduling schemes have identical fairness distributions for the simulated velocity scenarios and the proposed MMPF scheduler clearly outperforms them. The received fairness gains are in range of 36%-38% when compared to the RR and PF scheduling schemes. The effect of user velocity over scheduling decision is clearly seen for PF-based schemes, where a fairness distribution among UE is decreasing when increasing UE velocity.

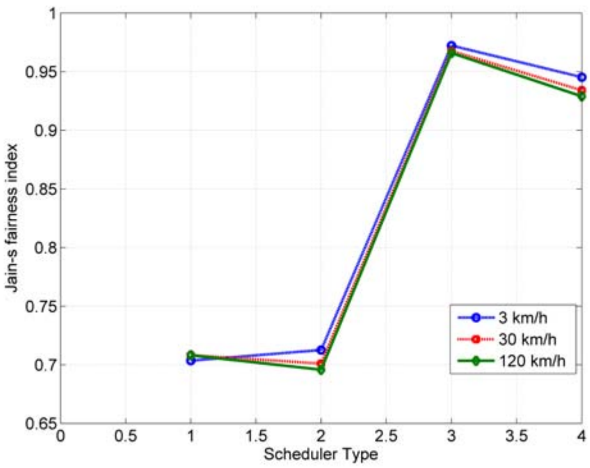


Figure 5. Jain's fairness index per scheduling scheme.

V. CONCLUSIONS

In this paper, we have studied the potential of advanced multi-user packet scheduling algorithms in OFDMA type radio system context, using UTRAN long term evolution (LTE) downlink in Macro cell best effort environment as practical example case. We demonstrated the benefits of multi-domains (time-, frequency- and spatial) scheduling metric in achieving increased user fairness and coverage taking into account instantaneous channel qualities (CQI's) as well as resource allocation fairness in different velocity scenarios. Overall, the achieved throughput performance together with coverage and fairness statistics were assessed, by using extensive system simulations, and compared against more traditional scheduling schemes. In high-velocity scenarios the schedulers are losing the diversity gains and thus affect reducing the overall system throughput by nearly 25 % for 30 km/h and 120km/h cases. The studied MMPF scheduling metric calculations based on combined UE channel feedback and throughput estimation metric offers better control over the ratio between the achievable cell/UE throughput and coverage increase, as well as increased UE fairness. As a practical example, the fairness in resource allocation together with cell coverage can be increased significantly (more than 25%) by keeping the cell throughput nearly constant.

VI. REFERENCES

[1] 3GPP Technical Specification Group Radio Access Network Technical Specification 36.300 v8.0.0: Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (EUTRAN); Overall description, march 2007

[2] Physical Layer Aspects for Evolved UTRA", 3GPP Technical Report TR 25.814, ver. 7.1.0, Oct.2006.

[3] ITU global standard for international mobile telecommunications "IMT-Advanced", IMT-ADV/1-E,Mar. 2008.

[4] WiMAX Forum™, http://www.wimaxforum.org, Aug. 2006.

[5] E. Dahlman et al., 3G Evolution: HSPA and LTE for Mobile Broadband, Academic Press, 2007.

[6] S. Nonchev and M. Valkama, "Efficient packet scheduling schemes for multiantenna packet radio downlink", in Proc. Fifth Advanced Int. Conf. Telecommunications (AICT-2009), Venice, Italy, May 2009.

[7] Nonchev and M. Valkama, "Advanced radio resource management for multiantenna packet radio systems", submitted to International Journal of Wireless & Mobile Networks (IJWMN), 14 pages

[8] Y. Sun et al., "Multi-user scheduling for OFDMA downlink with limited feedback for evolved UTRA", in Proc. IEEE VTC-2006, Montreal, Canada, September 2006.

[9] N. Wei, et al., "Performance of MIMO with frequency domain packet scheduling", in Proc. IEEE VTC-2007, Dublin, Ireland, May 2007

[10] C. Wengerter, J. Ohlhorst, and A.G.E Von Elbwert, "Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA", in Proc. IEEE VTC-2005, Stockholm, Sweden, May 2005.

[11] G. Monghal, et al., QoS Oriented Time and Frequency Domain Packet Schedulers for The UTRAN Long Term Evolution, in Proc. VTC2008-Spring, vol.3, pp. 2532-2536, May 2008

[12] S. Nonchev and M. Valkama, "A new fairness-oriented packet scheduling scheme with reduced channel feedback for OFDMA packet radio systems", Int. Journal of Communications, Network and System Sciences (IJCNS), 2009, vol. 7, pp. 608-618.

[13] P. Kela, J. Puttonen, N. Kolehmainen, T. Ristaniemi, T. Henttonen, and M. Moisio, "Dynamic Packet Scheduling Performance in UTRA Long Term Evolution Downlink," in Proceedings of the International Symposium on Wireless Pervasive Computing (ISWPC'08), May 2008.

[14] J. Puttonen, N. Kolehmainen, T. Henttonen, M. Moisio and M. Rinne, "Mixed Traffic Packet Scheduling in UTRAN Long Term Evolution Downlink", in Proceedings of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Cannes, France, 15-18th Sep. 2008.

[15] D. Chui and R. Jain, "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks", Computer Networks and ISDN Systems, 1989.

Publication 7

S. Nonchev and M. Valkama, “QoS-oriented packet scheduling for efficient video support in OFDMA-based packet radio systems,” in *Proceedings of the 4th International Conference on Multiple Access Communications* (MACOM-2011), pp. 168-180, Trento, Italy, September 2011.

QoS-Oriented Packet Scheduling for Efficient Video Support in OFDMA-Based Packet Radio Systems

Stanislav Nonchev and Mikko Valkama

Department of Communications Engineering, Tampere University of Technology,Tampere, Finland
stanislav.nonchev@tut.fi, mikko.e.valkama@tut.fi

Abstract. Next-generation mobile networks will provide users with high data rates, increased mobility and various services. The initial step made in LTE and LTE-A is adopting the OFDMA air interface and utilization of dynamic resource allocation techniques maximizing the cell throughput and coverage as part of enhanced radio resource management (RRM) functionality. Moreover, the control of the network resource division among users is performed by packet scheduler and different scheduling strategies are applied according to traffic scenario. Typically, video traffic (real-time video streaming, mobile TV etc.) requires higher data rates and certain quality of service (QoS) constrains, i.e. packet size, arrival rate, head-of-line (HOL) packet delay, etc. In this article, we propose a flexible and fairness-oriented packet scheduling approach for real-time video delivery, built on advanced QoS-aware multiuser proportional fair (PF) scheduling principle. The performance of the overall scheduling process is investigated in details in terms of cellular system capacity, resource allocation fairness and video traffic QoS guarantees. Experimental results reveal the upper bounds of real-time video traffic support in downlink LTE multiuser network.

Keywords: Cellular system performance, radio resource management, packet scheduling, QoS; proportional-fair, fairness.

1 Introduction

The Third Generation Partnership Project (3GPP) organization developed the Long Term Evolution (LTE) standard with support to high-data-rates, increased mobility, low-latency and packet optimized radio access. LTE uses single-carrier frequency-division multiple access (SC –FDMA) for the uplink (UL) and orthogonal FDMA in downlink (DL). Scalable bandwidth operation, exploitation of diverse MIMO technologies and advanced convergence techniques are some of the key benefits in OFDMA-based developments. [1] – [3]. The available spectrum is divided into large number of orthogonal subcarriers forming the basic time-frequency transmission resource - physical resource block (PRB). This allows multi user access and efficient reduction of the effects of inter-symbol interference (ISI) and inter-carrier interference (ICI). Therefore, increased spectral efficiency and high data rates are achieved [5], [6].

The mobile traffic boost due to increased usage of video applications such as video streaming, mobile TV and multimedia online gaming requires increased system performance and user QoS guarantees. On the other hand, performance improvements are typically obtained through proper radio resource management functionalities and exploiting the available multi-user diversity in both time and frequency as well as spatial domains [6]–[12]. Another requirement for achieving such performance improvements is obtaining accurate channel feedback from each mobile station (MS) within the serving cell. Particularly, each MS can measure the effective signal-to-interference-plus noise-ratio (SINR), per active subcarrier or block of subcarriers, and send back the obtained channel state to the base station (BS) in terms of channel quality information (CQI) reports. Moreover, considering multiple-input multiple-output (MIMO) systems, the stream wise feedback is provided as a codeword containing both MIMO ranking information and CQI measurement. Most of the performance studies in OFDMA based mobile networks are based on simulations for best-effort traffic [9] and VoIP traffic [13] with advanced PS strategies but streaming video traffic has not been extensively studied with such scheduling techniques. Recently, Luo et al. in [14] propose quality-driven cross-layer optimized video delivery scheduling strategy, while Basukala et al. in [15] demonstrated the performance of packet scheduler schemes serving video streaming users. The initial potential of LTE video capacities is also demonstrated in [16] with simple frequency diversity scheme.

Clearly, the overall radio system performance in terms of throughput, coverage and fairness, depends heavily on PS functionality, being the key ingredient in the radio resource management process. Most of the literature studies on different multi-user packet scheduling techniques demonstrate that the well-known proportional fair (PF) scheduling principle is the right choice for OFDMA based systems [7], [8]. On the other hand, despite increased throughput and fairness the PF scheduler cannot guarantee the packet delay constraint for video services by default. Thus, higher system throughput does not guarantee higher video quality and therefore video performance metric should be taken into account in PS decision. Only a few scheduling strategies have considered user's QoS requirements together with practical system and application constrains [17]-[21]. Stemming now from our previous work in advanced PS scheduling developments reported in [7], [20], [21], we extend our studies here to incorporate QoS requirements into scheduling decisions by effectively controlling user fairness and BS's RRM process for increased video support.

Furthermore, we apply different simulation cases for video traffic applications investigating the limits of achieved gains from time-frequency-spatial domain packet scheduling with limited feedback and QoS constrains [22], [23]. The system model used for the performance evaluations of the proposed scheduling methods presented in this paper is according to the 3GPP evaluation criteria [2]. The overall outcome is measured in terms of capacity – the number of supported video streams of different users per cell.

The rest of the paper is organized as follows: Section 2 gives an overview of the packet scheduling process and describes the proposed QoS aware multi-stream PF (QoS-MSPF) scheduling scheme. Section 3, in turn, presents the overall system model and simulation assumptions. The simulation results and detailed analysis are presented in Section 4, while the conclusions are drawn in Section 5.

2 Packet Scheduling Process

Fig. 1 illustrates the overall RRM framework. The key entities in it are PS, LA and HARQ manager. Located in the BS, the PS functionality consists of selecting the users (UEs) to be scheduled on transmission time interval (TTI) basis and allocating the required frequency resources (PRBs). In more details, the scheduling decision is based on priority metric calculation for individual UEs depending on the selected scheduling strategy. Some of the advanced PF based scheduling techniques require users' CQIs per given TTI and per frequency domain PRB. MIMO functionality, in turn, requires both single-stream and dual-stream CQI feedback by each UE. In addition, PS is interacting with LA entity for choosing the modulation and coding schemes (MCS) for individual PRBs and obtaining information for new transmissions or retransmissions from HARQ manager. BS buffer information is required for verification of keeping with the set packet delay budget.

The proposed QoS-aware multi-stream PF (QoS-MSPF) scheduler is based on widely used two-stage PF approach (see e.g. [9], [21]) with additional QoS enabled guarantees for video traffic. In terms of the actual metric calculations, the proposed QoS-aware multi–stream PF (QoS-MSPF) scheduler uses the following metric.

$$\overline{\gamma}_{i,k,s} = \arg\max_{i} \left\{ \delta_i(n) \left(\frac{CQI_{i,k,s}(n)}{CQI_i^{avg}(n)} \right)^{\alpha_1} \times \left(\frac{T_i^{se}(n)}{T_{i,k,s}(n)} \frac{T_i(n)}{T_{tot}(n)} \right)^{-\alpha_2} \right\} \tag{1}$$

which can be understood as an extension of the authors' earlier work in [20]. In the above metric, α_1 and α_2 are scheduler optimization parameters ranging basically from 0 to infinity. In expression (1), $T_i^{se}(n)$ is an estimate of the user throughput if user i is scheduled on sub-frame basis according to [18] and $T_{i,k,s}(n)$ is the estimated achievable throughput of user i at PRB k and stream s. $T_i(n)$ corresponds to average delivered throughput to the UE over the past and T_{tot} is the average delivered throughput (during the recent past) to all users ranked in TD stage served by the BS. $CQI_{i,k,s}$ is the CQI of user i at PRB k and stream s, and CQI^{avg}_i is the average CQI of user i calculated by traditional recursive method [24].

The $\delta_i(n)$ is, in turn QoS delay function factor defined as:

$$\delta_i(n) = \frac{max\left(d_i / B_i(n)\right)}{d_{max}} \tag{2}$$

where d_i is the delay of the packet of user i in the transmit buffer $B_i(n)$ at time instant n, and d_{max} is the maximum delay allowed.

The scheduling metric in (1) is essentially composed of three parts affecting the overall scheduling decisions. The first term takes into account the QoS requirements, i.e. packet delay budget. The second ratio is the relative instantaneous quality of the individual user's radio channels over their own average channel qualities. The third ratio is divided into two parts. The first one takes into account the estimated throughputs of individual UE's and the second one the achievable over total throughputs. The power coefficients α_1 and α_2 are additional adjustable parameters.

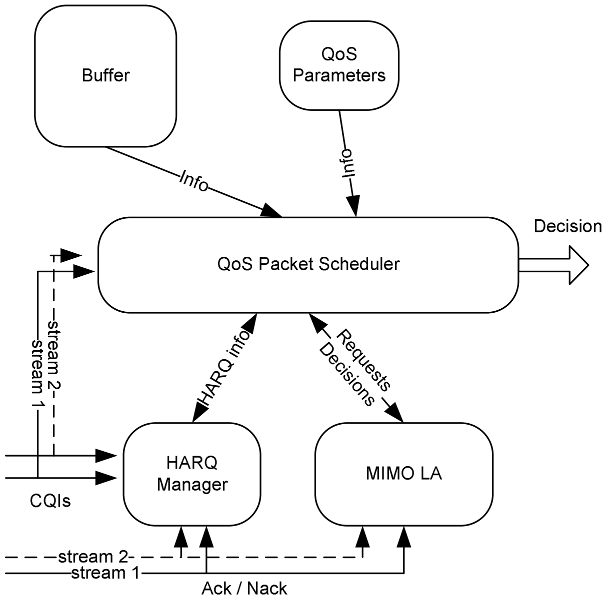


Fig. 1. RRM entities and packet scheduling process

In more details, in the first stage, which is the time-domain (TD) scheduling step executed within each TTI, UEs are ranked based on the full bandwidth channel state information, QoS parameters and the corresponding throughput calculations. These UEs form Group 1 of the whole Scheduling Candidates Set (SCS). Group 2 consists of users with pending retransmissions and Group 3 is formed by users having packet delay close to the delay budget. Furthermore, discard timer is used so that buffered packets that are already late will not be transmitted over the air interface. Thus, the accumulated delay for subsequent packet is also reduced.

In the second stage, the scheduling functionality is expanded in frequency-domain (FD) where the actual PRB allocation takes place.

Initially, the needed PRB's for pending re-transmissions (on one stream-basis only) signaled through HARQ channels are reserved – Group 2 UEs. The remaining PRBs are given to the delay sensitive users from Group 3 and the rest to the first transmission users from Group 1 determined through available buffer information. The actual priority metric in FD/SD stage is evaluated at PRB-level taking into account the available stream-wise channel state information, QoS parameters and the corresponding throughput calculations.

In general, dynamic resource allocation methods increase scheduling flexibility so that enough PRBs will be mapped to the UEs with good channel conditions and more resources will be available for the other SCS users. Consequently, prioritizing Group1 UEs will further decrease packet delays and reduce retransmissions. On the other hand, combined QoS- and channel-aware scheduling can offer performance enhancements at the expense of increased scheduling complexity, in terms of scheduling metric calculations and increased signaling overhead.

Keeping these at reasonable levels requires thus some constraints on the scheduling algorithm, so for simplicity we assume here that only one MIMO mode (SU or MU) and fixed modulation and coding scheme (MCS) is allowed per user within one scheduling element. Moreover, we limit the number of users for multiplexing in TD stage to further reduce the signaling overhead and complexity of FD/SD scheduling. Consequently, further decrease is examined by exploiting different reduce-feedback reporting schemes.

3 System Simulation Model

3.1 Video Traffic

Real –time video services are modeled as follows:

- Each frame of video data arrives at a regular interval determined by the number frames per second.
- Each frame is decomposed into a fixed number of slices, each transmitted as a single packet. The size of these packets/slices is modeled to have a truncated Pareto distribution.
- The video encoder introduces encoding delay intervals (modeled by a truncated Pareto distribution) between the packets of a frame.

In our studies, two different video streaming services with 128 kbps (Scenario 1) and 256 kbps (Scenario 2) constant bit rates (CBR) source video data are used in the simulations. The corresponding mean packet size with truncated Pareto distribution and mean inter-arrival packet time are 100 bytes - 6ms for Scenario 1, and 200 bytes - 4ms for Scenario 2 [23]. Packet delay budget, as well as discard timer threshold is set to 20 ms.

3.2 Video Capacity Estimation

The main target is to estimate the number of UEs that can be supported by the system based on deployed scenarios. Following 3GPP evaluation metrology, a user is considered to be in outage if 2% of the video packets for the user are erroneous or discarded due to exceeding delay limit when monitored over the whole video session duration. On the other hand, video capacity is defined as number of supported users per cell without exceeding the system saturation point. Here, the system saturation point is set to 5% of the cell outage level, i.e. to the point where 95% of the users in the cell are satisfied (having maximum of 2% packet loss rate as described above).

3.3 Simulation Environment

Quasistatic system level simulator is used to evaluate the proposed scheduling scheme in video traffic scenarios for LTE downlink. It includes detailed traffic modeling, multiuser packet scheduling and link adaptation including HARQ, following the 3GPP evaluation criteria [2].

The chosen Micro Case 1 deployment scenario consists of 10 MHz system bandwidth divided into 50 physical resource blocks (PRB) containing 600 data sub-carriers. In a single simulation run mobile stations are randomly distributed over standard hexagonal cellular layout with altogether 19 cells each having 3 sectors. Fast fading is updated per TTI according to Typical Urban (20 taps) radio channel, while shadow fading and distance dependent path loss remain constant during the whole simulation. Moreover, the UE velocity is fixed to 3 km/h. The main simulation parameters and assumptions are summarized in Table I.

The RRM functionalities are controlled by the packet scheduler together with link adaptation and HARQ entities. Moreover, link adaptation functionality consist of removing CQI imperfections, estimating supported data rates and MCS's, and stabilizing the 1st transmission Block Error Probability (BLEP) to the target range (typically 10-20%). HARQ is based on SAW protocol and a maximum of three retransmissions is allowed. MIMO functionality requires individual HARQ entry per stream which is also implemented. Link-to-system level mapping is based on the effective SINR mapping (EESM) principle [2].

The actual effective SINR calculations rely on estimated subcarrier-wise channel gains (obtained using reference symbols) and depend in general also on the assumed receiver topology. Here we assume per-antenna rate control (PARC) MIMO case, i.e. two transmits antennas at the BS and two receive antennas at each UE and the receivers are equipped with LMMSE detectors.

4 Simulation Results

In this section, we present the results obtained from the system simulations using the RRM algorithms described in the paper.

The system-level performance is generally measured and evaluated in terms of:

- Capacity
- Throughput
- CDF of the number of users scheduled per TTI
- CDF of the number of PRBs per UE
- CDF of packed delay
- Fairness distribution

In general, the video capacity depends on the video data rates, packet size and the choice of outage criteria. The user outage criteria is defined as 2% of the video packets are erroneous or discarded during the whole simulation. Capacity in turn corresponds to maximum number of supported users not exceeding 5 % cell outage level. The cell throughput is defined as the number of successfully delivered user bits per unit time. Fairness is measured using the Jain's fairness index [25].

The performance of the proposed QoS-MSPF scheduler is compared against reference PF scheduler with a delay dependent component for video traffic support [12] for the video traffic models as discussed in Section 3 A and evaluated using different power coefficients α_1 and α_2. Here α_2 is fixed here to 1 and the used values for α_1 coefficient are defined as $\alpha_1 = \{1,2\}$. Moreover, the power coefficient values are presented as index M, where M1 represents the first couple, i.e., α_1=1, α_2=1 and M2: α_1=2, α_2=1.

Table 1. Default Simulation Parameters

Parameter	Assumption
Carrier Frequency / Bandwidth	2000MHz / 10 MHz
Number of active sub-carriers	600
Sub-carrier spacing	15kHz
Sub-frame duration	0.5 ms
Channel estimation	Ideal
PDP	ITU Typical Urban 20 paths
Minimum distance between UE and cell	>= 35 meters – Macro
Max. number of frequency multiplexed UEs	10
UE receiver type	LMMSE
Shadowing standard deviation	8 dB
UE speed	3km/h
Total BS TX power (Ptotal)	46dBm
Fast Fading Model	Jakes Spectrum
CQI reporting schemes	Full CQI
CQI log-normal error std.	1 dB
CQI reporting time	5 TTI
CQI delay	2 TTIs
CQI quantization	1 dB
CQI std error	1 dB
MCS rates	QPSK (1/3, 1/2, 2/3), 16QAM (1/2, 2/3, 4/5), 64QAM (1/2, 2/3, 4/5)
ACK/NACK delay	2ms
Number of SAW channels	6
Maximum number of retransmissions	3
HARQ model	Ideal chase combining (CC)
1^{st} transmission BLER target	20%
Scheduler forgetting factor	0.0025
Scheduling schemes used	RPF, QoS -MSPF (proposed)
Simulation duration (one drop)	120 seconds
Number of drops	10

Fig. 2 illustrates the number of supported video mobile users and average cell throughput for the different QoS-aware schedulers in Scenario 1 system simulation case. The obtained results with the proposed scheduler schemes are compared with the reference PF scheduler achieving video capacity of 94 users per sector and average sector throughput of 13,2 Mbps (subfigures (a) and (b)). By using the first term (M1) of the new metric calculation for QoS-MSPF we achieve video capacity gain in the order of 13% and additional 8% throughput increase. The achieved gain mainly comes from better utilization of the resources due to the scheduler priority metric calculation, as well as the increase in frequency-spatial domain multiplexed users.

For coefficient α_1 equal to 2 (M2) the new QoS-MSPF scheduler achieves video capacity gains in the order of 15% and throughput increase of 7% compared to reference PF scheduler.

The performance statistics obtained for Scenario 2 demonstrate similar trends, as in the previous case, as shown in Fig. 3. Starting from the reference scheduling case and full CQI reporting scheme the obtained video capacity is 52 users. The reduced number of supported users is due to increased video packet size reflecting on fulfilling the

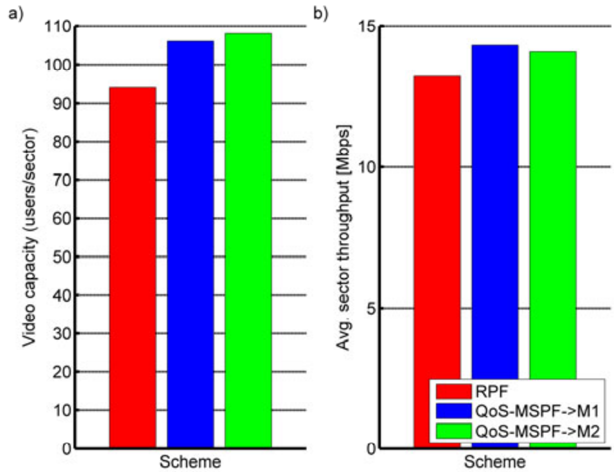


Fig. 2. Video capacity and average sector throughput for different scheduling schemes (a, b) in Scenario 1. M1-M2 refers to the proposed scheduler with different power coefficient values.

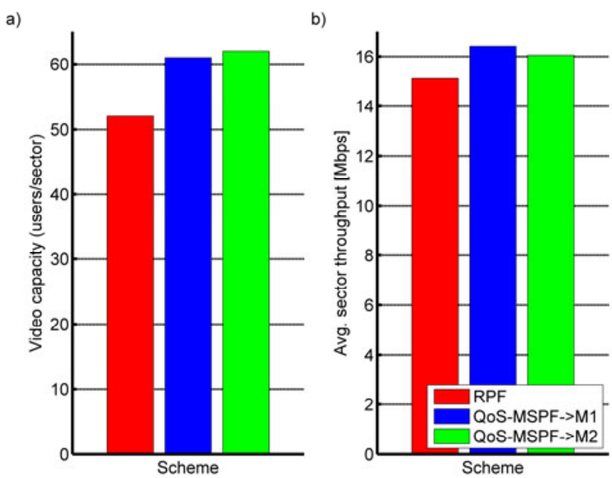


Fig. 3. Video capacity and average sector throughput for different scheduling schemes (a, b) in Scenario 2. M1-M2 refers to the proposed scheduler with different power coefficient values.

QoS requirements with more PRBs scheduled per user. In primary case M1, with full CQI, we obtain a 16% gain in video capacity correspond to 61 users and 9% throughput improvement. In the case of M2, additional 2% increase of the video capacity and throughput gains are achieved.

The performance results for both cases are summarized in Table 2.

Table 2. Obtained performance statistics compared to reference PF scheduler with different power coefficients (M1-M2) for the proposed scheduler

QoS -MSPF	Performance Statistics			
	Video Capacity Gain [%]		**Throughput Gain [%]**	
	Scenario 1 CBR 128 kbps	Scenario 2 CBR 256 kbps	Scenario 1 CBR 128 kbps	Scenario 2 CBR 256 kbps
M1	13	16	8	9
M2	15	18	7	6

Continuing on the evaluation of relative system performance using the proposed scheduler in different simulation cases are presented in Fig. 4 in terms of cumulative distribution function (CDF) of scheduled PRBs per TTI for individual scheduling scheme. Noticeably, the proposed scheduling scheme allocates more resources to the users shown on the Fig. 4. with right shift in CDF curve. Identical trends in PS functionality are seen in both simulation cases, which clearly correspond to increased system capacity.

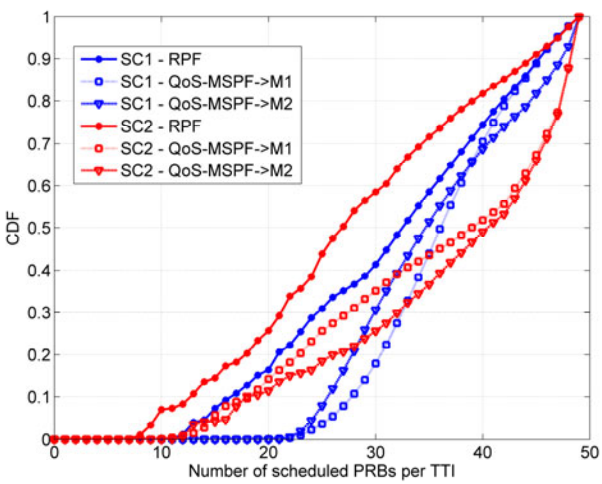


Fig. 4. CDF of scheduled PRB per TTI for simulated scenarios

Further illustrations on the obtainable system performance are illustrated on Fig. 5. with CDF of the number of scheduled users per TTI. We can clearly see from the figure that with the proposed scheduling scheme allows more users to be scheduled in each TTI and the achieved capacity gains are due to the increased usage of PRB resources as already concluded above. Similarity of the QoS-MSPF scheduler functionality in simulated scenarios is also seen here.

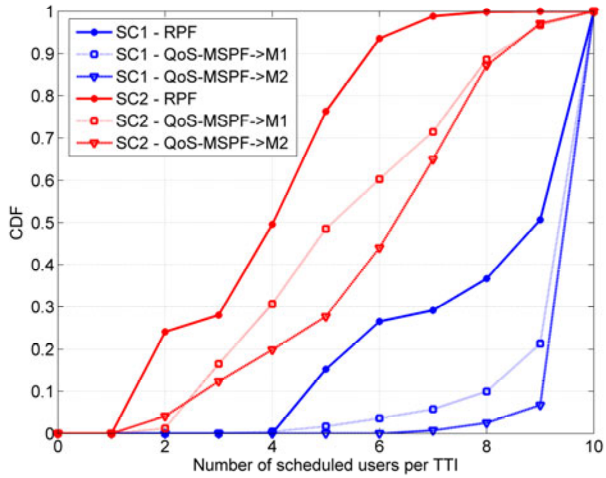


Fig. 5. CDF of scheduled users per TTI for simulated scenarios

Fig. 6 shows the CDF of packet delay for different schedulers. The CQI reporting process has a direct impact on the delay performance and therefore a combined approach of HOL and CQI criteria in priority metric calculation will benefit in such scenarios. The delay performance is strictly within the bounds of 20 ms for the simulated video traffic scenarios.

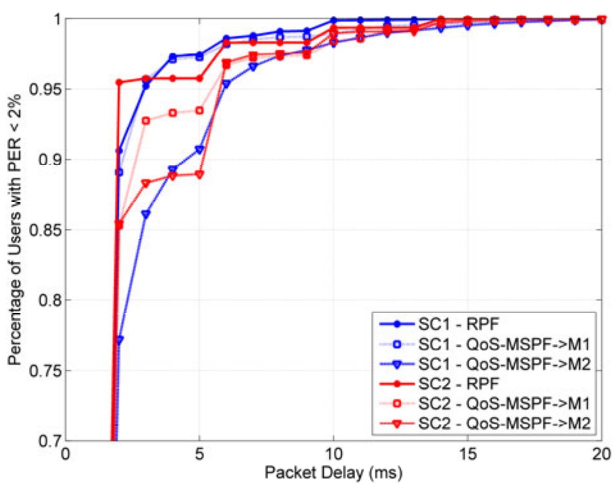


Fig. 6. CDF of packet delay for simulated scenarios

Fig. 7 illustrates the Jain's fairness indexes [24] for scheduling scheme based on the number of supported users.

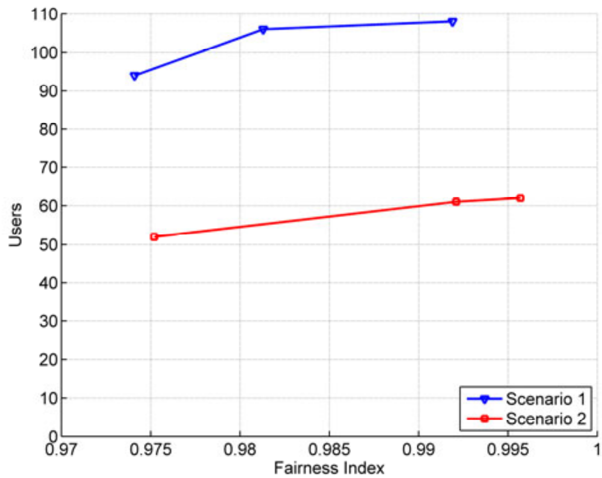


Fig. 7. Jian's fairness indexes of supported video users for simulated scenarios

The value of each curve corresponds to the used scheduler type (1st values refer to Reference PF scheduler, 2nd values refer to QoS-MSPF – M1 and 3rd values refer to QoS-MSPF – M2).We observe that the reference scheduling scheme has lower fairness indexes for all the simulated cases and the proposed QoS-MSPF scheduler clearly obtains better fairness. Having video traffic model with strict QoS requirements implies increased user fairness and thus fairness indexes above 0,90 (1 refers to complete fairness) are expected.

5 Conclusions

In this article, we have studied the potential of advanced QoS-aware multi-user packet scheduling algorithms for video traffic support in OFDMA type radio system context, using UTRAN long term evolution (LTE) downlink in Macro cell environment as practical example case. New video-optimized multi-stream proportional fair scheduler metric covering time-, frequency- and spatial domains was proposed that takes into account video traffic QoS requirements, instantaneous channel qualities (CQI's) as well as resource allocation fairness. Furthermore, detailed set of performance results for real-time video support are presented. As a practical example over 10 MHz bandwidth, more than 100 users can be supported for video traffic with source data rate of 128 kbps and more than 60 users can be supported in 256 kbps case.

Acknowledgements. This research has been financially supported by by the Finnish Funding Agency for Technology and Innovation (Tekes), under the project "Energy and Cost Efficiency for Wireless Access (ECEWA)".

References

1. 3GPP Technical Specification Group Radio Access Network Technical Specification 36.300 v8.0.0: Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (EUTRAN); Overall description, (March 2007)
2. Physical Layer Aspects for Evolved UTRA, 3GPP Technical Report TR 25.814, ver. 7.1.0 (October 2006)
3. ITU global standard for international mobile telecommunications IMT-Advanced, IMT-ADV/1-E (March 2008)
4. WiMAX Forum™ (August 2006), `http://www.wimaxforum.org`
5. Dahlman, E., et al.: 3G Evolution: HSPA and LTE for Mobile Broadband. Academic Press, London (2007)
6. Tripathi, N.D., et al.: Radio Resource management in Cellular Systems. Springer, Heidelberg (2001)
7. Nonchev, S., Valkama, M.: Advanced radio resource management for multiantenna packet radio systems. International Journal of Wireless & Mobile Networks (IJWMN) 2(2), 1–14 (2010)
8. Sun, Y., et al.: Multi-user scheduling for OFDMA downlink with limited feedback for evolved UTRA. In: Proc. IEEE VTC 2006, Montreal, Canada (September 2006)
9. Pedersen, K.I., Monghal, G., Kovacs, I.Z., Kolding, T.E., Pokhariyal, A., Frederiksen, F., Mogensen, P.: Frequency domain scheduling for OFDMA with limited and noisy channel feedback. In: Proc. IEEE Vehicular Technology Conference (VTC 2007 Fall), Baltimore, MD, pp. 1792–1796 (September 2007)
10. Wei, N., et al.: Performance of MIMO with frequency domain packet scheduling. In: Proc. IEEE VTC 2007, Dublin, Ireland (May 2007)
11. Wengerter, C., Ohlhorst, J., Von Elbwert, A.G.E.: Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA. In: Proc. IEEE VTC 2005, Stockholm, Sweden (May 2005)
12. Elsayed, K., Khattab, A.: Channel-aware earliest deadline due fair scheduling for wireless multimedia networks. Wireless Personal Communications 38, 233–252 (2008)
13. Fan, Y., Kuusela, M., Lundén, P., Valkama, M.: Performance of VoIP on EUTRA Downlink with Limited Channel Feedback. In: Proc. IEEE ISWCS 2008, Reykjavik, Iceland (October 2008)
14. Haiyan, L., et al.: Quality-Driven Cross-Layer Optimazed Video Delivery over LTE. IEEE Communication magazine 48, 102–109: doi:10.1109/MCOM.2010.5402671
15. Basukala, R., et al.: Impact of CQI Feedback Rate/Delay on Scheduling Video Streaming Services in LTE Downlink. In: Proc. International Conference on Communication Technology ICCT 2010, Nanjing, China (November 2010)
16. Talukdar, A.K., Mondal, B., Cudak, M., Ghosh, A., Wang, F.: Streaming Video Capacity Comparisons of Multi-Antenna LTE Systems. In: Proc. VTC Spring, pp. 1–5 (2010)
17. Monghal, G., et al.: QoS Oriented Time and Frequency Domain Packet Schedulers for The UTRAN Long Term Evolution. In: Proc. VTC 2008 Spring, vol. 3, pp. 2532–2536 (May 2008)

18. Kolding, T.E.: QoS-Aware Proportional Fair Packet Scheduling with Required Activity Detection. In: VTC2006 Fall (September 2006)
19. Nguyen, T.-D., Han, Y.: A Proportional Fairness Algorithm with QoS Provision in Downlink OFDMA Systems. IEEE Communication Letters 10(11) (November 2006)
20. Nonchev, S., Valkama, M.: A new fairness-oriented packet scheduling scheme with reduced channel feedback for OFDMA packet radio systems. Int. Journal of Communications, Network and System Sciences (IJCNS) 7, 608–618 (2009)
21. Nonchev, S., Valkama, M.: Efficient packet scheduling schemes with built-in fairness control for multiantena packet radio systems. Int. Journal on Advances in Networks and Services, IJANS (2009)
22. Talukdar, A., Cudak, M., Ghosh, A.: Streaming Video Capacities of LTE Air-Interface. In: Proceedings of Conference (ICC)...
23. 3GPP, Feasibility Study for Orthogonal Frequency Division Multiplexing (OFDM) for UTRAN enhancement. 3GPP TR25.892 (2004)
24. Jalali, A., Padovani, R., Pankaj, R.: Data throughput of CDMA-HDR high efficiency-high data rate personal communication wireless system. In: Proc. IEEE VTC 2000 Spring, Tokyo (May 2000)
25. Chui, D., Jain, R.: Analysis of the increase and decrease algorithms for congestion avoidance in computer networks. Computer Networks and ISDN Systems (1989)

Publication 8

S. Nonchev, M. Valkama and R. Hamila, “Advanced packet scheduling for efficient video support with limited channel feedback on MIMO LTE downlink,” in *Proceedings of IEEE GLOBECOM Workshop on Enabling Green Wireless Multimedia Communications,* pp.766-771, Houston, TX, December 2011.

Advanced Packet Scheduling for Efficient Video Support with Limited Channel Feedback on MIMO LTE Downlink

Stanislav Nonchev and Mikko Valkama
Tampere University of Technology
Department of Communications Engineering
P.O.Box 553, FI-33101, Tampere, Finland
stanislav.nonchev@tut.fi, mikko.e.valkama@tut.fi

Ridha Hamila
Qatar University
College of Engineering
P.O.Box 2713, Doha, Qatar
hamila@qu.edu.qa

Abstract — **Next-generation mobile networks will provide users with high data rates, increased mobility and various services. The initial step made in LTE and LTE-A is adopting the OFDM(A) air interface and utilization of dynamic resource allocation techniques maximizing the cell throughput and coverage as part of enhanced radio resource management functionality. Moreover, the control of the network resource division among users is performed by packet scheduler and different scheduling strategies are applied according to traffic scenario. Typically, video traffic (real-time video streaming, mobile TV etc.) requires higher data rates and certain quality of service (QoS) constrains, i.e. packet size, arrival rate, head-of-line (HOL) packet delay, etc. In this article, we propose a flexible and fairness-oriented packet scheduling approach for real-time video delivery, built on advanced QoS-aware multiuser proportional fair (PF) scheduling principle. Special emphasis is put on practical feedback reporting schemes, including the effects of mobile measurements and estimation errors, reporting delays, and feedback quantization and compression. The performance of the overall scheduling and feedback reporting process is investigated in details in terms of cellular system capacity, resource allocation fairness and video traffic QoS guarantees. Experimental results reveal the bounds of real-time video traffic support on LTE downlink.**

Keywords- cellular system performance; radio resource management; packet scheduling; QoS; proportional-fair; channel quality feedback; fairness;

I. INTRODUCTION

The Third Generation Partnership Project (3GPP) organization developed the Long Term Evolution (LTE) standard with support to high data-rates, increased mobility, low latency and packet optimized radio access. LTE uses single-carrier frequency-division multiple access (SC-FDMA) for the uplink (UL) and orthogonal FDMA in downlink (DL). Scalable bandwidth operation, exploitation of diverse MIMO technologies and advanced convergence techniques are some of the key benefits in OFDMA-based developments [1]–[3]. The available spectrum is divided into large number of orthogonal subcarriers forming the basic time-frequency transmission resource - physical resource block (PRB). This allows flexible multiuser access, efficient mitigation of multipath effects, and efficient channel-aware link adaptation and resource allocation. Therefore, increased spectral efficiency and high data rates are achieved [5], [6].

The mobile traffic boost due to increased usage of video applications such as video streaming, mobile TV and multimedia online gaming requires increased system performance and user QoS guarantees. Performance improvements are typically obtained through proper radio resource management (RRM) functionalities and exploiting the available multi-user diversity in both time and frequency as well as spatial domains [6]–[12]. Another important aspect for achieving such performance improvements is obtaining accurate channel feedback from each mobile station (MS) within the serving cell. As a practical example, each MS can measure the effective signal-to-interference-plus-noise -ratio (SINR), per active subcarrier or block of subcarriers, and send back the obtained channel state to the base station (BS) in terms of channel quality information (CQI) reports. Moreover, considering multiple-input multiple-output (MIMO) systems, the stream wise feedback is provided as a codeword containing both MIMO ranking information and CQI measurement. This, in turn, can easily lead to considerable control signaling overhead if not designed and implemented properly. Furthermore, the amount of the feedback information is also subject to different errors and delays, affecting the overall system-level performance and needs to be limited [8], [9]. Most of the performance studies of CQI feedback reduction techniques are based on best-effort traffic [9] or VoIP traffic [13] but streaming video traffic has not been extensively studied so far. Recently, Luo et al. in [14] proposed quality-driven cross-layer optimized video delivery scheduling strategy, while Basukala et al. in [15] demonstrated the performance of a packet scheduler serving video streaming users with delayed and aperiodic CQI reporting schemes.

Clearly, the overall radio system performance in terms of throughput, coverage and fairness, depends heavily on packet scheduling (PS) functionality, being the key ingredient in the RRM process. Most of the literature studies on different multi-user packet scheduling techniques demonstrate that the well-known proportional fair (PF) scheduling principle is the right choice for OFDMA based systems [7], [8]. On the other hand, despite increased throughput and fairness, the PF scheduler cannot guarantee the packet delay constraint for video services by default. Thus, higher system throughput does not necessarily guarantee higher video quality or higher number of video streams and therefore video performance metrics should be taken into account in PS metrics and decisions. Only a few scheduling strategies have considered user's QoS requirements together with practical system and application constrains [16]-[20]. Stemming now from our previous work in advanced PS scheduling developments reported in [7], [19], [20], we extend our studies here to incorporate QoS requirements into scheduling decisions by effectively controlling user fairness and BS's RRM process. Furthermore, we apply different simulation cases for video traffic applications investigating the limits of achieved gains from time-frequency-spatial domain packet scheduling with limited feedback and QoS constrains [21], [22]. The system model used for the performance evaluations of the proposed scheduling methods presented in this paper is according to the 3GPP evaluation criteria [2]. The overall outcomes are measured, e.g., in terms of video capacity which means here the number of supported video streams per cell.

This research has been financially supported by the Finnish Funding Agency for Technology and Innovation (Tekes), under the project "Energy and Cost Efficiency for Wireless Access (ECEWA)", and Qatar University.

The rest of the paper is organized as follows: Section II gives an overview of the packet scheduling process and describes the proposed QoS aware multi-stream PF (QoS-MSPF) scheduling scheme. Section III, in turn, addresses different feedback reporting schemes in the scheduling context. Section IV presents the overall system model and detailed simulation assumptions. The simulation results and analysis are then presented in Section V, while conclusions are drawn in Section VI.

II. Packet Scheduling

Figure 1 illustrates the overall RRM framework. The key entities in it are packet scheduler (PS), link adaptation (LA) and HARQ manager. Located in the BS, the PS functionality consists of selecting the users (UEs) to be scheduled on transmission time interval (TTI) basis and allocating the required resources (PRBs). In more details, the scheduling decision is based on priority metric calculation for individual UEs depending on the selected scheduling strategy. Some of the advanced PF based scheduling techniques require users' CQIs per given TTI and per frequency domain PRB. MIMO functionality, in turn, requires both single-stream and dual-stream CQI feedback by each UE. In addition, PS is interacting with LA entity for choosing the modulation and coding schemes (MCS) for individual PRBs and obtaining information for new transmissions or retransmissions from HARQ manager. BS buffer information is required for verification of keeping with the set packet delay budget.

The proposed QoS-aware multi-stream PF (QoS-MSPF) scheduler is based on widely used two-stage PF approach (see e.g. [9], [20]) with additional QoS enabled guarantees for video traffic. In terms of the actual metric calculations, the proposed QoS-MSPF scheduler uses the following metric:

$$\overline{\gamma}_{i,k,s} = \arg\max_{i}\left\{\delta_i(n)\left(\frac{CQI_{i,k,s}(n)}{CQI_i^{avg}(n)}\right)^{\alpha_1} \times \left(\frac{T_i^{se}(n)}{T_{i,k,s}(n)}\frac{T_i(n)}{T_{tot}(n)}\right)^{-\alpha_2}\right\} \quad (1)$$

which can be understood as an extension of the authors' earlier work in [19]. In the above metric, α_1 and α_2 are scheduler optimization parameters ranging basically from 0 to infinity. In expression (1), $T_i^{se}(n)$ is an estimate of the user throughput if user i is scheduled on sub-frame basis according to [17] and $T_{i,k,s}(n)$ is the estimated achievable throughput of user i at PRB k and stream s. $T_i(n)$ corresponds to average delivered throughput to the UE over the past and T_{tot} is the average delivered throughput (during the recent past) to all users ranked in TD stage served by the BS. $CQI_{i,k,s}$ is the CQI of user i at PRB k and stream s, and $CQI^{avg}{}_i$ is the average CQI of user i calculated by traditional recursive method [23]. The component $\delta_i(n)$ is, in turn, *QoS delay function factor* defined as

$$\delta_i(n) = \frac{\max\{d_i / B_i(n)\}}{d_{\max}} \quad (2)$$

where d_i is the delay of the packet of user i in the transmit buffer $B_i(n)$ at time instant n, and d_{max} is the maximum delay allowed. The scheduling metric in (1)-(2) is essentially composed of three parts affecting the overall scheduling decisions. The first term takes into account the QoS requirements i.e,, packet delay budget. The second ratio is the relative instantaneous quality of the individual users' radio channels over their own average channel qualities. The third ratio is divided into two parts. The first one takes into account the estimated throughputs of individual UE's and the second one the achievable over total throughputs. The power coefficients α_1 and α_2 are additional adjustable parameters.

In more details, in the first stage, which is the time-domain (TD) scheduling step executed within each TTI, UEs are ranked based on the full bandwidth channel state information, QoS parameters and the corresponding throughput calculations. These UEs form Group 1 of the whole Scheduling Candidates Set (SCS). Group 2 consists of users with pending retransmissions and Group 3 is formed by users having packet delay close to the delay budget. Furthermore, discard timer is used so that buffered packets that are already late will not be transmitted over the air interface. Thus, the accumulated delay for subsequent packet is also reduced. In the second stage, the scheduling functionality is expanded in frequency-domain (FD) where the actual PRB allocation takes place. Initially, the needed PRB's for pending retransmissions (on one stream-basis only) signaled through HARQ channels are reserved, forming Group 2 UEs. The remaining PRBs are given to the delay sensitive users from Group 3 and the rest to the first transmission users from Group 1 determined through available buffer information. The actual priority metric in FD/SD stage is evaluated at PRB-level taking into account the available stream-wise channel state information, QoS parameters and the corresponding throughput calculations, as described in (1)-(2).

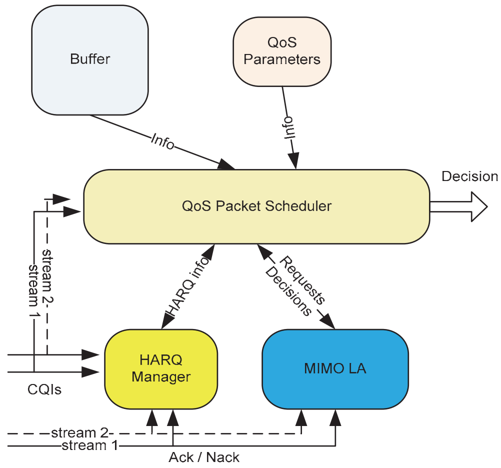


Figure 1. RRM entities and packet scheduling process.

In general, dynamic resource allocation methods increase scheduling flexibility so that enough PRBs will be mapped to the UEs with good channel conditions and more resources will be available for the other SCS users. Consequently, prioritizing Group1 UEs will further decrease packet delays and reduce retransmissions. On the other hand, combined QoS- and channel-aware scheduling can offer performance enhancements at the expense of increased scheduling complexity, in terms of scheduling metric calculations and increased signaling overhead. Keeping these at reasonable levels requires thus some constraints on the scheduling algorithm, so for simplicity we assume here that only one MIMO mode (SU or MU) and fixed modulation and coding scheme (MCS) is allowed per user within one scheduling element. Moreover, we limit the number of users for multiplexing in TD stage to further reduce the signaling overhead and complexity of FD/SD scheduling. Consequently, further decrease in signaling overhead is examined by exploiting different reduced-feedback reporting schemes, explained in the following.

III. Feedback Reporting

The overall reporting process between UE's and BS is illustrated in Figure 2. Within each time window of length t_r, each mobile sends channel quality indicator (CQI) reports to BS, formatted in finite number of bits and possibly compressed. Additional reporting delay of t_d seconds is introduced due to both UE sending time and BS decoding time. In our studies here, the starting point (reference case) is that the CQI reports are quantized SINR measurements across the entire bandwidth (wideband CQI reporting), to take advantage of the time and frequency variations of the radio channels for the different users. Considering MIMO scenario, the reported feedback contains rank adaptation information about the possible multi-stream transmissions options.

Naturally, each report is subject to errors due to imperfect decoding of the received signal. Consequently, the CQI reporting frequency-resolution has a direct impact on the achievable multi-user frequency diversity and thereon to the overall system performance and the efficiency of RRM functionality [8]-[9]. As an example, the CQI block can be formed from two or more consecutive PRBs. The CQI for k-th PRB on stream s is modeled here as:

$$CQI(k,s) = Quantize\{Q;10\log_{10} SINR(k,s) + \delta(k)\} \quad (3)$$

where $\delta(k)$ is a zero mean Gaussian distributed variable with standard deviation 1 dB representing the errors in the SINR measurements and Q is quantization step (in bits) for finite rate reporting. The full CQI and alternative reduced feedback schemes are described and evaluated, as discussed below in more details.

A. Full CQI Reporting

Every UE reports all the measured CQIs across the entire system bandwidth for each TTI. In a general OFDMA radio system, the overall system bandwidth is assumed to be divided into v measurement blocks and with quantizing the CQI values to q bits, the overall full CQI report size is $S_{full} = q \times v$ bits. In case of LTE, with 10 MHz system bandwidth and grouping 2 physical resource blocks into 1 measurement block, it follows that $v = 25$. Assuming further that quantization is carried with $q = 5$ bits, then each UE is sending a CQI word of 25 x 5 = 125 bits for every 1ms (TTI length) on single stream.

B. Best-m CQI Reporting

The best-m scheme is illustrated graphically in Figure 3. Reducing the reporting and feedback signaling information is obtained as follows:

- Select only $m < v$ different CQI measurements and reporting them together with their frequency positions to the serving cell. The evaluation criteria for choosing those m sub-bands for reporting is based on the highest measured SINR values (hence the name best-m).
- Set the unclaimed v - m CQI measurements to the lowest reported CQI value in the BS.

The resulting report size is then given by

$$S_{best-m} = q \times m + \left\lceil \log_2 \left(\frac{v!}{m!(v-m)!} \right) \right\rceil \quad (4)$$

As an example, with $v = 25$, $q = 5$ bits and $m = 10$, it follows that $S_{best\text{-}m} = 72$ bits, while $m = 15$ results to $S_{best\text{-}m} = 97$ bits. In our performance evaluations, we set the values of m={10,15}, which correspond to reported CQI values per stream of 20 and 30 PRBs.

IV. System Simulation Model

A. Video traffic model

Real –time video services are modeled as follows:

- Each frame of video data arrives at a regular interval determined by the number frames per second.
- Each frame is decomposed into a fixed number of slices, each transmitted as a single packet. The size of these packets is modeled to have a truncated Pareto distribution.
- The video encoder introduces encoding delay intervals (modeled by a truncated Pareto distribution) between the packets of a frame.

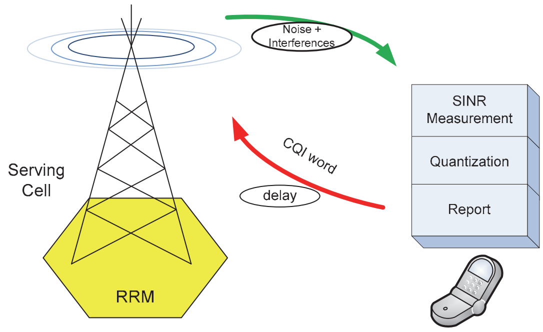


Figure 2. Feedback reporting mechanism.

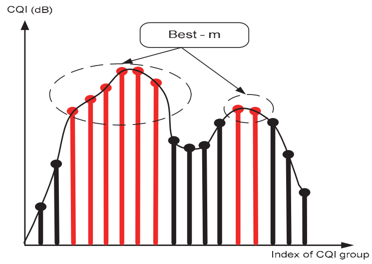


Figure 3. Best-m reporting process.

In our studies reported here, a principal average-rate video-streaming service with 256 kbps constant bit rate (CBR) source video data is assumed. With 256 kbps streaming rate, the corresponding mean packet size with truncated Pareto distribution and mean inter-arrival packet time are 200 bytes and 4 ms [22]. Packet delay budget, as well as discard timer threshold, are set to 20 ms.

B. Video capacity estimation

The main target is to estimate the number of video streaming UE's that can be supported by the system based on deployed scenarios. Following 3GPP evaluation metrology, a user is considered to be in outage if 2% of the video packets for the user are erroneous or discarded due to exceeding delay limit when monitored over the whole video session duration. On the other hand, video capacity is defined as number of supported users per cell without exceeding the system saturation point. Here, the system saturation point is set to 5% of the cell outage level, i.e. to the point where 95% of the users in the cell are satisfied (having maximum of 2% packet loss rate as described above).

C. Simulation environment

Quasistatic system level simulator is used to evaluate the proposed scheduling scheme in video traffic scenarios for LTE downlink. It includes detailed traffic modeling, multiuser packet scheduling and link adaptation including HARQ, following the 3GPP evaluation criteria [2]. The chosen Micro Case 1 deployment scenario consists of 10 MHz system bandwidth divided into 50 physical resource blocks (PRB) containing 600 data subcarriers. In a single simulation run, mobile stations are randomly distributed over standard hexagonal cellular layout with altogether 19 cells each having 3 sectors. Fast fading is updated per TTI according to Typical Urban (20 taps) radio channel, while shadow fading and distance dependent path loss remain constant during the whole simulation. The main simulation parameters and assumptions are summarized in Table I.

The RRM functionalities are controlled by the packet scheduler together with link adaptation (LA) and HARQ entities. Moreover, LA functionality consists of removing CQI imperfections, estimating supported data rates and MCS's, and stabilizing the 1st transmission Block Error Probability (BLEP) to the target range (typically 10-20%). HARQ is based on SAW protocol and a maximum of three re-transmissions is allowed. MIMO functionality requires individual HARQ entry per stream which is also implemented. Link-to-system level mapping is based on the effective SINR mapping (EESM) principle [2].

The actual effective SINR calculations rely on estimated sub-carrier-wise channel gains (obtained using reference symbols) and depend in general also on the assumed receiver topology. Here we assume per-antenna rate control (PARC) MIMO case, i.e. two transmit antennas at the BS and two receive antennas at each UE and the receivers are equipped with LMMSE detectors. The detector structures and SINR modeling for different transmission modes are described in detail below.

1) Single– Stream Single-User Case

Here, only one of the two BS transmit antennas is used to transmit one stream. At individual time instant (time-index dropped here), the received spatial 2x1 signal vector of UE i at sub-carrier c is then $\mathbf{y}_{i,c} = \mathbf{h}_{i,c}x_{i,c} + \mathbf{n}_{i,c} + \mathbf{z}_{i,c}$ where $x_{i,c}$, $\mathbf{h}_{i,c}$, $\mathbf{n}_{i,c}$ and $\mathbf{z}_{i,c}$ denote the transmit symbol, 2x1 channel vector, 2x1 received noise vector and 2x1 inter-cell interference vector, respectively. Then the LMMSE detector $\hat{x}_{i,c} = \mathbf{w}_{i,c}^H \mathbf{y}_{i,c}$ is given by

$$\mathbf{w}_{i,c} = (\mathbf{h}_{i,c}\sigma_{x,i}^2\mathbf{h}_{i,c}^H + \Sigma_{n,i} + \Sigma_{z,i})^{-1}\sigma_{x,i}^2\mathbf{h}_{i,c} \tag{5}$$

where $\sigma_{x,i}^2$, $\Sigma_{n,i}$ and $\Sigma_{z,i}$ denote the transmit power, noise covariance matrix and inter-cell interference covariance matrix, respectively. Now the SINR is given by

$$\gamma_{i,c} = \frac{\left|\mathbf{w}_{i,c}^H\mathbf{h}_{i,c}\right|^2\sigma_{x,i}^2}{\mathbf{w}_{i,c}^H\Sigma_{n,i}\mathbf{w}_{i,c} + \mathbf{w}_{i,c}^H\Sigma_{z,i}\mathbf{w}_{i,c}} \tag{6}$$

Noise is assumed spatially white (diagonal $\Sigma_{n,i}$) while the more detailed modeling of inter-cell interference (structure of $\Sigma_{z,i}$) takes into account the distances and channels from neighboring base stations (for more details, see e.g. [19]).

2) Dual –Stream Single-User Case

In this case, both of the two BS transmit antennas are used for transmission, on one stream per antenna basis. At individual time instant, the received spatial 2x1 signal vector of UE i at sub-carrier c is now given by $\mathbf{y}_{i,c} = \mathbf{H}_{i,c}\mathbf{x}_{i,c} + \mathbf{n}_{i,c} + \mathbf{z}_{i,c}$ where $\mathbf{x}_{i,c}$ and $\mathbf{H}_{i,c} = [\mathbf{h}_{i,c,1}, \mathbf{h}_{i,c,2}]$ denote the 2x1 transmit symbol vector and 2x2 channel matrix, respectively. Now the LMMSE detector $\hat{\mathbf{x}}_{i,c} = \mathbf{W}_{i,c}\mathbf{y}_{i,c}$ is given by

$$\mathbf{W}_{i,c} = \Sigma_{x,i}\,\mathbf{H}_{i,c}^H(\mathbf{H}_{i,c}\,\Sigma_{x,i}\,\mathbf{H}_{i,c}^H + \Sigma_{n,i} + \Sigma_{z,i})^{-1} = \begin{bmatrix}\mathbf{w}_{i,c,1}^H \\ \mathbf{w}_{i,c,2}^H\end{bmatrix} \tag{7}$$

where $\Sigma_{x,i} = diag\{\sigma_{x,i,1}^2, \sigma_{x,i,2}^2\} = diag\{\sigma_{x,i}^2/2, \sigma_{x,i}^2/2\}$ denotes the 2x2 covariance matrix (assumed diagonal) of the transmit symbols. Compared to single-stream case, the overall BS transmit power is now divided between the two antennas, as indicated above. Then the SINR's for the two transmit symbols are given by

$$\begin{aligned}\gamma_{i,c,1} &= \frac{\left|\mathbf{w}_{i,c,1}^H\mathbf{h}_{i,c,1}\right|^2\sigma_{x,i,1}^2}{\left|\mathbf{w}_{i,c,1}^H\mathbf{h}_{i,c,2}\right|^2\sigma_{x,i,2}^2 + \mathbf{w}_{i,c,1}^H\Sigma_{n,i}\mathbf{w}_{i,c,1} + \mathbf{w}_{i,c,1}^H\Sigma_{z,i}\mathbf{w}_{i,c,1}} \\ \gamma_{i,c,2} &= \frac{\left|\mathbf{w}_{i,c,2}^H\mathbf{h}_{i,c,2}\right|^2\sigma_{x,i,2}^2}{\left|\mathbf{w}_{i,c,2}^H\mathbf{h}_{i,c,1}\right|^2\sigma_{x,i,1}^2 + \mathbf{w}_{i,c,2}^H\Sigma_{n,i}\mathbf{w}_{i,c,2} + \mathbf{w}_{i,c,2}^H\Sigma_{z,i}\mathbf{w}_{i,c,2}}\end{aligned} \tag{8}$$

3) Dual-Stream Multi-User Case

In this case, the transmission principle and SINR modeling are similar to subsection above, but the two spatially multiplexed streams belong now to two different UE's, say i and i'. Thus the SINR's in (8) are interpreted accordingly.

V. Simulation results and Analysis

In this section, we present the results obtained from the system simulations using the RRM algorithms described in the paper. The system-level performance is generally measured and evaluated in terms of:

- Video capacity in terms of number of streams
- System throughput
- CDF of the number of users scheduled per TTI
- CDF of the number of PRBs per UE
- CDF of packed delay
- Fairness distribution

In general, the video capacity depends on the video data rates, packet size and the choice of outage criteria. The user outage criteria is defined as 2% of the video packets are erroneous or discarded during the whole simulation. Capacity in turn corresponds to maximum number of supported users not exceeding 5 % cell outage level. The cell throughput is defined as the number of successfully delivered user bits per unit time. Fairness is measured using the Jain's fairness index [24].

The performance of the proposed QoS-MSPF scheduler is compared against reference PF scheduler with a delay dependent component for video traffic support [12] for the video traffic models as discussed in Section IV and evaluated using different power coefficients α_1 and α_2. Here α_2 is fixed to 1 and the used values for α_1 coefficient are defined as $\alpha_1 = \{1,2\}$. Moreover, the power coefficient values are presented as index M, where M1 represents the first couple, i.e., $\alpha_1=1$, $\alpha_2=1$ and M2: $\alpha_1=2$, $\alpha_2=1$.

Figure 4 (left column) illustrates the number of supported video mobile users and average cell throughput for the different QoS-aware schedulers obtained using the quasi-static system simulator. The obtained results with the proposed scheduler scheme are compared with the reference PF scheduler achieving video capacity of 52 users per sector and average sector throughput of 15 Mbps (subfigures (a) and (b)). By using the first term (M1) of the new metric calculation for QoS-MSPF, in combination with full CQI reporting scheme, we achieve video capacity gain in the order of 16% and additional 9% throughput increase. The achieved gain mainly comes from better utilization of the resources due to the scheduler priority metric calculation, as well as the increase in frequency-spatial domain multiplexed users. In the case of full CQI feedback and coefficient α_1 equal to 2 (M2), the new QoS-MSPF scheduler achieves video capacity gains in the order of 18% and throughput increase of 11% compared to reference PF scheduler. Continuing on the evaluation of relative system performance using the proposed scheduler together with the choice of the *Best-m* CQI reporting scheme, we clearly see the relatively similar trend illustrated in Figure 4. In the case of *Best-m* (m=10) and *Best-m* (m=15) reporting schemes presented in Figure 4 (c) and (d), and Figure 4 (e) and (f), we have video capacity and throughput gains of 18-11% and 22-14% when compared to reference PF case.

Further illustrations on the obtainable system performance are presented in Figure 4 (right column) in terms of cumulative distribution function (CDF) of scheduled PRBs per TTI for individual scheduling scheme. The impact of different CQI reporting schemes on packet scheduling functionality results in small decrease of utilizing PRB resources indicated as shift on the left of CDF curve. Noticeably, the proposed scheduling scheme allocates more resources to the users shown in subfigures (b) and (c) with right shift in CDF curve. Clearly, reduced CQI measurement causes the loss of frequency selective information and therefore more PRBs should be allocated to individual users for maintaining the capacity.

Figure 5 (left) shows the CDF of the number of scheduled users per TTI. We can clearly see from the figure that with the proposed scheduling scheme allows more users to be scheduled in each TTI and the achieved capacity gains are due to the increased usage of PRB resources as already concluded above. Figure 5 (middle) shows the CDF of packet delay for different schedulers and applied feedback reporting schemes. The CQI reporting process has a direct impact on the delay performance and therefore a combined approach of HOL and CQI criteria in priority metric calculation will benefit in such scenarios. Figure 5 (right) illustrates the Jain's fairness indexes [24] for scheduling scheme based on the number of supported users per chosen feedback reporting mechanism. The value of each curve corresponds to the used scheduler type (1st values refer to Reference PF scheduler, 2nd values refer to QoS-MSPF – M1 and 3rd values refer to QoS-MSPF – M2). We observe that the reference scheduling scheme has lower fairness indexes for all the simulated cases and the proposed QoS-MSPF scheduler clearly obtains better fairness. Having video traffic model with strict QoS requirements implies increased user fairness and thus fairness indexes above 0,85 (1 refers to complete fairness) are expected.

Table 1. Default simulation parameters

Parameter	Assumption
Carrier Frequency / Bandwidth	2000MHz / 10 MHz
Number of active sub-carriers	600
Sub-carrier spacing	15kHz
Sub-frame duration	0.5 ms
Channel estimation	Ideal
PDP	ITU Typical Urban 20 paths
Min. distance between UE and cell	>= 35 meters – Macro
Max. number of frequency multiplexed UEs	10
UE receiver type	LMMSE
Shadowing standard deviation	8 dB
UE speed	3km/h
Total BS TX power (Ptotal)	46dBm
Fast Fading Model	Jakes Spectrum
CQI reporting schemes	Full CQI, Best -m
CQI log-normal error std.	1 dB
CQI reporting time	5 TTI
CQI delay CQI quantization CQI std error	2 TTIs 1 dB 1 dB
MCS rates	QPSK (1/3, 1/2, 2/3), 16QAM (1/2, 2/3, 4/5), 64QAM (1/2, 2/3, 4/5)
ACK/NACK delay	2ms
Number of SAW channels	6
Max. number of retransmissions	3
HARQ model	Ideal chase combining (CC)
1st transmission BLER target	20%
Scheduler forgetting factor	0.0025
Scheduling schemes used	RPF, QoS -MSPF (proposed)
Simulation duration (one drop)	120 seconds
Number of drops	10

VI. CONCLUSIONS

In this article, we have studied the potential of advanced QoS-aware multi-user packet scheduling algorithms for video traffic support in OFDMA packet radio system context, using UTRAN long term evolution (LTE) type downlink in Macro cell environment as practical example case. New video-optimized multi-stream proportional fair scheduler metric covering time-, frequency- and spatial domains was proposed that takes into account video traffic QoS requirements, instantaneous channel qualities (CQI's) as well as resource allocation fairness. Furthermore, detailed set of performance results with different practical CQI reporting schemes for real-time video support were presented. As a practical low-rate streaming example, more than 60 users can be supported for video traffic with source data rate of 256 kbps over the assumed 10 MHz bandwidth.

REFERENCES

[1] 3GPP Technical Specification Group Radio Access Network Technical Specification 36.300 v8.0.0: Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (EUTRAN); Overall description, march 2007

[2] Physical Layer Aspects for Evolved UTRA", 3GPP Technical Report TR 25.814, ver. 7.1.0, Oct. 2006.

[3] ITU global standard for international mobile telecommunications "IMT-Advanced", IMT-ADV/1-E,Mar. 2008.

[4] WiMAX Forum™, http://www.wimaxforum.org, Aug. 2006.

[5] E. Dahlman et al., *3G Evolution: HSPA and LTE for Mobile Broadband*, Academic Press, 2007.

[6] N. D. Tripathi et al., *Radio Resource management in Cellular Systems*, Springer, 2001.

[7] S. Nonchev and M. Valkama, "Efficient packet scheduling schemes with built-in fairness control for multiantenna packet radio systems," Int. J. Advances in Networks and Services, vol. 2, pp. 182-194, 2009.

[8] Y. Sun et al., "Multi-user scheduling for OFDMA downlink with limited feedback for evolved UTRA", in *Proc. IEEE VTC-2006*, Montreal, Canada, September 2006.

[9] K.I. Pedersen, G. Monghal, I.Z. Kovacs, T.E. Kolding, A. Pokhariyal, F. Frederiksen, P. Mogensen, "Frequency domain scheduling for OFDMA with limited and noisy channel feedback," in *Proc. IEEE VTC-2007* Fall, Baltimore, MD, Sept. 2007, pp. 1792-1796

[10] N. Wei, et al., "Performance of MIMO with frequency domain packet scheduling", in *Proc. IEEE VTC-2007*, Dublin, Ireland, May 2007.

[11] C. Wengerter, J. Ohlhorst, and A.G.E Von Elbwert, "Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA", in *Proc. IEEE VTC-2005*, Stockholm, Sweden, May 2005.

[12] K. Elsayed and A. Khattab, "Channel-aware earliest deadline due fair scheduling for wireless multimedia networks," *Wireless Personal Communications*, vol. 38, pp. 233-252, 2008.

[13] Y. Fan, M. Kuusela, P. Lundén, M. Valkama, "Performance of VoIP on EUTRA Downlink with Limited Channel Feedback", in *Proc. IEEE ISWCS-2008*, Reykjavik, Iceland, October 2008.

[14] L. Haiyan, et al.,"Quality-Driven Cross-Layer Optimazed Video Delivery over LTE", *IEEE Communication magazine*, vol. 48, pp. 102-109, DOI 10.1109/MCOM.2010.5402671.

[15] R.Basukala, et al., "Impact of CQI Feedback Rate/Delay on Scheduling Video Streaming Services in LTE Downlink", in *Proc. International Conf.. Communication Technology ICCT–2010*, Nanjing, China, November 2010.

[16] G. Monghal, et al., "QoS Oriented Time and Frequency Domain Packet Schedulers for The UTRAN Long Term Evolution", in *Proc. VTC2008-Spring*, vol.3, pp. 2532-2536, May 2008.

[17] T.E. Kolding, "QoS-Aware Proportional Fair Packet Scheduling with Required Activity Detection", in *Proc. IEEE VTC2006-Fall*, September 2006

[18] T.-D. Nguyen, Y. Han, A Proportional Fairness Algorithm with QoS Provision in Downlink OFDMA Systems, *IEEE Communication Letters*, vol. 10, No. 11, November 2006.

[19] S. Nonchev and M. Valkama, "A new fairness-oriented packet scheduling scheme with reduced channel feedback for OFDMA packet radio systems", *Int. Journal of Communications, Network and System Sciences (IJCNS),* 2009, vol. 7, pp. 608-618.

[20] S. Nonchev and M. Valkama, “Advanced radio resource management for multi antenna packet radio systems,” *Int. Journal of Wireless & Mobile Networks (IJWMN)*, vol. 2, pp. 1-14, May 2010.
[21] A. Talukdar, M. Cudak, and A. Ghosh, “Streaming Video Capacities of LTE Air-Interface,” in *Proc. IEEE ICC 2010*, Cape Town, South Africa, May 2010.
[22] 3GPP, "Feasibility Study for Orthogonal Frequency Division Multiplexing (OFDM) for UTRAN enhancement," 3GPP TR25.892, 2004.
[23] A. Jalali, R. Padovani, and R. Pankaj, “Data throughput of CDMA-HDR high efficiency-high data rate personal communication wireless system,” in *Proc. IEEE VTC-2000 spring*, Tokyo, May 2000.
[24] D. Chui and R. Jain, “Analysis of the increase and decrease algorithms for congestion avoidance in computer networks”, *Computer Networks and ISDN Systems*, 1989.

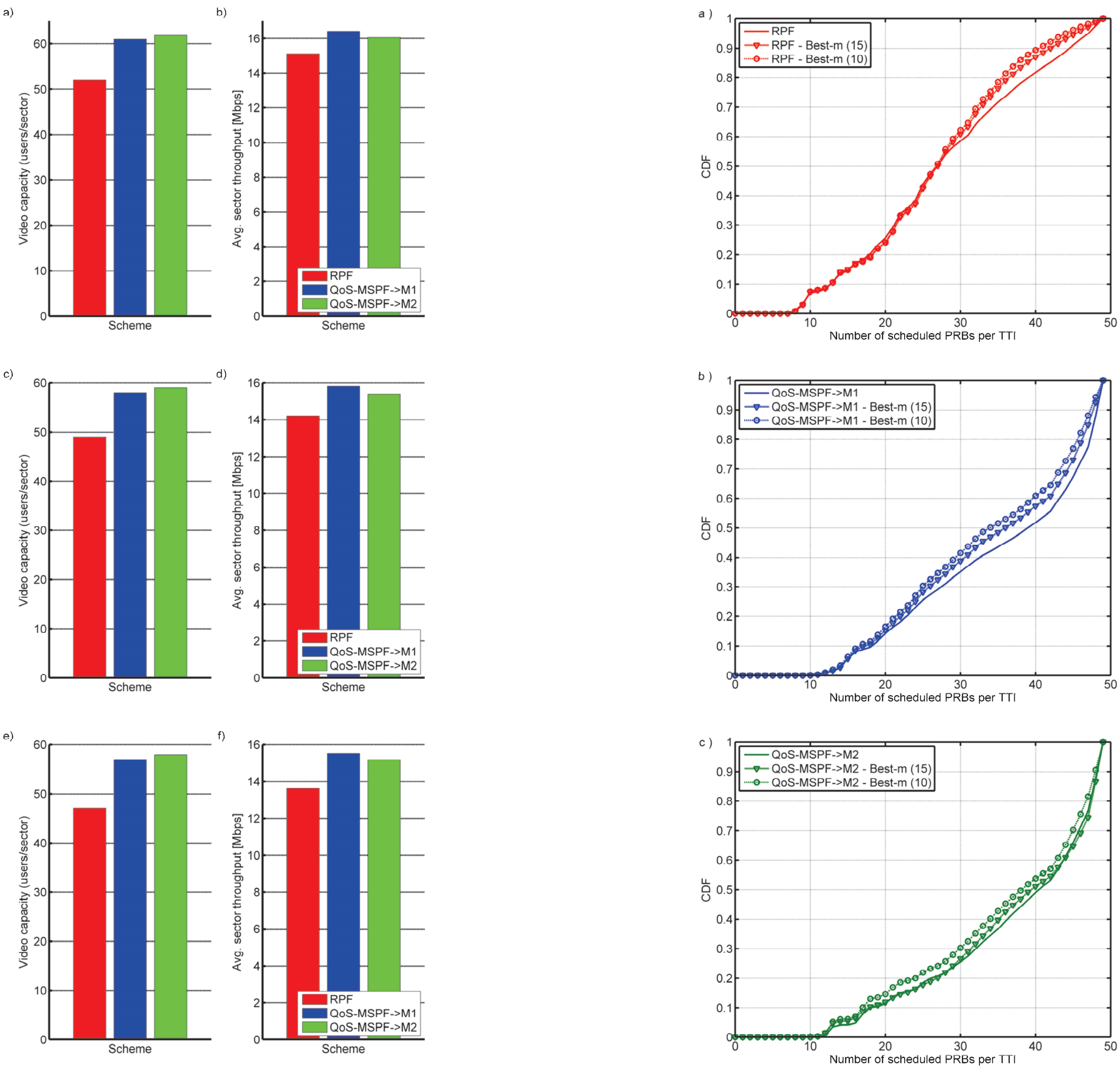


Figure 4. Left column: Video capacity and average sector throughput for different scheduling schemes with *full* CQI feedback (a, b), *Best -m (m=15)* CQI feedback (c, d) and *Best -m (m=10)* CQI feedback (e, f). M1-M2 refer to the proposed scheduler with power coefficient values. Right column: CDF of scheduled PRB per TTI for different CQI reporting schemes based scheduling principle: (a) RPF, (b) QoS-MSPF - M1, (c) QoS-MSPF – M2

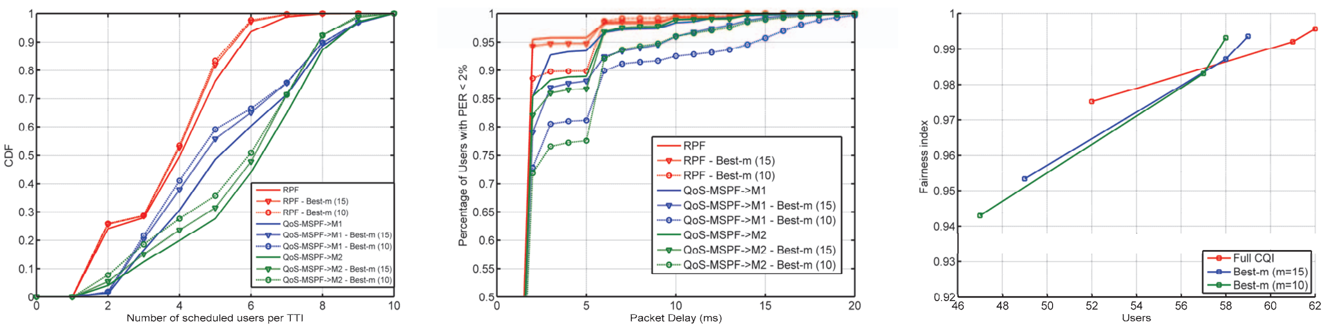


Figure 5. Left column: CDF of number of scheduled users per TTI for different feedback reporting schemes. Middle column: CDF of packet delay for different feedback reporting schemes. Right column: Jain’s fairness indexes of supported video users for different feedback reporting schemes.

Publication 9

S. Nonchev and M. Valkama, “A new fairness-oriented packet scheduling scheme with reduced channel feedback for OFDMA packet radio systems,” *International Journal of Communications, Network and System Sciences* (IJCNS), vol. 2, no.7, pp. 608-618, October 2009.

Int. J. Communications, Network and System Sciences, **2009, 7, 608-618**
doi:10.4236/ijcns.2009.27068 Published Online October 2009 (http://www.SciRP.org/journal/ijcns/).

A New Fairness-Oriented Packet Scheduling Scheme with Reduced Channel Feedback for OFDMA Packet Radio Systems

Stanislav NONCHEV, Mikko VALKAMA
Department of Communications Engineering, Tampere University of Technology, Tampere, Finland
Email: {*stanislav.nonchev, mikko.e.valkama*}*@tut.fi*
Received June 26, 2009; *revised August* 13, 2009; *accepted September* 26, 2009

ABSTRACT

In this paper, we propose a flexible and fairness-oriented packet scheduling approach for 3GPP UTRAN long term evolution (LTE) type packet radio systems, building on the ordinary proportional fair (PF) scheduling principle and channel quality indicator (CQI) feedback. Special emphasis is also put on practical feedback reporting mechanisms, including the effects of mobile measurement and estimation errors, reporting delays, and CQI quantization and compression. The performance of the overall scheduling and feedback reporting process is investigated in details, in terms of cell throughput, coverage and resource allocation fairness, by using extensive quasi-static cellular system simulations in practical OFDMA system environment with frequency reuse of 1. The performance simulations show that by using the proposed modified PF approach, significant coverage improvements in the order of 50% can be obtained at the expense of only 10-15% throughput loss, for all reduced feedback reporting schemes. This reflects highly improved fairness in the radio resource management (RRM) compared to other existing schedulers, without essentially compromising the cell capacity. Furthermore, we demonstrate the improved functionality increase in radio resource management for UE's utilizing multi-antenna diversity receivers.

Keywords: Radio Resource Management, Packet Scheduling, Proportional-Fair, Channel Quality Feedback, Throughput, Fairness

1. Introduction

Development of new radio interface technologies for beyond 3G cellular radio systems with support to high data rates, low latency and packet-optimised radio access has led to the use of OFDM/OFDMA. One good example of such developments is e.g. the UTRAN long term evolution (LTE), being currently standardized by 3GPP [1–3]. In general, performance improvements over the existing radio systems are basically obtained through proper deployment of fast link adaptation and new packet scheduling algorithms, exploiting the available multi-user diversity in both time and frequency domains [4–6]. On the other hand, achieving such performance improvements typically requires relatively accurate channel state feedback in terms of CQI reports from mobile stations (MS) to the base station (BS) [6–12]. As a practical example, each mobile station can measure the effective signal-to-interference-plus-noise-ratio (SINR), per active subcarrier or block of subcarriers, and send back the obtained channel state to the base station for downlink radio resource management. This, in turn, can easily lead to considerable control signalling overhead if not designed and implemented properly. Thus in general, the amount of the feedback information needs to be limited and is also subject to different errors and delays, affecting the overall system-level performance. Another important aspect in scheduling and resource allocation process is fairness, implying that also users with less favourable channel conditions should anyway be given some reasonable access to the radio spectrum [4–6,13–18]. This is especially important in serving users at, e.g., cell edges in cellular networks.

In this paper, we address the packet scheduling and channel state reporting tasks in OFDMA-based cellular packet radio systems. Stemming from ordinary proportional fair (PF) scheduling principle, a modified PF scheduler is first proposed having great flexibility to tune the exact scheduling characteristics in terms of capacity, coverage and fairness. More specifically, the proposed scheduler can offer greatly improved fairness among the users in a cell, measured in terms of coverage and other established fairness measures, like Jain's index [19], without essentially compromising the overall cell capacity. This is verified using extensive quasi-static cellular system simulations, conforming to the current LTE downlink specifications [1–3]. In the performance studies, different realistic CQI reporting schemes are also addressed and incorporated in the system simulations.

In general, the research on novel packet scheduling algorithms and channel state reporting schemes has been very active in the recent years, see e.g. [8,10,11,13–18] and the references therein. Using [13–17] as starting points for LTE type packet radio systems, it has been reported that frequency domain packet scheduling (FDPS) algorithms are always a compromise between the overall cell throughput and resource fairness among users. Here we propose a modified proportional fair algorithm, which in general offers an attractive balance between cell throughput, coverage and user fairness. Compared to plain frequency domain scheduling, we extend the studies by deploying both time domain and frequency domain scheduling steps, together with proper metrics, that as a whole can more efficiently utilise the provided yet limited feedback information from all the user equipments (UEs). Furthermore, we apply different realistic CQI reporting schemes to thoroughly investigate the limits of achieved performance gains from enhanced scheduling. The cellular system model used for the performance evaluations is fully conforming to the 3GPP evaluation criteria [1–3]. The overall outcomes are measured in terms of average *cell throughput*, *coverage* and *fairness index*.

The rest of the paper is organised as follows: Section 2 reviews the reference proportional fair scheduler and proposes then a modified PF scheduling scheme. Section 3, in turn, addresses different feedback reporting schemes in the scheduling context. Section 4 presents then the overall system model and simulation assumptions, and the simulation results and analysis are presented in Section 5. Finally, the conclusions are drawn in Section 6.

2. Scheduling Process

2.1. General Scheduling and Link Adaptation Principles

In general, the task of a packet scheduler (PS) is to select the most suitable users to access the available radio spectrum at any given time window, in order to optimize the system performance in terms of 1) throughput, 2) resource fairness, and/or 3) delay [4–6]. Joint optimization of all the above features is generally known very difficult. In *fast* packet scheduling, new scheduling decisions are basically taken in each transmission time interval (TTI), which in LTE is 1ms.

To efficiently utilize the limited radio resources, the scheduler should consider the current state of the channel when selecting the user to be scheduled, by utilizing e.g. the ACK/NACK signalling information and CQI reports [4–6,8,10,11,14]. Depending on the selected CQI reporting scheme, the accuracy and resolution of the channel quality information can easily differ considerably. In OFDMA based radio systems, like LTE, the CQI information is not necessarily available for all the individual subcarriers but more likely for certain groups of subcarriers only [12,20–22]. In general, the channel state information is also used by link adaptation (LA) mechanisms to select proper modulation and coding scheme (MCS) for each scheduled mobile, and thereon to ensure that the individual link qualities conform to the corresponding target settings. This is typically measured in terms of block error rate (BLER) for the first transmission. Hybrid ARQ (HARQ) mechanisms are then commonly used to provide the necessary buffer information and transmission format for pending retransmissions [4–6,16]. A principal block-diagram of the overall RRM flow is given in Figure 1.

As a practical example of the available spectral resources, in the 10 MHz system bandwidth case of LTE [1–3], there are 50 physical resource blocks (PRB's or sub-bands), each consisting of 12 sub-carriers with subcarrier spacing of 15 kHz. This sets the basic resolution in frequency domain (FD) UE multiplexing (scheduling), i.e., the allocated individual UE bandwidths are multiples of the PRB bandwidth.

2.2. Ordinary Proportional Fair (PF) Scheduler

The well-known proportional fair scheduler [13,16] works in two steps: 1) time domain (TD) PF step and 2) frequency domain (FD) PF step. Such simplified sche-

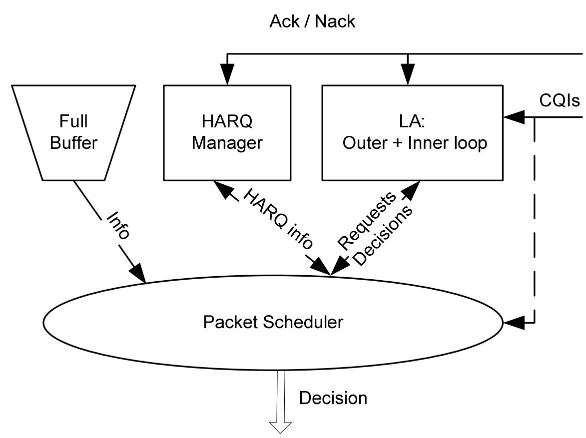


Figure 1. Principal RRM block diagram.

duling principle is beneficial from the complexity point of view, since the FD step considers a reduced number of UEs for frequency multiplexing in each TTI [17]. Thus in the first part, inside each TTI n, all the UE's are ranked according to the following priority metric

$$\gamma_i^{td}(n) = \frac{R_i(n)}{T_i(n)} \tag{1}$$

In above, the UE index $i = 1, 2, \ldots, I_{\text{TOT}}$, $R_i(n)$ denotes the estimated throughput to the UE i over the *full bandwidth* (provided by link adaptation unit) [13,16], and $T_i(n)$ in turn is the corresponding average delivered throughput to the UE i during the recent past and can be obtained, e.g., recursively by

$$T_i(n) = \left(1 - \frac{1}{t_c}\right) T_i(n-1) + \frac{1}{t_c} R_i'(n-1) \tag{2}$$

In (2), t_c controls the averaging window length over which the average delivered throughput is calculated and $R'_i(n\text{-}1)$ denotes the actually realized throughput to the UE i at the previous TTI.

In the next step, out of this ranked list of UE's, the first I_{BUFF} ($< I_{\text{TOT}}$) UE's with highest priority metric are picked to the actual frequency domain multiplexing or scheduling stage. In the following, this subset is called scheduling candidate set (SCS), and is denoted by $\Omega(n)$. Then, for each physical resource block $k = 1, 2, \ldots, K_{\text{TOT}}$, and for each i belonging to the SCS, the following final scheduling metric of the form

$$\gamma_{i,k}^{fd}(n) = \frac{R_{i,k}(n)}{T_i(n)} \tag{3}$$

is evaluated where now $R_{i,k}(n)$ denotes the *estimated* throughput to the UE i for the k-th PRB (provided by LA unit again), and $T_i(n)$ is again the corresponding average throughput *delivered* to the UE i during the recent past given in (2). Finally, the access to each PRB resource is granted for the particular user with the highest metric for the corresponding PRB.

2.3. Proposed Modified PF (MPF) Scheduler

In order to obtain a scheduler with yet increased fairness in the resource allocation, we proceed as follows. First the time domain priority metric is modified as

$$\overline{\gamma}_i^{td}(n) = CQI_i(n) \left(\frac{T_i(n)}{T_{tot}(n)} \right)^{-1} \tag{4}$$

where $CQI_i(n)$ denotes the full bandwidth channel quality report for UE i at TTI n and $T_i(n)$ is as defined in (2). $T_{tot}(n)$, in turn, denotes the averaged throughput over the past and over the scheduled users and can be calculated by

$$\begin{aligned} T_{tot}(n) &= \left(1 - \frac{1}{t_c}\right) T_{tot}(n-1) \\ &\quad + \frac{1}{t_c} \frac{1}{I_{BUFF}} \sum_{i \in \Omega(n-1)} R_i'(n-1) \end{aligned} \tag{5}$$

In (5), $R'_i(n\text{-}1)$ denotes the actual *delivered* throughput for UE i at the previous TTI.

Similar to the ordinary PF scheduler described in Subsection 2.2, this modified metric in (4) is used to rank the UE's inside each TTI, and the I_{BUFF} ($< I_{\text{TOT}}$) UE's with highest priority metric form a SCS. $\Omega(n)$ for the actual frequency domain resource allocation. Since estimated throughput in the link adaptation stage is based on reported CQI values, we assume that the substitution in (4) has the same weight in priority calculation. For mapping the users of the SCS into PRB's, the following modified frequency domain metric is then proposed:

$$\overline{\gamma}_{i,k}^{fd}(n) = \left(\frac{CQI_{i,k}(n)}{CQI_i^{avg}(n)} \right)^{s_1} \left(\frac{T_i(n)}{T_{tot}(n)} \right)^{-s_2} \tag{6}$$

Here s_1 and s_2 are adjustable parameters, and $CQI_{i,k}(n)$ is the channel quality report of user i for sub-band k at TTI n while $CQI_i^{avg}(n)$ is the corresponding average CQI over the past and over the sub-bands, and can be calculated using

$$\begin{aligned} CQI_i^{avg}(n) &= \left(1 - \frac{1}{t_c}\right) CQI_i^{avg}(n-1) \\ &\quad + \frac{1}{t_c} \frac{1}{K_{TOT}} \sum_{k=1}^{K_{TOT}} CQI_{i,k}(n) \end{aligned} \tag{7}$$

The access to each PRB resource is then granted for the particular user with the highest metric in (6) for the corresponding PRB.

Considering the re-transmissions, re-transmitting users are simply considered as additional users in the time domain scheduling part (step 1), and if qualified to the frequency domain SCS, the re-transmission users are given an additional priority to reserve exactly the same sub-bands used for the corresponding original transmissions. Even though this does not take the exact sub-band condition into account at re-transmission stage, the practical implementation is simplified, in terms of control signalling, and re-transmissions anyway always benefit from the HARQ combining gain [6].

Intuitively, the proposed scheduling metrics in (4) and (6) are composed of two elements, affecting the overall scheduling decisions. The first dimension measures the relative instantaneous quality of the individual user's radio channels against their own average channel qualities while the second dimension is related to measuring the achievable throughput of individual UE's against the corresponding average throughput of scheduled users. Consequently, by understanding the power coefficients s_1

and s_2 as additional adjustable parameters, the exact scheduler statistics can be tuned and controlled to obtain a desired balance between the throughput and fairness. This will be demonstrated in Section 5.

3. Feedback Reporting Process

The overall reporting process between UE's and BS is illustrated in Figure 2. Within each time window of length t_r, each mobile sends channel quality indicator (CQI) reports to BS, formatted and possibly compressed, with a reporting delay of t_d seconds [6,8,10,11]. Each report is naturally subject to errors due to imperfect decoding of the received signal. In general, the CQI reporting frequency-resolution has a direct impact on the achievable multi-user frequency diversity and thereon to the overall system performance and the efficiency of radio resource management (RRM), as described in general e.g. in [11]. In our studies here, the starting point (reference case) is that the CQI reports are quantized SINR measurements across the entire bandwidth (wideband CQI reporting), to take advantage of the time and frequency variations of the radio channels for the different users. Then also alternative reduced feedback schemes are described and evaluated, as discussed below.

3.1. Full CQI Reporting

In a general OFDMA radio system, the overall system bandwidth is assumed to be divided into v CQI measurement blocks. Then quantizing the CQI values to q bits, the overall full CQI report size is

$$S_{full} = q \times v \qquad \textbf{(8)}$$

bits which is reported by every UE for each TTI [1–3,11]. In case of LTE, with 10 MHz system bandwidth and grouping 2 physical resource blocks into 1 measurement

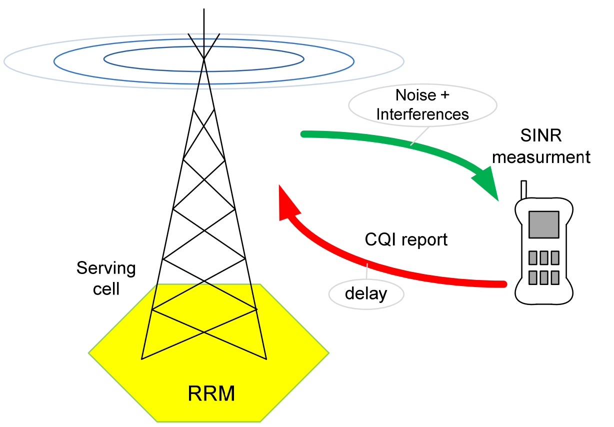


Figure 2. Reporting mechanism between UE and BS.

block, it follows that $v = 25$. Assuming further that quantization is carried with $q = 5$ bits, then each UE is sending 25x5 = 125 bits for every 1ms (TTI length).

3.2. Best-m CQI Reporting

One simple approach to reduce the reporting and feedback signalling is obtained as follows. The method is based on selecting only $m < v$ different CQI measurements and reporting them together with their frequency positions to the serving cell [8,11]. We assume here that the evaluation criteria for choosing those m sub-bands for reporting is based on the highest SINR values (hence the name best-m). The resulting report size in bits is then given by

$$S_{best-m} = q \times m + \left\lceil \log_2 \left(\frac{v!}{m!(v-m)!} \right) \right\rceil \qquad \textbf{(9)}$$

As an example, with $v = 25$, $q = 5$ bits and $m = 10$, it follows that $S_{best-m} = 72$ bits, while $S_{full} = 125$ bits. Furthermore, on the scheduler side, we assume that the PRBs which are not reported by the UE are allocated a CQI value equal to the lowest reported one.

3.3. Threshold Based CQI Reporting

This reporting scheme is a further simplification and relies on providing information on only the average CQI value above certain threshold together with the corresponding location (sub-band index) information. First the highest CQI value is identified within the full bandwidth, which sets an upper bound of the used threshold window. All CQI values within the threshold window are then averaged and only this information is sent to the BS together with the corresponding sub-band indexes. On the scheduler side, the missing CQI values can then be treated, e.g., as the reported averaged CQI value minus a given dB offset (e.g. 5 dB, the exact number is again a design parameter). The number of bits needed for reporting is therefore only

$$S_{threshold} = q + v \qquad \textbf{(10)}$$

As an example, with $v = 25$ and $q = 5$ bits (as above), it follows that $S_{threshold} = 30$ bits, while $S_{best-m} = 72$ bits and $S_{full} = 125$ bits. The threshold-based scheme is illustrated graphically in Figure 3 [10].

4. System Simulation Model and Assumptions

In order to evaluate the system-level performance of the proposed scheduling scheme in a practical OFDMA-based cellular system context, a comprehensive quasi-static system simulator for LTE downlink has been developed,

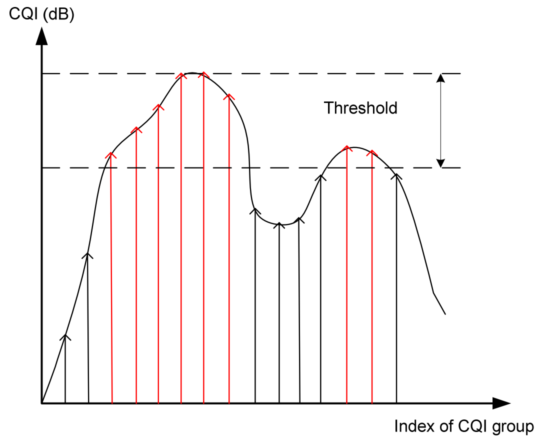


Figure 3. Basic principle of threshold-based CQI reporting.

conforming to the specifications in [1–3]. In the overall simulation flow, mobile stations are first randomly dropped or positioned over each sector and cell. Then based on the individual distances between the mobiles and the serving base station, the path losses for individual links are directly determined, while the actual fading characteristics of the radio channels depend on the assumed mobility and power delay profile. In updating the fading statistics, the time resolution in our simulator is set to one TTI (1ms). In general, a standard hexagonal cellular layout is utilized with altogether 19 cell sites each having 3 sectors. In the performance evaluations, statistics are collected only from the central cell site while the others simply act as sources of inter-cell interference.

As a practical example case, the 10 MHz LTE system bandwidth mode [1–3] is assumed. The main simulation parameters and assumptions are generally summarized in Table 1 for the so-called Macro cell case 1, following again the LTE working assumptions. As illustrated in Figure 1, the RRM functionalities are controlled by the packet scheduler and also link adaptation and HARQ mechanisms are modelled and implemented, as described in Table 1. As a practical example, the maximum number of simultaneously multiplexed users (I_{BUFF}) is set to 10 here. In general, we assume that the BS transmission power is equally distributed among all PRB's. In the basic simulations, 20 UE's are uniformly dropped within each sector and experience inter-cell interferences from the surrounding cells, in addition to path loss and fading. The UE velocity equals 3km/h, and the typical urban (TU) channel model standardized by ITU is assumed in modelling the power-delay spread of the radio channels. Infinite buffer traffic model is applied in the simulations, i.e. every user has data to transmit (when scheduled) for the entire duration of a simulation cycle. The length of a single simulation run is set to 5 seconds which is then repeated for 10 times to collect reliable statistics.

In general, every UE has an individual HARQ entity,

Table 1. Basic simulation parameters.

Parameter	Assumption
Cellular Layout	Hexagonal grid, 19 cell sites, 3 sectors per site
Inter-site distance	500 m
Carrier Frequency / Bandwidth	2000 MHz / 10 MHz
Number of active sub-carriers	600
Sub-carrier spacing	15 kHz
Sub-frame duration	0.5 ms
Channel estimation	Ideal
PDP	ITU Typical Urban 20 paths
Minimum distance between UE and cell	>= 35 meters
Average number of UE's per sector	20
Max. number of frequency multiplexed UEs (I_{BUFF})	10
UE receiver type	2-Rx MRC, 2-Rx IRC
Shadowing standard deviation	8 dB
UE speed	3km/h
Total BS TX power (P_{total})	46dBm
Traffic model	Full Buffer
Fast Fading Model	Jakes Spectrum
CQI reporting schemes	Full CQI Best-m (with m=10) Threshold based (with 5dB threshold)
CQI log-normal error std.	1 dB
CQI reporting time	5 TTI
CQI delay	2 TTIs
CQI quantization	1 dB
CQI std error	1 dB
MCS rates	QPSK (1/3, 1/2, 2/3), 16QAM (1/2, 2/3, 4/5), 64QAM (1/2, 2/3, 4/5)
ACK/NACK delay	2ms
Number of SAW channels	6
Maximum number of retransmisions	3
HARQ model	Ideal chase combining (CC)
1st transmission BLER target	20%
Scheduler forgetting factor	0.002
Scheduling schemes used	Ordinary PF (for reference) Modified PF (proposed
Simulation duration (one drop)	5 seconds
Number of drops	10

which operates the physical layer re-transmission functionalities. It is based on the stop-and-wait (SAW) protocol and for simplicity, the number of entries per UE is fixed to six. HARQ retransmissions are always transmitted with the same MCS and on the same PRB's (if scheduled in TD step) as the first transmissions. The supported modulation schemes are QPSK, 16QAM and 64QAM with variable rates for the encoder as shown in Table 1.

Link adaptation handles the received UE reports containing the channel quality information for the whole or sub-set of PRB's as described in Section 3. The implemented link adaptation mechanism consists of two separate elements – the inner loop (ILLA) and outer loop (OLLA) LA's – and are used for removing CQI imperfections and estimating supported data rates and MCS. As a practical example, it is assumed that the CQI report errors are log-normal distributed with 1dB standard deviation.

The actual effective SINR calculations rely on estimated subcarrier-wise channel gains (obtained using reference symbols in practice) and depend in general also on the assumed receiver topology. Here we assume the single-input-multiple-output (SIMO) diversity reception case, i.e. a single BS transmit antenna and multiple UE receiver antennas. Considering now an individual UE i, the SINR per active sub-carrier c at TTI n, denoted here by $\xi_{i,c}(n)$, is calculated according to

$$\xi_{i,c} = \frac{\left|\mathbf{w}_{i,c}^{H}\mathbf{h}_{i,c}\right|^{2}\sigma_{sig,i}^{2}}{\mathbf{w}_{i,c}^{H}\sum_{noise}\mathbf{w}_{i,c} + \mathbf{w}_{i,c}^{H}\sum_{int,i}\mathbf{w}_{i,c}} \quad \textbf{(11)}$$

where the time index n is dropped for notational simplicity. Here h_{ic} is an N_{RX} x 1 vector of the user i complex channel gains at subcarrier c from BS to N_{RX} receiver antennas and w_{ic} is the corresponding N_{RX} x 1 spatial filter used to combine the signals of different receiver antennas (more details below). $s_{sig,i}^{2}$, in turn, denotes the received nominal signal power per antenna while $\mathbf{\Sigma}_{noise}$ and $\mathbf{\Sigma}_{int,i}$ are the covariance matrices of the received (spatial) noise and interference vectors. The superscript $(.)^{H}$ denotes conjugate transpose. The noise covariance is assumed diagonal ($\mathbf{\Sigma}_{noise} = s_{noise}^{2}\mathbf{I}$) and independent of the user index i. The interference modeling, on the other hand, takes into account the interference from neighboring cells. Assuming a total of L_{int} interference sources, with corresponding path gain vectors $\mathbf{g}_{l,i,c}$, the overall interference covariance at receiving UE i is given by

$$\mathrm{S}_{int,i} = \sum_{l=1}^{L_{int}} \sigma_{int,l,i}^{2}\mathbf{g}_{l,i,c}\mathbf{g}_{l,i,c}^{H} \quad \textbf{(12)}$$

where $s_{int,l,i}^{2}$, denotes the received nominal interferer power per antenna and per interference source (l).

Concerning the actual UE receiver topologies (spatial filters), both maximum ratio combining (MRC) and interference rejection combining (IRC) receivers are deployed in the simulations. These are given by (see, e.g., [6] and the references therein)

$$\mathbf{w}_{i,c}^{MRC} = \frac{\mathbf{h}_{i,c}}{\left\|\mathbf{h}_{i,c}\right\|^{2}} \quad \textbf{(13)}$$

and

$$\mathbf{w}_{i,c}^{IRC} = \frac{\sum_{tot,i}^{-1}\mathbf{h}_{i,c}}{\mathbf{h}_{i,c}^{H}\sum_{tot,i}^{-1}\mathbf{h}_{i,c}} \quad \textbf{(14)}$$

where $\mathbf{\Sigma}_{tot,i}$ denotes the total noise plus interference covariance, i.e., $\mathbf{\Sigma}_{tot,i} = s_{noise}^{2}\mathbf{I} + \mathbf{\Sigma}_{int,i}$.

Using the above modeling and the selected UE receiver type, the effective SINR values are then calculated through exponential effective SINR mapping (EESM), as described in [1–3], for link-to-system level mapping purposes.

5. Results

In this section, we present the system-level performance results obtained using the previously described quasi-static radio system simulator. Both ordinary PF and modified (proposed) PF packet schedulers are used, together with the three different CQI reporting schemes. The system-level performance is generally measured and evaluated in terms of:

- Throughput statistics – the cumulative distribution function (CDF) of the total number of successfully delivered bits per time unit. Measured at both individual UE level as well as overall cell level.
- Coverage – the experienced data rate per UE at the 95% coverage probability (5% UE throughput CDF level).
- Jain's fairness index [19].

In addition to Jain's index, also the coverage and slope of the throughput CDF reflect the fairness of the scheduling algorithms.

With the proposed modified PF scheduler, different example values for the power coefficients s_1 and s_2 are used as shown in Table 2. To focus mostly on the role of the channel quality reporting, s_2 is fixed here to 1 and the effects of using different values for s_1 are then demonstrated. This way the impact of the different CQI reporting schemes is seen more clearly. For the cases of *Best–m* and *Threshold* based CQI reporting schemes, we fix the value of m equal to 10 and threshold to 5 dB, respectively. Similar example values have also been used by other authors in the literature earlier, see e.g. [11]. Complete performance statistics are gathered for both dual antenna MRC and dual antenna IRC UE receiver cases.

Table 2. Different power coefficient combinations used to evaluate the performance of the proposed scheduler.

Coefficient	Value						
s_1	1	2	4	6	8	10	20
s_2	1	1	1	1	1	1	1

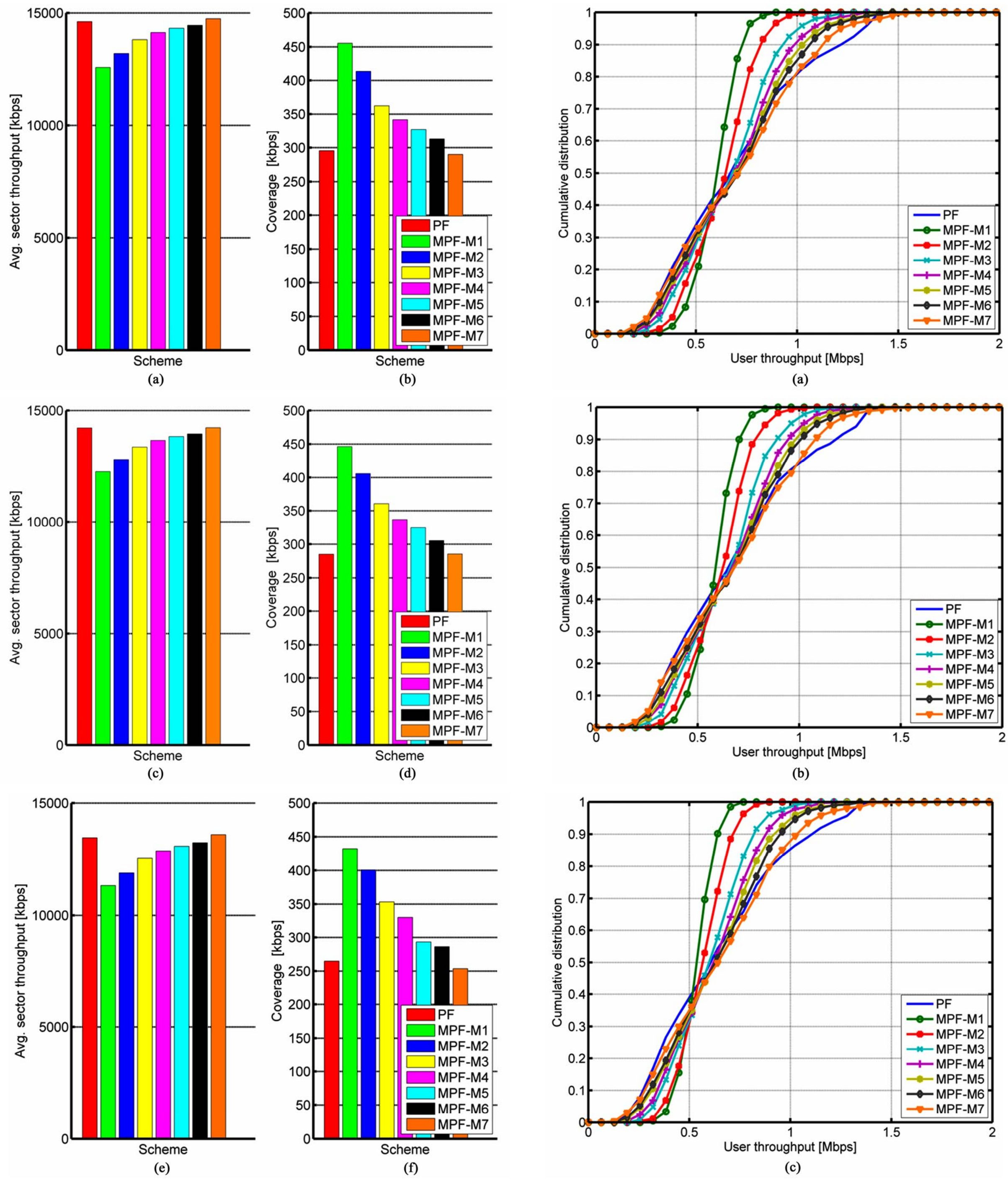


Figure 4. Left column: Average sector throughput and coverage for different scheduling schemes and assuming dual-antenna MRC UE receiver type with *full* CQI feedback (a, b), *Best -m* CQI feedback (c, d) and *Threshold* based CQI feedback (e, f). M1-M7 refer to the modified PF scheduler with power coefficient values as given in Table 2 (M1: s_1=1, s_2=1, etc.). Right column: CDF's of individual UE throughputs for different scheduling schemes and assuming dual-antenna MRC UE receiver type with *full* CQI feedback (a), *Best -m* CQI feedback (b) and *Threshold* based CQI feedback (c).

5.1. Dual Antenna MRC UE Receiver Case

Figure 4 (left column) illustrates the average sector throughput and coverage for the different schedulers, assuming dual antenna maximum ratio combining (MRC) UE receiver type. The power coefficient values from Table 2 are presented as index M, where M1 represents the first couple (s_1=1, s_2=1), etc, for the metric calculation of the modified PF scheduler. The used reference scheduler is the ordinary proportional fair approach. In the first coefficient case (M1), in combination with full CQI reporting scheme, we achieve coverage gain in the order of 50% at the expense of only 15% throughput loss as shown in Figure 4 (a) and (b). This sets the basic reference for comparisons in the other cases. In the case of best-m and threshold based reporting schemes presented in and (d), and Figure 4 (e) and (f), we have coverage increases by 57% and 63% with throughput losses of 16% and 19%, correspondingly.

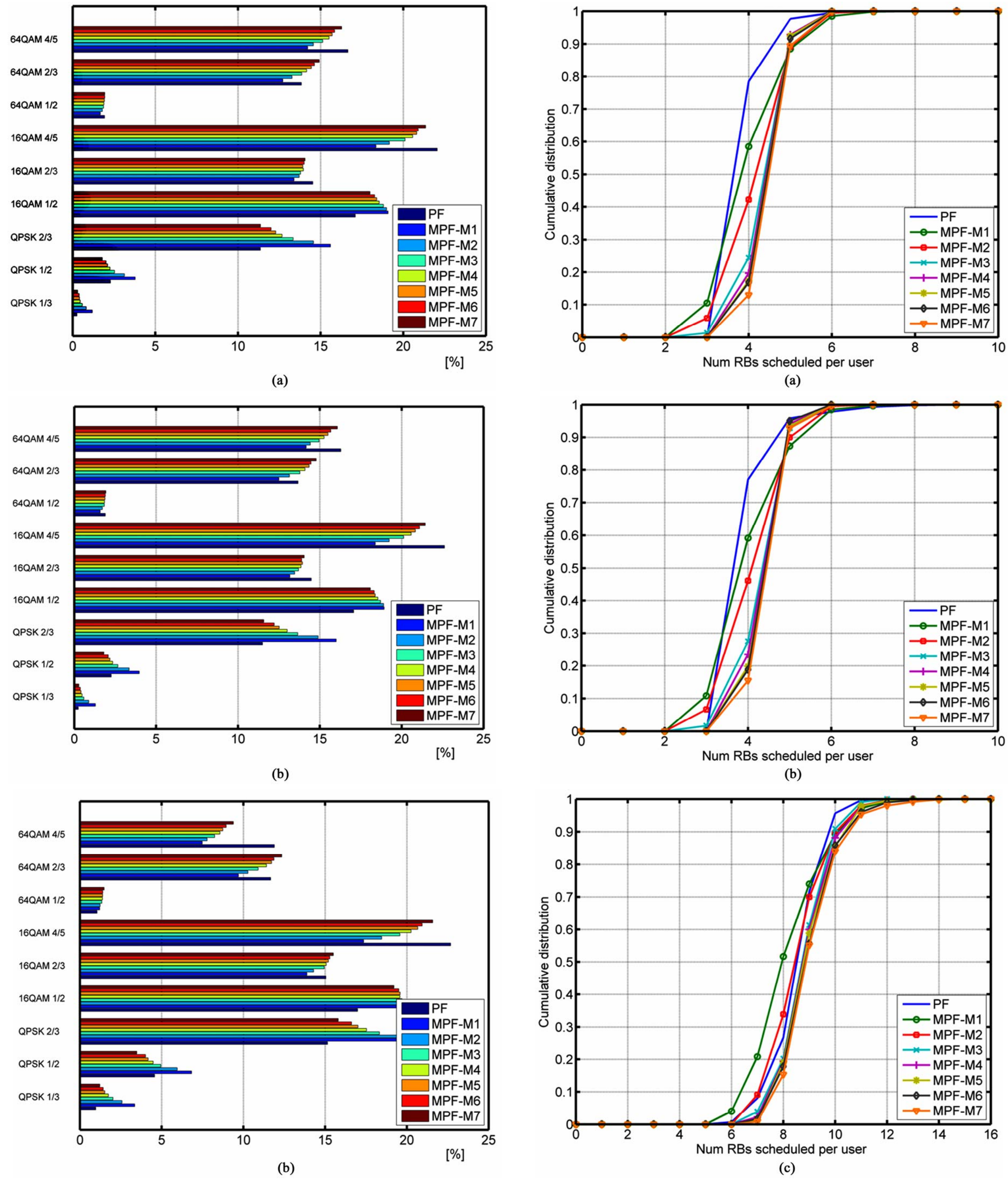


Figure 5. **<u>Left column:</u> MCS distributions [%] for different scheduling principles with (a) *Full* CQI reporting, (b) *Best-m* CQI reporting, and (c) *Threshold* based CQI reporting assuming dual-antenna MRC UE receiver. <u>Right column:</u> CDF's of scheduled PRB's per user for different schedulers with (a) *Full* CQI reporting, (b) *Best-m* CQI reporting, and (c) *Threshold* based CQI reporting assuming dual-antenna MRC UE receiver.**

Continuing on the evaluation of relative system performance using the modified PF scheduler, we clearly see a trade-off between average cell throughput and coverage for different power coefficient cases. The remaining power coefficient values shown in Table 2 are used for tuning the overall system behaviour together with the choice of the CQI reporting scheme. In the case of full CQI feedback and coefficient s_1 varying between 2 and 10 (M2–M6) the cell throughput loss is decreased to around 1%, while the coverage gain is reduced to around 6%. Similar behaviour is observed for the other feedback reporting schemes as well. The exact percentage values for the coverage gains and throughput losses are stated in Table 3 in the end.

Further illustrations on the obtainable system performance are presented in Figure 4 (right column) in terms of the statistics of individual UE data rates for the applied simulation scenarios. The slope of the CDF reflects generally the fairness of the algorithms. Therefore we aim to achieve steeper slope corresponding to algorithm fairness. This type of slope change behavior can clearly be established for each simulation scenario. Clearly, at 5% (coverage) point of the CDF curves, corresponding to users typically situated at the cell edges, we observe significant data rate increases indicated by shift to the right for all CQI feedback schemes when the coefficient s_1 is changed in the proposed metric. This indicates improved overall cell coverage at the expense of slight total throughput loss.

Figure 5 (left column) shows the modulation and coding scheme (MCS) distributions for different schedulers and with applied feedback reporting schemes, still assuming the case of 2 antenna MRC UE receiver type. The negligible decrease in higher order modulation usage (less than 3%) leads to the increase in the lower (more robust) ones for improving the cell coverage. In all the simulated cases, the MCS distribution behaviour has a relatively similar trend following the choice of the power coefficients in the proposed packet scheduling. In general, the use of higher-order modulations is affected mostly in the most coarse CQI feedback (threshold based) case while the other two reporting schemes behave fairly similarly.

Similarly, Figure 5 (right column) illustrates the CDF's of scheduled PRB's per UE for the different scheduler scenarios and reporting schemes. Clearly, the modified PF provides better resource allocation in the *full* and *best-m* feedback cases. Considering the 50% probability point for the resource allocation, and taking the case of M1, we have about 5% gain, while in case of M2 the gain is raised to 15% compared to ordinary PF. The average obtained improvement for the rest of the cases is about 33%. In the case of *threshold*-based feedback, the resource allocation is not as efficient, and even a small reduction in the RB allocation is observed with small power coefficients, compared to the reference PF scheduler. Starting from M3, the improvement is anyway noticeable and the achieved gain is about 20%.

Table 3. Obtained performance statistics compared to ordinary PF scheduler with different CQI reporting schemes and different power coefficients (M1-M7) for the proposed scheduler. Dual-antenna MRC UE receiver case.

	Coverage Gain [%]			Throughput Loss [%]		
	full	best-m	threshold	full	best-m	threshold
M1	54	57	63	16	16	19
M2	40	42	51	10	10	12
M3	23	26	33	6	6	7
M4	16	18	25	3	4	5
M5	11	14	11	2	3	3
M6	6	7	8	1	2	2
M7	-2	0	-4	0	0	0

Table 4. Obtained performance statistics compared to ordinary PF scheduler with different CQI reporting schemes and different power coefficients (M1-M7) for the proposed scheduler. Dual-antenna IRC UE receiver case.

	Coverage Gain [%]			Throughput Loss [%]		
	full	best-m	threshold	full	best-m	threshold
M1	56	58	64	15	15	18
M2	43	46	48	9	9	11
M3	26	30	32	6	6	8
M4	17	20	24	4	4	5
M5	10	12	13	2	3	3
M6	8	10	8	2	2	2
M7	-1	1	1	0	1	0

5.2. Dual Antenna IRC UE Receiver Case

Next similar performance statistics are obtained for dual antenna interference rejection combining (IRC) UE receiver case. Starting from the primary case M1, with full CQI, we obtain a 13% loss in throughput and 57% coverage improvement. For the reduced feedback reporting schemes – best-m and threshold based – we have 13% and 15% throughput losses and 58% and 62% coverage gains, respectively. Furthermore, resource allocation gains for full CQI feedback and best-m are 7% for M1 and 17% for M2 correspondingly. The average obtained improvement for the rest of the cases is about 34%. *Threshold* based reporting scheme leads to decrease of 12% for M1 and 7% for M2, and roughly 14% increase for the rest of simulated cases. The exact percentage

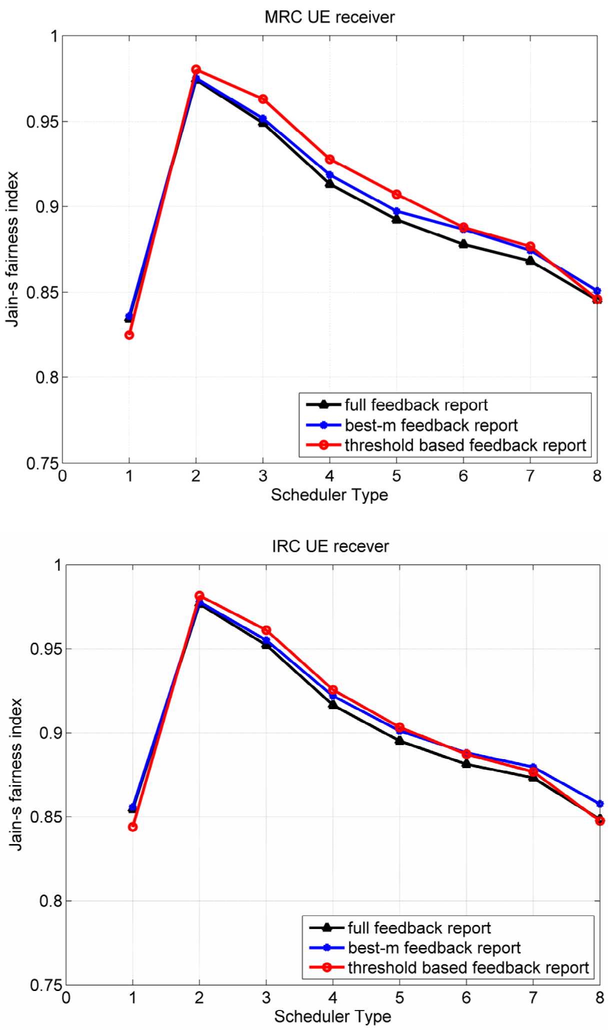


Figure 6. Jain's fairness index per feedback reporting scheme for dual-antenna MRC UE receiver case (up) and dual-antenna IRC UE receiver case (down). Scheduler type 1 means ordinary PF, while 2-8 means proposed modified PF with power coefficients as described in Table 2.

read from the figures are again stated in table format in Table 4 in the end.

5.3. Fairness Index

Figure 6 illustrates the Jain's fairness index per scheduler for the applied feedback reporting schemes, calculated over all the I_{TOT} = 20 UE's using the truly realized UE throughputs at each TTI and over all the simulation runs. The value on the x-axis corresponds to the used scheduler type, where 1 refers to the reference PF scheduler and 2-8 refer to the proposed modified PF schedulers with different power coefficients. The Jain's fairness index defined in [19] is generally in the range of [0…1], where the value of 1 corresponds to all users having the same amount of resources (maximum fairness). Clearly, the fairness distribution with the proposed modified PF scheduler outperforms the used reference PF scheduler for both UE receiver types. The received fairness gains are in range of 2%-17% for the MRC receiver case, and 1%-14% for the IRC receiver case, respectively.

6. Conclusions

In this article, we have studied the potential of advanced packet scheduling principles in OFDMA type radio system context, using UTRAN long term evolution (LTE) as a practical example system scenario. A modified proportional fair scheduler taking both the instantaneous channel qualities (CQI's) as well as resource allocation fairness into account was proposed. Also different practical CQI reporting schemes were discussed, and used in the system level performance evaluations of the proposed scheduler. All the performance evaluations were carried out with a comprehensive quasi-static system level simulator, conforming fully to the current LTE working assumptions. Also different UE receiver types were demonstrated in the performance assessments. In general, the achieved throughput and coverage gains were assessed against more traditional ordinary proportional fair scheduling. In the case of fixed coverage requirements and based on the optimal parameter choice for CQI reporting schemes, the proposed scheduling metric calculations based on UE channel feedback offers better control over the ratio between the achievable cell/UE throughput and coverage increase. As a practical example, even with limited CQI feedback, the cell coverage can be increased significantly (more than 30%) by allowing a small decrease (in the order of only 5-10%) in the cell throughput. This is seen to give great flexibility to the overall RRM process and optimization.

7. Acknowledgments

Fruitful discussions with Markku Kuusela, Nokia Devices, Helsinki, Finland, and Dr. Toni Huovinen, Tampere University of Technology, Tampere, Finland, are greatly acknowledged.

8. References

[1] 3GPP RAN Technical Specification Group, "E-UTRA/ E-UTRAN Overall description, stage 2," Technical Report TR 36.300, ver. 9.0.0, June 2009.

[2] 3GPP RAN Technical Specification Group, "E-UTRA/ LTE physical layer—General description," Technical Report TR 36.201, ver. 8.3.0, March 2009.

[3] 3GPP RAN Technical Specification Group, "Physical layer aspects for evolved UTRA," Technical Report TR 25.814, ver. 7.1.0, Oct. 2006.

[4] N. D. Tripathi, *et al*., "Radio resource management in cellular systems," Springer, 2001.

[5] H. Holma and A. Toskala, Eds., "HSDPA/HSUPA for UMTS–High speed radio access for mobile communications," Wiley, 2006.

[6] E. Dahlman, *et al*., "3G evolution: HSPA and LTE for mobile broadband," Academic Press, 2007.

[7] S. Yoon, C. Suh, Y. Cho, and D. Park, "Orthogonal frequency division multiple access with an aggregated subchannel structure and statistical channel quality measurements," in Proc. IEEE Vehicular Technology Conference (VTC'04 Fall), Los Angeles, CA, September 2004.

[8] Y. Sun, *et al*., "Multi-user scheduling for OFDMA downlink with limited feedback for evolved UTRA," in Proc. IEEE Vehicular Technology Conference (VTC'06 Fall), Montreal, Canada, September 2006.

[9] I. Toufik and H. Kim, "MIMO-OFDMA opportunistic beamforming with partial channel state information," in Proc. IEEE International Conference on Communications, Instanbul, Turkey, pp. 5389–5394, June 2006.

[10] T. E. Kolding, F. Frederiksen, and A. Pokhariyal, "Low-bandwidth channel quality indication for OFDMA frequency domain packet scheduling," in Proc. ISWCS'06, Spain, September 2006.

[11] K. I. Pedersen, G. Monghal, I. Z. Kovacs, T. E. Kolding, A. Pokhariyal, F. Frederiksen, and P. Mogensen, "Frequency domain scheduling for OFDMA with limited and noisy channel feedback," in Proc. IEEE Vehicular Technology Conference (VTC'07 Fall), Baltimore, MD, pp. 1792–1796, Sept. 2007.

[12] P. Svedman, D. Hammarwall, and B. Ottersten, "Subcarrier SNR estimation at the transmitter for reduced feedback OFDMA," in Proc. European Signal Processing Conf., Florence, Italy, September 2006.

[13] C. Wengerter, J. Ohlhorst, and A. G. E Von Elbwert," Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA," in Proc. IEEE Vehicular Technology Conference (VTC' 05 Spring), Stockholm, Sweden, May 2005.

[14] S. Nonchev, J. Venäläinen, and M. Valkama, "New frequency domain packet scheduling schemes for UTRAN LTE Downlink," in Proc. ICT Mobile Summit, Stockholm, Sweden, June 2008.

[15] S. Nonchev and M. Valkama, "Efficient packet scheduling schemes for multiantenna packet radio downlink," in Proc. Fifth Advanced Int. Conf. Telecommunications (AICT'09), Venice, Italy, May 2009.

[16] T. E. Kolding, "Link and system performance aspects of proportional fair scheduling in WCDMA/HSDPA," in Proc. IEEE Vehicular Technology Conference (VTC'03 Fall), Orlando, FL, pp. 1717–1723, Oct. 2003.

[17] A. Pokhariyal, K. I. Pedersen, G. Monghal, I. Z. Kovacs, C. Rosa, T. E. Kolding, and P. E. Mogensen, "HARQ aware frequency domain packet scheduler with different degrees of fairness for the UTRAN long term evolution," in Proc. IEEE Vehicular Technology Conference (VTC' 07 Spring), Dublin, Ireland, April 2007, pp. 2761–2765.

[18] A. Pokhariyal, T. E. Kolding, and P. E. Mogensen, "Performance of downlink frequency domain packet scheduling for the UTRAN long term evolution," in Proc. IEEE Personal, Indoor and Mobile Radio Communications Conference (PIMRC'06), Helsinki, Finland, Sept. 2006.

[19] D. Chui and R. Jain, "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks," Computer Networks and ISDN Systems, 1989.

[20] P. Svedman, S. K. Wilson, L. J. Cimini, and B. Ottersten, "A simplified opportunistic feedback and scheduling scheme for OFDMA," in Proc. IEEE Vehicular Technology Conference (VTC'04 Spring), pp. 1878–1882, May 2004.

[21] P. Svedman, L. J. Cimini, and B. Ottersten, "Using unclaimed sub-carriers in opportunistic OFDMA systems," in Proc. IEEE Vehicular Technology Conference (VTC' 06 Fall), Montreal, Canada, September 2006.

[22] S. Sanayei, A. Nosratinia, and N. Aldhahir, "Opportunistic dynamic sub-channel allocation in multiuser OFDM networks with limited feedback," in IEEE Proc. Inform. Theory Workshop, San Antonio, TX, pp. 182–186, October 2004.

[23] P. Morgensen, *et al*., "LTE capacity compared to the Shannon bound," in Proc. IEEE Vehicular Technology Conference (VTC'07 Spring), Dublin, Ireland, April 2007.

Publication 10

S. Nonchev and M. Valkama, “Efficient packet scheduling schemes with built-in fairness control for multiantenna packet radio systems,” *International Journal on Advances in Networks and Services (IJANS)*, vol.2, no. 2&3, pp, 182-194, 2009.

Efficient Packet Scheduling Schemes with Built-in Fairness Control for Multiantenna Packet Radio Systems

Stanislav Nonchev and Mikko Valkama
Tampere University of Technology
Department of Communications Engineering
P.O. Box 553, FI-33101, Tampere, FINLAND
Email: stanislav.nonchev@tut.fi, mikko.e.valkama@tut.fi

Abstract* – In this article, we propose fairness-oriented dual stream multiple-input multiple-output (MIMO) packet scheduling schemes with efficient utilization of channel quality indicator (CQI) feedback for emerging multiantenna packet radio systems. In general, multiuser multiantenna transmission schemes allow users to be scheduled on different parallel streams on the same time-frequency resource. Based on that, implementations of more intelligent scheduling schemes that are aware of the instantaneous state of the radio channel require utilization of time, frequency and spatial domain resources in an efficient manner. Stemming from the earlier advanced proportional fair (PF) scheduler studies, we extend the developments to dual stream MIMO packet radios with fairness-oriented scheduling metric and practical feedback reporting mechanisms, including the effects of mobile measurement and estimation errors, reporting delays, and CQI quantization and compression. Furthermore, we investigate the resulting fairness distribution among users together with the achievable radio system performance in terms of throughput and coverage, by simulating practical OFDMA cellular system environment with MIMO functionality in Micro and Macro cell scenarios. As a concrete example, we demonstrate that by using the proposed fairness-oriented multiuser scheduling schemes, significant coverage improvements in the order of 40% can be obtained at the expense of only 16% throughput loss for all feedback reporting schemes. Furthermore, the user fairness is also greatly increased, by more than 30%, when measured using Jain's fairness index.

***Keywords* - radio resource management; packet scheduling; proportional fair; channel quality feedback; fairness; coverage; throughput; multiantenna**

I. INTRODUCTION

The advancement of new radio technologies for beyond third generation (3G) cellular systems continues steadily. This includes, e.g., third generation partnership project (3GPP) long term evolution (LTE) [2], worldwide interoperability for microwave access (WiMAX) [3] and the work in various research projects, like WINNER [4]. Some common elements in most of these developments are orthogonal frequency division multiple access (OFDMA) based air interface, operating bandwidths of at least 10-20 MHz, and the exploitation of multiple-input multiple-output (MIMO) techniques and advanced channel-aware packet scheduling principles [2]. MIMO in terms of *spatial multiplexing* (SM), possibly combined with pre-coding, is considered as one core physical layer technology towards increased link spectral efficiencies compared to existing radio systems. In addition, it also provides the packet scheduler (PS) with an extra degree of freedom (spatial domain), by offering a possibility to multiplex multiple data streams of one or more users on the same physical time-frequency resource. The two principal concepts widely analyzed in literature (see, e.g., [5]-[6] and the references therein) in this context are single user (SU) and multiuser (MU) MIMO. The SU-MIMO approach allows only the streams of one individual UE to be scheduled at the same time-frequency resource block (RB), while MU-MIMO provides additional flexibility so that streams of multiple users can be scheduled over the same time-frequency RB. Assuming relatively accurate channel state feedback in terms of channel quality indicator (CQI) reports from mobile stations (MS) to base station (BS), together with fast link adaptation mechanisms, advanced channel-aware packet scheduling schemes have major impact on the system-level performance optimization in terms of, e.g., throughput and coverage. Practical CQI reporting mechanisms in this context are described, e.g., in [7]-[13]. Another important feature of multiuser radio systems related to scheduling, in addition to throughput and coverage, is fairness, implying that also users with less favorable channel conditions should anyway be given some reasonable access to the radio spectrum. This is especially important in serving users at, e.g., cell edges in cellular networks.

Recently, multiantenna oriented packet scheduling principles have started to be investigated in the literature, see, e.g., [14]-[19]. New scheduling algorithms have been proposed and their performance been evaluated in different simulator environments in [1], [16]-[19]. In this article, we concentrate on the proportional fair (PF) principle, which in general offers an attractive balance between cell throughput and user fairness, and extend it with spatial domain functionality for the needs of SU- and MU-MIMO operation. More specifically, we extend our earlier studies in [1], [18]-[19] on algorithm development by deploying SM functionality to frequency domain (FD) PF scheduling schemes that can efficiently utilize the provided feedback information from all the user equipments (UEs), in terms of quantized CQI reports. A new packet scheduler is proposed, called MIMO modified PF (MMPF), with built-in fairness control in its scheduling metric, which is shown to improve the system performance considerably, in terms of cell-edge coverage and overall scheduling fairness, when compared to existing reference schedulers. For practicality, all the performance evaluations are carried out in 3GPP LTE system context, covering both Micro and Macro Cell scenarios, and conforming to the 3GPP evaluation criteria [2]. The used radio system performance measures are cell throughput distribution, average throughput, cell-edge coverage and Jain's fairness index [21].

In general, while the increased flexibility of channel-aware scheduling can offer great performance enhancements, compared to fixed resource allocation, it also has some practical disadvantages. This includes, e.g., relatively higher scheduling complexity, in terms of scheduling metric calculations and increased signaling overhead to facilitate CQI reporting. Keeping these at reasonable levels requires thus some constraints on the scheduling algorithm, so for simplicity we assume here that only one MIMO mode (SU or MU) and fixed modulation and coding scheme (MCS) is allowed per user within one time-frequency scheduling element (RB). For simplicity, we also assume that the BS as well as all the UE's are equipped with 2 antennas (dual antenna TX and RX).

The rest of the article is organised as follows: Section II describes the MIMO channel-aware scheduling principles and the proposed fairness-oriented scheduling scheme. Section III, in turn, gives an overview of different feedback reporting schemes in packet scheduling context. The overall radio system model and simulation assumptions are then presented and discussed in Section IV. The corresponding simulation results and analysis are presented in Section V, while the conclusions are drawn in Section VI.

II. CHANNEL-AWARE MIMO SCHEDULING

In general, the task of the packet scheduler is to select the most suitable users to access the overall radio resources at any given time window, based on some selected priority metric calculations. Typically the scheduler also interacts with other radio resource management (RRM) units such as *link adaptation* (LA) and *Hybrid ARQ* (HARQ) manager as shown in Figure 1. The scheduling decision is based on received users' signaling information in terms of acknowledgements (ACK/NACK) and channel state information (CQI reports) per given transmission time interval (TTI) [9] and per frequency domain physical resource block (PRB). More specifically, in this article in multiantenna radio system context, we assume that both single stream and dual stream CQIs are reported to BS by each UE. Depending on the selected CQI reporting scheme, the accuracy and resolution of the channel quality information can then easily differ considerably [7], [8], [11], [12], as will be explained in Section III. Moreover, the CQI information is not necessarily available for all the individual subcarriers but more likely for certain groups of subcarriers only [13], [22], [23]. Based on this information, the BS scheduler then decides whether the particular time-frequency resource is used for (i) transmitting only one stream to a specific UE, (ii) two streams for a specific UE (SU-MIMO) or (iii) two streams to two different UEs (MU-MIMO, one stream per UE).

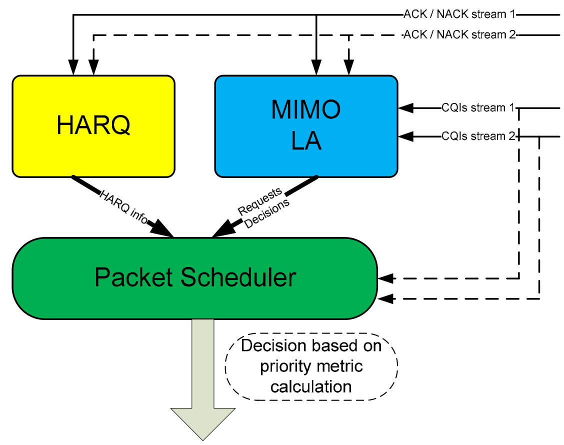


Figure 1: Joint time- and frequency-domain scheduling process.

A. Proposed Scheduler at Principal Level

In general, we use the well-known two-step PF approach with extended functionality in extra spatial dimension to enable MIMO operation [1], [5], [18]-[19]. The first step is time-domain (TD) scheduling in which the scheduler selects a sub-group of users in each TTI, called scheduling candidate set, based on the full bandwidth channel state information. More specifically, the TD step selects those I_{BUFF} UE's (out of the total number of UE's, say I_{TOT}) whose total instantaneous throughput, per TTI, calculated over the full bandwidth is highest [20]. In this stage, we also take the different spatial multiplexing possibilities (single stream, dual stream SU, dual stream MU) into account, in calculating all the possible full bandwidth reference throughputs.

The second step is then frequency-domain (FD) / spatial-domain (SD) scheduling in which the scheduler first reserves the needed PRB's for pending re-transmissions (on one stream-basis only for simplicity) and the rest available PRB's are allocated to the selected UE's of the scheduling candidate set obtained from the TD step. The actual metric in FD/SD allocation is based on the PRB-level and stream-wise channel state information, and the corresponding throughput calculations, as will be explained in more details below.

B. Exact Scheduling Metrics

Here we describe the actual scheduling metrics used in ranking users in the TD scheduling step as well as mapping the users to FD/SD resources in the second step. First a multistream extension of "ordinary" PF is described in sub-section II.B.1, used as a reference in the performance simulations, and then the actual proposed modified metric with increased fairness-control is described in sub-section II.B.2.

1) Multistream Proportional Fair:

For the PF scheduler, scheduling decision per TTI is based on the following priority metric

$$\gamma_{i,k,s} = \arg\max_i \left\{ \frac{R_{i,k,s}(n)}{T_i(n)} \right\} \tag{1}$$

in which $R_{i,k,s}(n)$ is the estimated instantaneous throughput of user i at sub-band k on stream s for the time instant (TTI) n (calculated based on the CQI reports through, e.g., EESM mapping [1]). $T_i(n)$, in turn, is the average delivered throughput to the UE i during the recent past and is calculated by

$$T_i(n) = \left(1 - \frac{1}{t_c}\right) T_i(n-1) + \frac{1}{t_c} R_i(n-1) \tag{2}$$

Here t_c controls the averaging window length over which the average delivered throughput is calculated [11]-[15] and $R_i(n\text{-}1)$ is the actually delivered throughput to user i at previous TTI n-1, calculated over all sub-bands k and possible streams s. In general, $1/t_c$ is also called the forgetting factor.

Considering the previous TD and FD/SD steps described earlier in Section II.A, the above metrics are used as follows:

a) TD: Metric (1) is evaluated over the full bandwidth and for different stream options to rank the I_{TOT} UE's. Out of these, $I_{BUFF} < I_{TOT}$ UE's with highest metric are picked to the following FD/SD stage. In the following, this subset is called scheduling candidate set (SCS), and is denoted by $\Omega(n)$.

b) FD/SD: The access to individual PRB and stream(s) is granted for the user(s) belonging to the above SCS with the highest metric (1) evaluated for the particular PRB and stream at hand.

2) Modified Multistream Proportional Fair (MMPF):

Stemming from the earlier work in [1], [11], the following modified multistream PF metric is proposed:

$$\bar{\gamma}_{i,k,s} = \arg\max_i \left\{ \left(\frac{CQI_{i,k,s}(n)}{CQI_i^{ave}(n)} \right)^{\alpha_1} \left(\frac{T_i(n)}{T_{tot}(n)} \right)^{-\alpha_2} \right\} \tag{3}$$

Here, α_1 and α_2 are scheduler optimization parameters ranging basically from 0 to infinity, $CQI_{i,k,s}$ is the CQI of user i at PRB k and stream s, and CQI^{ave}_i is the average CQI of user i calculated using

$$\begin{aligned} CQI_i^{avg}(n) &= \left(1 - \frac{1}{t_c}\right) CQI_i^{avg}(n-1) \\ &+ \frac{1}{t_c} \frac{1}{K_{TOT}} \frac{1}{S_{TOT}} \sum_{s=1}^{S_{TOT}} \sum_{k=1}^{K_{TOT}} CQI_{i,k,s}(n) \end{aligned} \tag{4}$$

In above, K_{TOT} is the total number of available PRB's while S_{TOT} denotes the maximum number of streams which is here two (max two streams). In (3), $T_{tot}(n)$ is the average delivered throughput (during the recent past) to all I_{BUFF} users served by the BS and is calculated as

$$\begin{aligned} &T_{tot}(n) = \\ &\left(1 - \frac{1}{t_c}\right) T_{tot}(n-1) + \frac{1}{t_c} \frac{1}{I_{BUFF}} \sum_{i \in \Omega(n-1)} T_i(n-1) \end{aligned} \tag{5}$$

Intuitively, the proposed scheduling metric in (3) is composed of two elements affecting the overall scheduling decisions. The first dimension measures in a stream-wise manner the relative instantaneous quality of the individual user's radio channels against their own average channel qualities while the second dimension is related to measuring the achievable throughput of individual UE's against the corresponding average throughput of scheduled users. In this way, and by understanding the power coefficients α_1 and α_2 as additional adjustable parameters, the exact scheduler statistics can be tuned and controlled to obtain a desired balance between the throughput and fairness. This will be demonstrated later using radio system simulations. Considering finally the actual TD and FD/SD steps described at general level in Section II.A, the same approach as in sub-section II.B.1 is deployed but the metrics (1)-(2) are of course here replaced by the metrics in (3)-(5).

III. Feedback Reporting Process

The overall channel state reporting process between UE's and BS is illustrated in Figure 2. Within each time window of length t_r, each mobile sends channel quality indicator (CQI) reports to BS, formatted and possibly compressed, with a reporting delay of t_d seconds [7] , [8], [11], [12]. Each report is naturally subject to errors due to imperfect decoding of the received signal. In general, the CQI reporting frequency-resolution has a direct impact on the achievable multiuser frequency diversity and thereon to the overall system performance and the efficiency of radio resource management, as described in general, e.g., in [12]. In our studies here, the starting point (reference case) is that the CQI reports are quantized SINR measurements across the entire bandwidth (wideband CQI reporting), to take advantage of the time and frequency variations of the radio channels for the different users. Then also alternative reduced feedback schemes are described and evaluated, as discussed below.

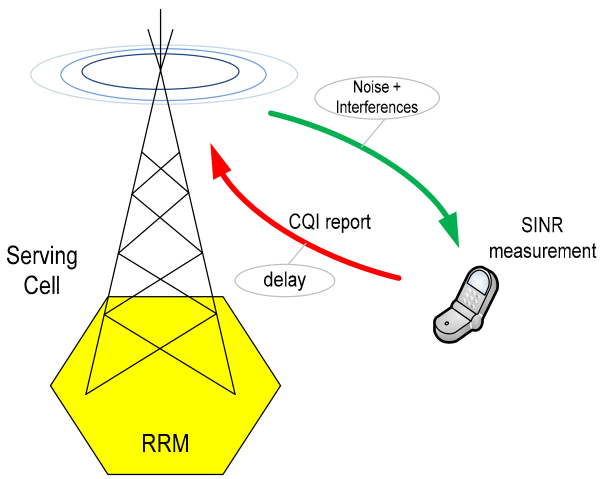


Figure 2: Reporting mechanism between UE and BS.

A. Full CQI Reporting

In a general OFDMA radio system, the overall system bandwidth is assumed to be divided into v CQI measurement blocks. Then quantizing the CQI values to say q bits, the overall full CQI report size is

$$S_{full} = q \times v \tag{6}$$

bits which is reported by every UE for each TTI [2]-[4], [12]. In case of LTE, with 10 MHz system bandwidth and grouping 2 physical resource blocks into 1 measurement block, it follows that $v = 25$. Assuming further that quantization is carried with $q = 5$ bits, then each UE is sending 25x5 = 125 bits for every 1ms (TTI length).

B. Best-m CQI Reporting

One simple approach to reduce the reporting and feedback signaling is obtained as follows. The method is based on selecting only $m < v$ different CQI measurements and reporting them together with their frequency positions to the serving cell [9], [12]. We assume here that the evaluation criteria for choosing those m sub-bands for reporting is based on the highest SINR values (hence the name *Best-m*). The resulting report size in bits is then given by

$$S_{best-m} = q \times m + \left\lceil \log_2 \left(\frac{v!}{m!(v-m)!} \right) \right\rceil \tag{7}$$

As an example, with $v = 25$, $q = 5$ bits and $m = 10$, it follows that $S_{best\text{-}m} = 72$ bits, while $S_{full} = 125$ bits. Furthermore, on the scheduler side, we assume that the PRBs which are not reported by the UE are allocated a CQI value equal to the lowest reported one.

C. Threshold based CQI Reporting

This reporting scheme is a further simplification and relies on providing information on only the average CQI value above certain threshold together with the corresponding location (sub-band index) information. First the highest CQI value is identified within the full bandwidth, which sets an upper bound of the used threshold window. All CQI values within the threshold window are then averaged and only this information is sent to the BS together with the corresponding sub-band indexes. On the scheduler side, the missing CQI values can then be treated, e.g., as the reported averaged CQI value minus a given dB offset (e.g., 5 dB; the exact number is again a design parameter). The number of bits needed for reporting is therefore only

$$S_{threshold} = q + v \tag{8}$$

As an example, with $v = 25$ and $q = 5$ bits (as above), it follows that $S_{threshold} = 30$ bits, while $S_{best\text{-}m} = 72$ bits and $S_{full} = 125$ bits. The threshold-based scheme is illustrated graphically in Figure 3 [11].

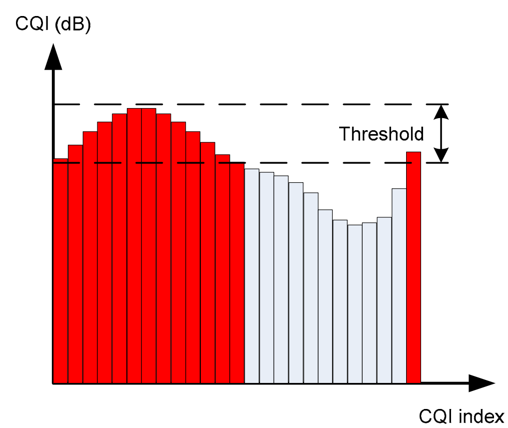


Figure 3: Basic principle of threshold-based CQI reporting.

IV. Quasi-Static Radio System Simulator

A. Basic Features

System-level performance of the proposed scheduling scheme is evaluated based on a quasi-static radio system simulator for LTE downlink, providing traffic modeling, multiuser packet scheduling and link adaptation [2]. As a practical example, the 10 MHz system bandwidth case of LTE is assumed, meaning that there are 50 physical resource blocks each consisting of 12 sub-carriers with sub-carrier spacing of 15 kHz. This sets also the basic resolution in FD/SD UE multiplexing (scheduling), i.e., the allocated individual UE bandwidths are multiples of the PRB bandwidth. The actual reported CQI's are based on received signal-to-interference-and-noise ratios (SINR), calculated by the UE's for each PRB. Here the UE's are assumed to deploy dual antenna linear MMSE (LMMSE) receiver principle, and utilize in the SINR calculations the actual radio channel response, the received noise level, and the structure of the detector (like described in more details below). Furthermore, like mentioned already earlier, the UE's always report single stream as well as both SU and MU dual stream SINR's (at the corresponding detector output).

In a single simulation run, mobile stations are randomly dropped or positioned over each sector and cell. Then based on the individual distances between the mobile and the serving base station, the path losses for individual links are directly determined, while the actual fading characteristics of the radio channels depend on the assumed mobility and power delay profile. In updating the fading statistics, the time resolution in our simulator is set to one TTI (1ms). In general, a standard hexagonal cellular layout is utilized with altogether 19 cell sites each having 3 sectors in Macro case and 1 sector in Micro case as show in Figure 4. In the performance evaluations, statistics are collected only from the central cell site while the others simply act as sources of inter-cell interference.

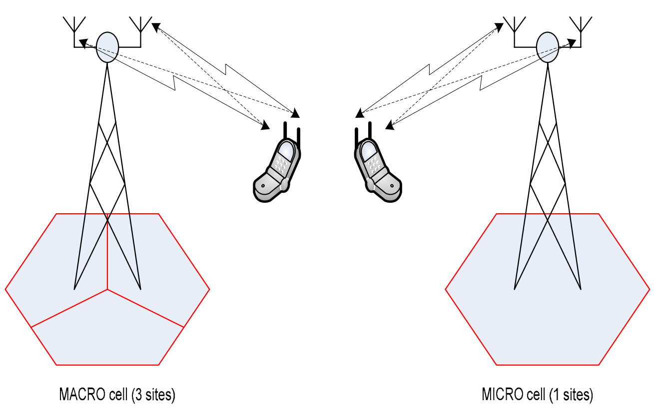


Figure 4: Macro and Micro cell scenarios.

The main simulation parameters and assumptions are generally summarized in TABLE 1 for both Macro and Micro cell scenarios, following again the LTE working assumptions. The used MIMO scheme is *per-antenna rate control* (PARC) with two transmit antennas at the BS and two receive antennas at the UE's and the receivers are equipped with LMMSE detectors. As illustrated in Figure 1, the RRM functionalities are controlled by the packet scheduler together with link adaptation and HARQ mechanisms. Notice that the maximum number of simultaneously multiplexed users (I_{BUFF}) is set to 10 here while the total number of UE's (I_{TOT}) is 15. In general, we assume that the BS transmission power is equally distributed among all PRB's. In the basic simulations, 15 UE's are uniformly dropped within each cell and experience inter-cell interferences from the surrounding cells, in addition to path loss and fading. The UE velocities are 3km/h, and the typical urban (TU) channel model standardized by ITU is assumed in modeling the power-delay spread of the radio channels. Infinite buffer traffic model is applied in the simulations, i.e. every user has data to transmit (when scheduled) for the entire duration of a simulation cycle. Exponential effective SINR mapping (EESM) is used for link-to-system level mapping (throughput calculations), as described in [2]. The length of a single simulation run is set to 5 seconds which is then repeated for 10 times to collect reliable statistics.

Considering MIMO functionality, every UE has an individual HARQ entry per stream, which operates the physical layer re-transmission functionalities. It is based on the stop-and-wait (SAW) protocol and for simplicity, the number of entries per UE is fixed to six. HARQ retransmissions are always transmitted with the same MCS and on the same PRB's (if scheduled) as the first transmissions in a single stream mode. The supported modulation schemes are QPSK, 16QAM and 64QAM with variable rates for the encoder as shown in TABLE 1.

Link adaptation handles the received UE reports containing the channel quality information for each PRB based on single and dual stream MIMO modes. The implemented link adaptation mechanism consists of two separate elements – the inner loop (ILLA) and outer loop (OLLA) LA's – and are used for removing CQI imperfections and estimating supported data rates and MCS. It is assumed that the CQI reporting errors are log-normal distributed with 1dB standard deviation.

B. *Detectors and SINR Modeling*

The actual effective SINR calculations rely on subcarrier-wise complex channel gains (estimated using reference symbols in practice) and depend in general also on the assumed receiver (detector) topology. Here we assume that the LMMSE detector, properly tailored for the transmission mode (1-stream SU, 2-stream SU or 2-stream MU) is deployed. The detector structures and SINR modeling for different transmission modes are described in detail below.

TABLE 1. Basic simulation parameters.

Parameter	Assumption
Cellular layout	Hexagonal grid, 19 cell sites, 3 sectors per site for Macro / 1 sector per site for Micro
Inter-site distance	500 m - Macro / 1732 m - Micro
Carrier frequency / Bandwidth	2000 MHz / 10 MHz
Number of active sub-carriers	600
Sub-carrier spacing	15 kHz
Sub-frame duration	0.5 ms
Channel estimation	Ideal
PDP	ITU Typical Urban 20 paths
Minimum distance between UE and cell	>= 35 meters - Macro >= 10 meters - Micro
Average number of UE's per sector	15
Max. number of frequency multiplexed UEs (I_{BUFF})	10
UE receiver type	LMMSE
Shadowing standard deviation	8 dB
UE speed	3km/h
Total BS TX power (Ptotal)	46 dBm
Traffic model	Full Buffer
Fast fading model	Jakes Spectrum
CQI reporting schemes	Full CQI, Best – m (with m=10), Threshold based (with 5dB threshold)
CQI log-normal error std.	1 dB
CQI reporting time	5 TTI
CQI delay	2 TTIs
CQI quantization	1 dB
CQI std error	1 dB
MCS rates	QPSK (1/3, 1/2, 2/3), 16QAM (1/2, 2/3, 4/5), 64QAM (1/2, 2/3, 4/5)
ACK/NACK delay	2 ms
Number of SAW channels	6
Maximum number of retransmissions	3
HARQ model	Ideal chase combining (CC)
1st transmission BLER target	20%
Scheduler forgetting factor	0.0025
Scheduling schemes used	Ordinary PF Modified PF (proposed)
Simulation duration (one drop)	5 seconds
Number of drops	10

1) *Single Stream SU Case*

In this case, only one of the two BS transmit antennas is used to transmit one stream. At individual time instant (time-index dropped here), the received spatial 2x1 signal vector of UE *i* at sub-carrier *c* is of the form

$$\mathbf{y}_{i,c} = \mathbf{h}_{i,c} x_{i,c} + \mathbf{n}_{i,c} + \mathbf{z}_{i,c} \tag{9}$$

where $x_{i,c}$, $\mathbf{h}_{i,c}$, $\mathbf{n}_{i,c}$ and $\mathbf{z}_{i,c}$ denote the transmit symbol, 2x1 channel vector, 2x1 received noise vector and 2x1 inter-cell interference vector, respectively. Then the LMMSE detector $\hat{x}_{i,c} = \mathbf{w}_{i,c}^H \mathbf{y}_{i,c}$ is given by

$$\mathbf{w}_{i,c} = \sigma_{x,i}^2 (\mathbf{h}_{i,c}^H \sigma_{x,i}^2 \mathbf{h}_{i,c} + \mathbf{\Sigma}_{n,i} + \mathbf{\Sigma}_{z,i})^{-1} \mathbf{h}_{i,c} \tag{10}$$

where $\sigma_{x,i}^2$, $\mathbf{\Sigma}_{n,i}$ and $\mathbf{\Sigma}_{z,i}$ denote the transmit power (per the used antenna), noise covariance matrix and inter-cell interference covariance matrix, respectively. Now the SINR is given by

$$\gamma_{i,c} = \frac{\left|\mathbf{w}_{i,c}^H \mathbf{h}_{i,c}\right|^2 \sigma_{x,i}^2}{\mathbf{w}_{i,c}^H \mathbf{\Sigma}_{n,i} \mathbf{w}_{i,c} + \mathbf{w}_{i,c}^H \mathbf{\Sigma}_{z,i} \mathbf{w}_{i,c}} \tag{11}$$

The noise variables at different receiver antennas are assumed uncorrelated (diagonal $\mathbf{\Sigma}_{n,i}$) while the more detailed modeling of inter-cell interference (structure of $\mathbf{\Sigma}_{z,i}$) takes into account the distances and channels from neighboring base stations (for more details, see, e.g., [19]).

2) *Dual Stream SU Case*

In this case, both of the two BS transmit antennas are used for transmission, on one stream per antenna basis. At individual time instant, the received spatial 2x1 signal vector of UE *i* at sub-carrier *c* is now given by

$$\mathbf{y}_{i,c} = \mathbf{H}_{i,c} \mathbf{x}_{i,c} + \mathbf{n}_{i,c} + \mathbf{z}_{i,c} \tag{12}$$

where $\mathbf{x}_{i,c}$ and $\mathbf{H}_{i,c} = [\mathbf{h}_{i,c,1}, \mathbf{h}_{i,c,2}]$ denote the 2x1 transmit symbol vector and 2x2 channel matrix, respectively. Now the LMMSE detector $\hat{\mathbf{x}}_{i,c} = \mathbf{W}_{i,c} \mathbf{y}_{i,c}$ is given by

$$\mathbf{W}_{i,c} = \mathbf{\Sigma}_{x,i} \mathbf{H}_{i,c}^H (\mathbf{H}_{i,c} \mathbf{\Sigma}_{x,i} \mathbf{H}_{i,c}^H + \mathbf{\Sigma}_{n,i} + \mathbf{\Sigma}_{z,i})^{-1} = \begin{bmatrix} \mathbf{w}_{i,c,1}^H \\ \mathbf{w}_{i,c,2}^H \end{bmatrix} \tag{13}$$

where $\mathbf{\Sigma}_{x,i} = diag\{\sigma_{x,i,1}^2, \sigma_{x,i,2}^2\} = diag\{\sigma_{x,i}^2/2, \sigma_{x,i}^2/2\}$ denotes the 2x2 covariance matrix (assumed diagonal) of the transmit symbols. Note that compared to single stream case, the overall BS transmit power is now divided between the two antennas, as indicated above. Then the SINR's for the two transmit symbols are given by

$$\gamma_{i,c,1} = \frac{\left|\mathbf{w}_{i,c,1}^H \mathbf{h}_{i,c,1}\right|^2 \sigma_{x,i,1}^2}{\left|\mathbf{w}_{i,c,1}^H \mathbf{h}_{i,c,2}\right|^2 \sigma_{x,i,2}^2 + \mathbf{w}_{i,c,1}^H \mathbf{\Sigma}_{n,i} \mathbf{w}_{i,c,1} + \mathbf{w}_{i,c,1}^H \mathbf{\Sigma}_{z,i} \mathbf{w}_{i,c,1}}$$

$$\gamma_{i,c,2} = \frac{\left|\mathbf{w}_{i,c,2}^H \mathbf{h}_{i,c,2}\right|^2 \sigma_{x,i,2}^2}{\left|\mathbf{w}_{i,c,2}^H \mathbf{h}_{i,c,1}\right|^2 \sigma_{x,i,1}^2 + \mathbf{w}_{i,c,2}^H \mathbf{\Sigma}_{n,i} \mathbf{w}_{i,c,2} + \mathbf{w}_{i,c,2}^H \mathbf{\Sigma}_{z,i} \mathbf{w}_{i,c,2}} \tag{14}$$

3) Dual Stream MU Case

In this case, the transmission principle and SINR modeling are similar to subsection 2) above, but the two spatially multiplexed streams belong now to two different UE's, say *i* and *i*'. Thus the SINR's in (14) are interpreted accordingly.

Finally, for link-to-system level mapping purposes, the exponential effective SINR mapping (EESM), as described in [2-4], is deployed.

V. Numerical Results

In this section, we present the results obtained from the quasi-static radio system simulations using the PS algorithms described in the article. The system-level performance is generally measured and evaluated in terms of:

a) *Throughput* – the total number of successfully delivered bits per unit time. Usually measured either in kbps or Mbps.

b) *Coverage* – the experienced data rate per UE at the 95% coverage probability (5% throughput CDF point).

c) *Fairness* – measures the resource allocation fairness among all UE's from the average throughput point of view. Evaluated using Jain's fairness index [21] which is calculated here as

$$fairness = \frac{\left(\sum_{i=1}^{I_{\mathrm{TOT}}} \overline{\mu}_i\right)^2}{I_{\mathrm{TOT}} \sum_{i=1}^{I_{\mathrm{TOT}}} \overline{\mu}_i^{\,2}} \tag{15}$$

where $\overline{\mu}_i$ denotes the average throughput value of user *i* across different simulation realizations (here ten). In single simulation run, the corresponding throughput of user *i* is given by

$$\mu_i = \frac{1}{N_{TTI}} \sum_{n=1}^{N_{TTI}} R_i(n) \tag{16}$$

where N_{TTI} denotes the length of a single simulation run in TTI's while $R_i(n)$ is the actually delivered throughput to user *i* at individual TTI *n*.

In the following, we illustrate the behavior of the proposed MMPF scheduler with using different power coefficients α_1 and α_2 as shown in TABLE 2. To focus mostly on the role of the channel quality reporting in the priority metric calculation in equation (3), α_2 is fixed here to 1 and different values are then demonstrated for α_1. More specifically, with large α_2 values, the effect of second term in priority metric calculation based on throughput estimation would be emphasized and the scheduling algorithm would behave like maximum throughput scheduler, which in turn would imply reduced fairness distribution. Consequently, the impact of the different CQI reporting schemes is seen now more clearly. For the cases of *Best–m* and *Threshold* based CQI reporting schemes, we fix the value of *m* equal to 10 and threshold to 5 dB, respectively. Similar example values have also been used by other authors in the literature earlier, see, e.g., [12]. Complete performance statistics are gathered for both Macro cell and Micro cell case scenarios.

A. Macro Cell Case

Figure 5 (left column) illustrates the average user throughput and coverage for the different schedulers. The power coefficient values from TABLE 2 are presented as index M, where M1 represents the first couple, etc. The obtained results are compared with the reference PF scheduler described also in Section II. By using the first term (M1) of the new metric calculation for MMPF, in combination with full CQI reporting scheme, we achieve coverage gain in the order of 63% at the expense of 20% throughput loss as shown in Figure 5 (a) and (b). This sets the basic reference for comparisons in the other cases. In the case of *Best-m* and *Threshold* based reporting schemes presented in Figure 5 (c) and (d), and Figure 5 (e) and (f), we have coverage increases by 69% and 74% with throughput losses of 15% and 20%, correspondingly.

Continuing on the evaluation of relative system performance using the proposed scheduler, we clearly see a trade-off between average cell throughput and coverage for different power coefficient cases. Furthermore, the remaining power coefficient values shown in TABLE 2 are used for tuning the overall system behavior together with the choice of the CQI reporting scheme. In the case of full CQI feedback and coefficient α_1 varying between 2 and 8 (M2 –M5) the cell throughput loss is decreased to around 7%, while the coverage gain is reduced to around 21%. Similar results are obtained for the other feedback reporting schemes as well. Consequently, an obvious trade-off between average cell throughput and coverage is clearly seen. In order to preserve the average sector throughput and still to gain from the coverage increase, coefficient values should thus be properly chosen. The exact percentage values for the coverage gains and throughput losses are stated in TABLE 3 at the end of the article.

Further illustrations on the obtainable system performance are presented in Figure 5 (right column) in terms of the statistics of individual UE data rates for the applied simulation scenarios. The slope of the CDF reflects generally the fairness of the algorithms. Therefore we aim to achieve steeper slope corresponding to algorithm fairness. This type of slope change behavior can clearly be established for each simulation scenario. Clearly, at 5% (coverage) point of the throughput CDF curves, corresponding to users typically situated at the cell edges, we observe significant data rate increases indicated by shift to the right for all CQI feedback schemes when the coefficient α_1 is changed in the proposed metric. This indicates improved overall cell coverage at the expense of slight total throughput loss.

Figure 6 (left column) shows the modulation and coding scheme (MCS) distributions for different schedulers and with applied feedback reporting schemes. The negligible decrease in higher order modulation usage (less than 3%) leads to the increase in the lower (more robust) ones for improving the cell coverage. In all the simulated cases, the MCS distribution behavior has a relatively similar trend following the choice of the power coefficients in the proposed packet scheduling. In general, the use of higher-order modulations is affected mostly in the case of *Best-m* and *Threshold* based reporting schemes.

Figure 6 (right column) illustrates the HARQ distributions for the different scheduler scenarios and reporting schemes. Clearly, the 20% BLER target rate is achieved in all simulated cases. Moreover, the MMPF scheduler provides slight increase in probability of successful first transmission of around 2% for the *Best-m* and *Threshold* based feedback cases.

TABLE 2. DIFFERENT POWER COEFFICIENT COMBINATIONS

Coefficient	Value				
α_1	1	2	4	6	8
α_2	1	1	1	1	1

B. Micro Cell Case

The performance statistics obtained for Micro cell case demonstrate similar trends, as in the previous Macro case, as shown in Figure 7. Starting from the primary case M1, with full CQI, we obtain a 17% loss in throughput and 92% coverage improvement. For the reduced feedback reporting schemes – *Best-m* and *Threshold* based – we have 21% and 19% throughput losses and 100% and 96% coverage gains, respectively. Furthermore, similar behavior is observed in the CDF's of individual UE throughputs, as well as MCS and HARQ distributions. The exact percentage values read from the figures are again stated in table format in TABLE 4 at the end of the article.

C. Jain's Fairness Index

Figure 8 illustrates the Jain's fairness index per scheduling scheme for Micro and Macro cell scenarios, calculated over all the I_{TOT} = 15 UE's. The value on the x axis corresponds to used scheduler type, where 1 refers to the reference PF scheduler, 2 refers to MMPF with index M1, etc. The value of Jain's fairness index is generally in the range of [0,1], where value of 1 corresponds to all users having the same amount of resources. Clearly, the fairness distribution with MMPF outperforms the used reference PF scheduler for both cases. The received fairness gains are in range of 13%-37% with *full* CQI feedback, 15-32% with *Best-m* CQI feedback and 17-35% with *Threshold* based CQI feedback in the Macro case scenario. The corresponding fairness gains in Micro case scenario are 25-46%, 32-41% and 34-43% for *full* CQI, *Best –m* and *Threshold* based reporting schemes, correspondingly. The exact percentage values read from the figures are again stated in table format in Table 5 in the end of the article.

VI. CONCLUSIONS

In this article, we have studied the potential of advanced multiuser packet scheduling algorithms in OFDMA type radio system context, using UTRAN long term evolution (LTE) downlink in Macro and Micro cell environment as practical example cases. New multistream proportional fair scheduler metric covering time-, frequency- and spatial domains was proposed that takes into account both the instantaneous channel qualities (CQI's) as well as resource allocation fairness. Also different practical CQI reporting schemes were discussed, and used in the system level performance evaluations of the proposed scheduler. Overall, the achieved throughput performance together with coverage and fairness statistics were assessed, by using extensive radio system simulations, and compared against more traditional proportional fair scheduling with multiantenna spatial multiplexing functionality. In the case of fixed coverage requirements and based on the optimal parameter choice for CQI reporting schemes, the proposed scheduling metric calculations based on UE channel feedback offers better control over the ratio between the achievable cell/UE throughput and coverage increase, as well as increased UE fairness. As a practical example, even with limited CQI feedback, the fairness in resource allocation together with cell coverage can be increased significantly (more than 40%) by allowing a small decrease (in the order of only 10-15%) in the cell throughput.

ACKNOWLEDGMENT

The authors would like to thank Markku Kuusela, Nokia Devices, Helsinki, Finland, and Dr. Toni Huovinen, Tampere University of Technology, Tampere, Finland, for fruitful discussions.

REFERENCES

[1] S. Nonchev and M. Valkama, "Efficient packet scheduling schemes for multiantenna packet radio downlink," in Proc. Fifth Advanced Int. Conf. Telecommunications (AICT-2009), Venice, Italy, May 2009.

[2] "Physical Layer Aspects for Evolved UTRA", 3GPP Technical Report TR 25.814, ver. 7.1.0, Oct. 2006.

[3] WiMAX Forum™, http://www.wimaxforum.org, Aug. 2006.

[4] "Wireless World Initiative New Radio (WINNER)," IST-507581, European Information Society Technologies, Technical Report, Dec. 2006.

[5] N. Wei, B. Talha, T. B. Sørensen, T. E. Kolding, and P. E. Mogensen, "Spectral efficiency of closed-loop transmit diversity with limited feedback for UTRA Long Term Evolution," in Proc. IEEE Int. Symp. Personal, Indoor and Mobile Radio Communications (PIMRC-2006), Helsinki, Sept. 2006.

[6] R. W. Heath, M. Airy, and A. Paulraj, "Multiuser diversity for MIMO wireless systems with linear receivers," in Proc. Asilomar Conf. Signals, Systems and Computers, Pacific Grove, CA, 2001.

[7] E. Dahlman et al., 3G Evolution: HSPA and LTE for Mobile Broadband, Academic Press, 2007.

[8] Y. Sun et al., "Multi-user scheduling for OFDMA downlink with limited feedback for evolved UTRA," in Proc. IEEE Vehicular Technology Conf. (VTC-2006 Fall), Montreal, Canada, Sept. 2006.

[9] S. Yoon, C. Suh, Y. Cho, D. Park, "Orthogonal frequency division multiple access with an aggregated sub-channel structure and statistical channel quality measurements," in Proc. IEEE Vehicular Technology Conf. (VTC-2004 Fall), Los Angeles, CA, Sept. 2004.

[10] I. Toufik and H. Kim, "MIMO-OFDMA opportunistic beamforming with partial channel state information," in Proc. IEEE Int. Conf. Communications (ICC-2006), Instanbul, Turkey, June 2006.

[11] T.E. Kolding, F. Frederiksen, A. Pokhariyal, "Low-bandwidth channel quality indication for OFDMA frequency domain packet scheduling," in Proc. IEEE Int. Symp. Wireless Communication Systems (ISWCS-2006), Valencia, Spain, Sept. 2006.

[12] K.I. Pedersen, G. Monghal, I.Z. Kovacs, T.E. Kolding, A. Pokhariyal, F. Frederiksen, P. Mogensen, "Frequency domain scheduling for OFDMA with limited and noisy channel feedback," in Proc. IEEE Vehicular Technology Conf. (VTC-2007 Fall), Baltimore, MD, Sept. 2007.

[13] P. Svedman, D. Hammarwall, B. Ottersten, "Sub-carrier SNR estimation at the transmitter for reduced feedback OFDMA," in Proc. European Signal Processing Conf., Florence, Italy, Sept. 2006.

[14] C. Wengerter, J. Ohlhorst, and A.G.E Von Elbwert, "Fairness and Throughput Analysis for Generalized Proportional Fair Frequency Scheduling in OFDMA," in Proc. IEEE Vehicular Technology Conf. (VTC-2005 Spring), Stockholm, Sweden, May 2005.

[15] T. E. Kolding, "Link and system performance aspects of proportional fair scheduling in WCDMA/HSDPA," in Proc. IEEE Vehicular Technology Conf. (VTC-2003 Fall), Orlando, FL, USA, Oct. 2003.

[16] P. Svedman, S.K.Wilson, L.J. Cimini, and B. Ottersten, "A simplified opportunistic feedback and scheduling scheme for OFDMA," in Proc. IEEE Vehicular Technology Conf. (VTC-2004 Spring), Milan, Italy, May 2004.

[17] I. Z. Kovács et al, "Effects of non-ideal channel feedback on dual-stream MIMO OFDMA system performance," in Proc. IEEE Vehicular Technology Conf. (VTC-2007 Fall), Baltimore, MD, Oct. 2007.

[18] S. Nonchev, J. Venäläinen and M. Valkama, "New frequency domain packet scheduling schemes for UTRAN LTE downlink," in Proc. ICT-MobileSummit, Stockholm, Sweden, 2008.

[19] S. Nonchev and M. Valkama, "A new fairness-oriented packet scheduling scheme with reduced channel feedback for OFDMA packet radio systems," Int. Journal of Communications, Network and System Sciences (IJCNS), vol. 2, No. 7, Oct. 2009.

[20] N. Wei, et al., "Performance of MIMO with frequency domain packet scheduling," in Proc. IEEE Vehicular Technology Conf. (VTC-2007 Spring), Dublin, Ireland, May 2007.

[21] D. Chui and R. Jain, "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks," Computer Networks and ISDN Systems, 1989.

[22] P. Svedman, L.J. Cimini, B. Ottersten, "Using unclaimed sub-carriers in opportunistic OFDMA systems," in Proc. IEEE Vehicular Technology Conf. (VTC-2006 Fall), Montreal, Canada, September 2006.

[23] S. Sanayei, A. Nosratinia, N. Aldhahir, "Opportunistic dynamic sub-channel allocation in multiuser OFDM networks with limited feedback," in Proc IEEE. Information Theory Workshop, San Antonio, TX, October 2004, pp. 182-186.

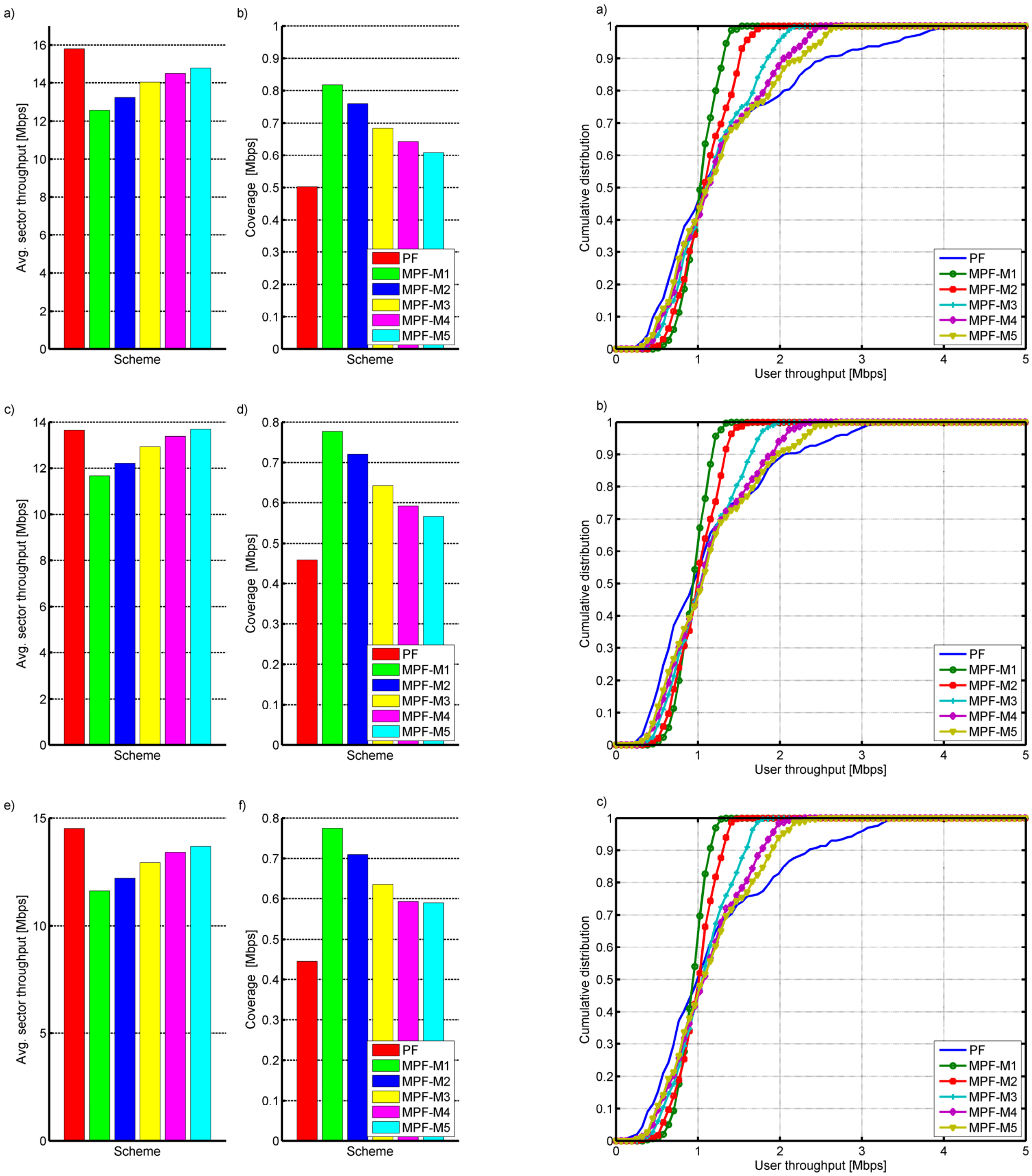


Figure 5: Left column: Average sector throughput and coverage for different scheduling schemes for Macro cell scenario with *full* CQI feedback (a, b), *Best -m* CQI feedback (c, d) and *Threshold* based CQI feedback (e, f). M1-M5 refer to the proposed scheduler with power coefficient values as given in TABLE 2. Right column: CDF's of individual UE throughputs for different scheduling schemes for Macro cell scenario with *full* CQI feedback (a), *Best -m* CQI feedback (b) and *Threshold* based CQI feedback (c).

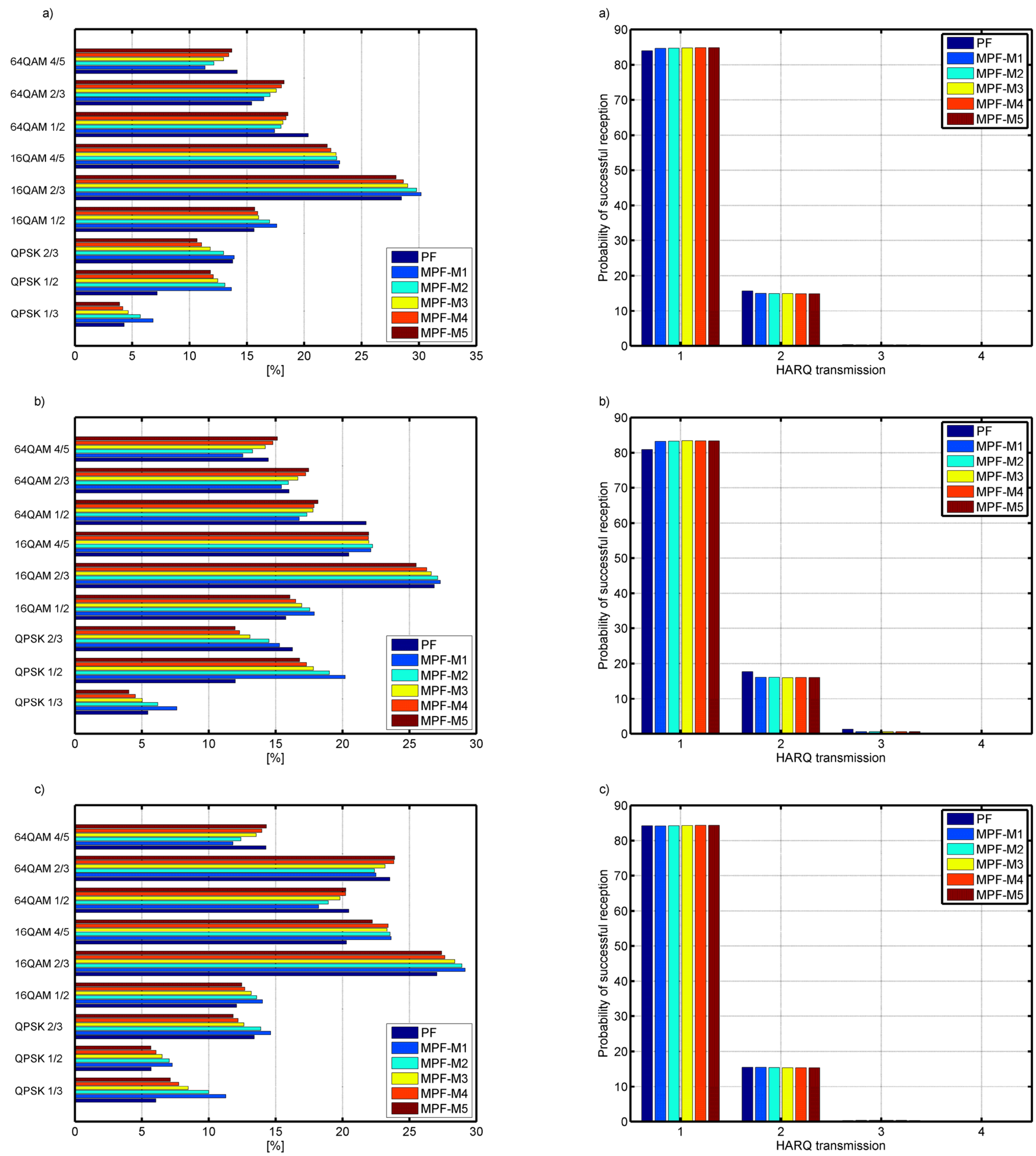


Figure 6: Left column: MCS distributions [%] for different scheduling principles for Macro cell scenario with *full* CQI feedback (a), *Best -m* CQI feedback (b) and *Threshold* based CQI feedback (c). Right column: HARQ distributions for different scheduling schemes for Macro cell scenario with *full* CQI feedback (a), *Best -m* CQI feedback (b) and *Threshold* based CQI feedback (c).

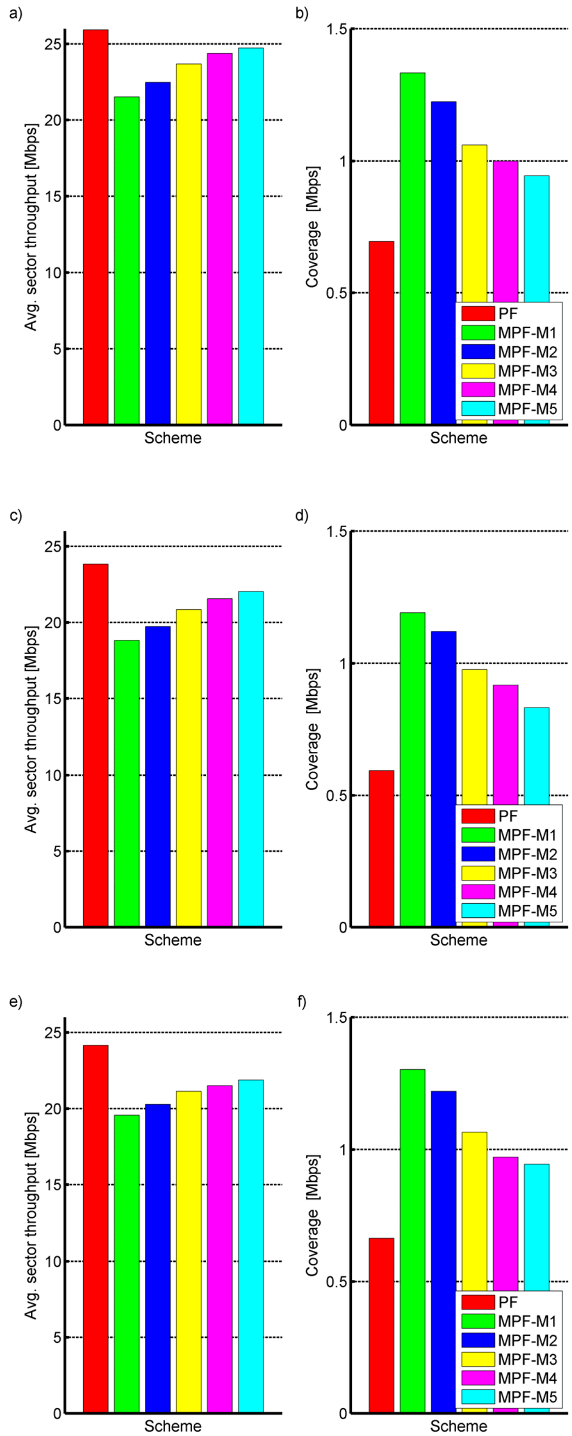


Figure 7: Average cell throughput and coverage gain over the reference PF scheduling scheme for Micro cell simulation scenario. The schemes M1-M5 refer to the new proposed scheduler with power coefficient values as given in TABLE 2.

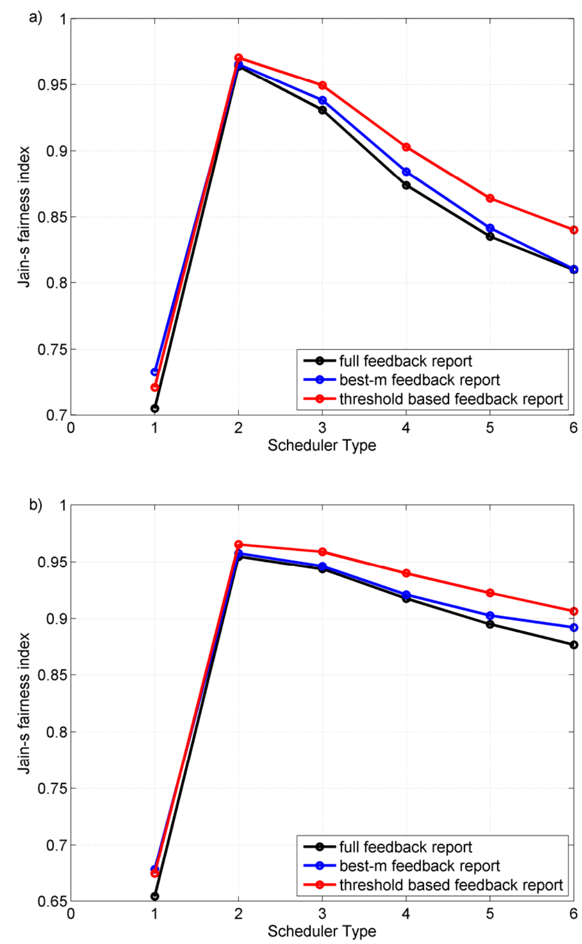


Figure 8: Jain's fairness index per feedback reporting scheme for the different simulation scenarios – Macro cell (a) and Micro cell (b). Scheduler type 1 means ordinary PF, while 2-6 means proposed modified PF with power coefficients as described in TABLE 2.

TABLE 3.
OBTAINED PERFORMANCE STATISTICS COMPARED TO ORDINARY PF SCHEDULER WITH DIFFERENT CQI REPORTING SCHEMES AND DIFFERENT POWER COEFFICIENTS (M1-M5) FOR THE PROPOSED SCHEDULER. MACRO CELL CASE SCENARIO.

	Coverage Gain [%]			Throughput Loss [%]		
	full	best-m	threshold	full	best-m	threshold
M1	63	69	74	20	15	20
M2	51	57	60	16	11	16
M3	36	40	43	13	6	12
M4	28	29	33	9	2	8
M5	21	24	32	7	0	6

TABLE 4.
OBTAINED PERFORMANCE STATISTICS COMPARED TO ORDINARY PF SCHEDULER WITH DIFFERENT CQI REPORTING SCHEMES AND DIFFERENT POWER COEFFICIENTS (M1-M5) FOR THE PROPOSED SCHEDULER. MICRO CELL CASE SCENARIO.

	Coverage Gain [%]			Throughput Loss [%]		
	full	best-m	threshold	full	best-m	threshold
M1	92	100	96	17	21	19
M2	76	88	84	13	17	16
M3	53	64	60	9	14	14
M4	44	54	46	6	11	12
M5	36	40	42	5	8	10

TABLE 5.
OBTAINED JAIN'S FAIRNESS INDEXES COMPARED TO ORDINARY PF SCHEDULER WITH DIFFERENT CQI REPORTING SCHEMES AND DIFFERENT POWER COEFFICIENTS (M1-M5) FOR THE PROPOSED SCHEDULER. MACRO AND MICRO CELL CASE SCENARIOS.

	Macro case Fairness Gain [%]			Micro case Fairness Gain [%]		
	full	best-m	threshold	full	best-m	threshold
M1	37	32	35	46	41	43
M2	32	28	32	44	40	42
M3	19	21	25	29	36	39
M4	16	15	20	27	33	37
M5	13	15	17	25	32	34

Publication 11

S. Nonchev and M. Valkama, “Advanced radio resource management for multiantenna packet radio systems,” *International Journal on Wireless and Mobile Networks (IJWMN)* vol 2.2, pp. 1-14, 2010.

ADVANCED RADIO RESOURCE MANAGEMENT FOR MULTIANTENNA PACKET RADIO SYSTEMS

Stanislav Nonchev and Mikko Valkama
Department of Communications Engineering, Tampere University of Technology, Tampere, Finland
stanislav.nonchev@tut.fi
mikko.e.valkama@tut.fi

ABSTRACT

In this paper, we propose fairness-oriented packet scheduling (PS) schemes with power-efficient control mechanism for future packet radio systems. In general, the radio resource management functionality plays an important role in new OFDMA based networks. The control of the network resource division among the users is performed by packet scheduling functionality based on maximizing cell coverage and capacity satisfying, and certain quality of service requirements. Moreover, multi-antenna transmit-receive schemes provide additional flexibility to packet scheduler functionality. In order to mitigate inter-cell and co-channel interference problems in OFDMA cellular networks soft frequency reuse with different power masks patterns is used. Stemming from the earlier enhanced proportional fair scheduler studies for single-input multiple-output (SIMO) and multiple-input multiple-output (MIMO) systems, we extend the development of efficient packet scheduling algorithms by adding transmit power considerations in the overall priority metrics calculations and scheduling decisions. Furthermore, we evaluate the proposed scheduling schemes by simulating practical orthogonal frequency division multiple access (OFDMA) based packet radio system in terms of throughput, coverage and fairness distribution among users. In order to completely reveal the potential of the proposed schemes we investigate the system performance of combined soft frequency reuse schemes with advanced power-aware packet scheduling algorithms for further optimization. As a concrete example, under reduced overall transmit power constraint and unequal power distribution for different sub-bands, we demonstrate that by using the proposed power-aware multi-user scheduling schemes, significant coverage and fairness improvements in the order of 70% and 20%, respectively, can be obtained, at the expense of average throughput loss of only 15%.

KEYWORDS

Radio Resource Management, Packet Scheduling, Soft Frequency Reuse, Proportional-fair, Power Masks, Channel Quality Feedback, Fairness, Throughput

1. INTRODUCTION

Development of new advanced packet radio systems continues progressively. Relevant work in this direction includes, e.g., 3GPP long term evolution (LTE) [1], WiMAX [2], IMT-Advanced [3] and the related work in various research projects, like WINNER [4]. These developments rely on OFDMA air interface, scalable bandwidth operation, exploitation of MIMO technologies and advanced convergence techniques. The operating bandwidths are divided into large number of orthogonal subcarriers for multi-access purposes and efficient reduction of the effects of inter-symbol interference (ISI) and inter-carrier interference (ICI).The major driving forces behind these developments are the increased radio system performance, in terms of average and peak cell throughputs, low latency and reduced operating expenditures. While the average and peak throughputs are typically emphasized, also the fairness and cell-edge coverage are equally important quality measures of cellular radio systems [5], [6]. Yet another key measure becoming all the time more and more important is the energy consumption of the radio access network.

Another important issue in these deployments is co-channel interference (CCI) that users at the cell borders are experiencing based on low signal to noise ratio (SNR) and high power emitters from neighbouring cells' base stations in their communication channel. Solution for this problem in OFDMA cellular networks is utilization of controlled frequency reuse schemes. Recent literature studies indicate that soft frequency reuse scheme has a capacity gain over the other hard frequency reuse schemes [7]. The overall functionality of frequency reuse schemes is based on applying specific power masks (fraction of the maximum transmission power level) over the whole system bandwidth. In soft frequency reuse (SFR) [8],[9] the frequency band, shared by all base stations (BS) (reuse factor is equal to 1) is divided into sub-bands with predefined power levels as illustrated in Figure 1. Consequently, the users near to and far away from the BS are allocated with different powers, limiting the impact of and to the neighbouring cells.

In general, all user equipments (UEs) within base station (BS) coverage are sharing the available radio resource (radio spectrum). This is controlled by the packet scheduler (PS) utilizing selected scheduling metrics. In both uplink (UL) and downlink (DL), the packet scheduler works in a centralized manner being located at the BS. It is generally fairly well understood that the overall radio system performance, in terms of throughput and fairness, depends heavily on the PS functionality. The PS operation can build on instantaneous radio channel conditions, quality of service (QoS) requirements and traffic situations of the served UEs [5], [6]. In the emerging radio systems, multi-antenna MIMO technologies increase spectral efficiency and also provide the PS with extra degree of freedom by offering a possibility to multiplex data of one or more users on the same physical resource (spatial domain multiplexing, SDM) [10], [11]. Such functionality requires additional feedback information provided by mobile stations (MS) to the base station, typically in terms of channel quality information (CQI) reports. The main disadvantage of MIMO-SDM technique is increased signaling overhead and scheduling complexity. Finally, an additional important aspect in scheduling functionality is the interaction with other radio resource management (RRM) entities, namely fast link adaptation (LA) and reliable re-transmission mechanisms.

Packet scheduling principles are generally rather widely investigated in the literature and many new scheduling algorithms have been proposed, see, e.g., [12]-[16]. Most of them are considering equal BS transmission power distribution among all physical resource blocks (PRBs), which is not necessarily a practical or optimum case as already indicated above. In practice, unequal power allocation between different PRB's can be used, e.g., to control the interference between neighboring cells or sectors in frequency re-use 1 radio systems. Stemming now from our previous work in advanced PS algorithms reported in [15]-[18], we extend the studies here to incorporate energy efficiency and cell power bounds considerations in the scheduling decisions based on different SFR power mask configurations. Our starting point is the advanced modified proportional fair scheduler for MIMO reported in [16], which was shown to improve cell-edge coverage and user fairness compared to state-of-the-art. In this paper, we introduce a new energy- or power-aware scheduling scheme taking into account the applied power pattern within the overall radio spectrum. This is called MIMO power-aware modified proportional fair (MPMPF) scheduler in the continuation. The performance of the proposed scheduler is investigated using both single-user MIMO (spatial multiplexing for individual UE streams) and multi-user MIMO (spatial multiplexing also for streams of different UEs) transmit-receive schemes in 3GPP LTE system context. The used system-level figures of merit are cell throughput distribution, cell-edge coverage and Jain's fairness index [19]. For simplicity and illustration purposes, 1x2 (SIMO) and 2x2 (MIMO) multi-antenna scenarios are assumed in the continuation.

The rest of the paper is organized as follows: Section 2 describes the SFR scheme and power mask configurations, while Section 3 is dedicated to the MIMO scheduling principles and proposed power-aware scheduling scheme. Section 4 gives an overview of the overall system model and simulation assumptions. The simulation results and analysis are presented in Section 5, while the conclusions are drawn in Section 6.

2. SOFT FREQUENCY REUSE SCHEMES

In general, SFR scheme reserves part of the frequency band for the cell-edge users and uses the power bound specified for it by the power mask. The rest of the unallocated sub-bands are dedicated to the near to BS users. One practical solution for sub-band division reported in the literature studies is 3 and therefore we are considering it in our case. Generally there is no restriction on the soft reuse factor [7],[8]. The same apply for the choice of the power mask. We have chosen for our evaluation the following power mask configurations: PM1 (0dB, -4dB, -4dB) and PM2 (0dB, -1dB, -4dB). The values in the brackets represent nominal transmission power values in dB as shown in Figure 1. Notice that if the sub-band allocated to the cell-edge UEs is not fully occupied, it can be still used by the other UEs.

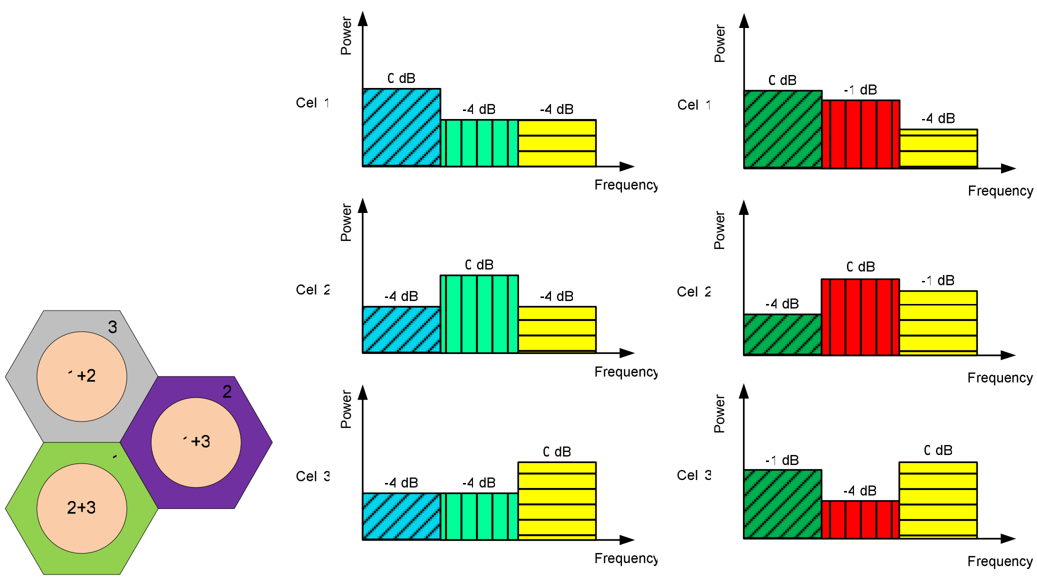


Figure 1: Soft Frequency reuse scheme with 3 sub-bands division.

3. POWER-AWARE MIMO SCHEDULING

In general, packet scheduler is part of the overall radio resource management mechanism at BS. Therefore, PS functionality depends heavily on the collaboration with other RRM units such as *link adaptation* (LA) and *Hybrid ARQ* (HARQ) manager, as depicted in Figure 2. The scheduling decisions are based on the selected scheduling metric, utilizing typically the CQI reports and acknowledgements from UEs, per given transmission time interval (TTI) in time domain and per physical resource block (PRB) in frequency domain [5]. Moreover, we assume here that the LA unit can feed the PS with power allocation information on a PRB-per-PRB basis. MIMO functionality, in turn, requires both single-stream and dual-stream CQI feedback by each UE. Consequently, the BS decides whether the particular time-frequency resource is used for transmitting (*i*) only one stream to a specific UE, (*ii*) two streams to a specific UE (SU-MIMO) or (*iii*) 1+1 streams to two different UEs (MU-MIMO) [6]. BS also handles proper transmit power allocation in all the cases (*i*)-(*iii*), respectively, such that target packet error rate is reached with selected modulation and coding scheme (MCS).

3.1. Proposed Scheduler at Principal Level

The proposed scheduler developments are stemming from the widely used two-stage (see e.g. [15], [16]) multi-stream PF approach. In the first stage, within each TTI, UEs are ranked based on the full bandwidth channel state information and throughput calculations. This is called time-domain (TD) scheduling step. For MIMO case, different spatial multiplexing possibilities (one-stream, dual-stream SU, dual-stream MU) are taken into account, in calculating all the possible reference throughputs. In the second stage, the scheduling functionality is expanded in frequency-domain (FD) and spatial-domain (SD) where the actual PRB allocation takes place. First the needed PRB's for pending re-transmissions (on one stream-basis only) are reserved and the rest available PRB's are allocated to the selected UE's from the first stage. The actual priority metric in FD/SD stage is evaluated at PRB-level taking into account the available stream-wise channel state information, the transmit power allocations and the corresponding throughput calculations. The exact scheduling metrics are described below.

3.2. Scheduling Metrics

Here we describe the actual scheduling metrics used in ranking users in the TD scheduling stage as well as mapping the users to FD/SD resources in the second stage. First a power-aware extension of ordinary multi-stream PF scheduler is described in sub-section 3.2.1, used as a reference in the performance simulations. Then the actual proposed modified power-aware metric is described in sub-section 3.2.2.

3.2.1 Power-Aware Multistream Proportional Fair

Stemming from ordinary multi-stream PF principle, the following power-aware PF (PPF) scheduling metric evaluated at each TTI is deployed

$$\gamma_{i,k,s} = arg\ \underset{i}{max}\left\{\frac{P_{k,s}}{P_{\max}}\frac{R_{i,k,s}(n)}{T_i(n)}\right\} \tag{1}$$

Here $R_{i,k.s}(n)$ is the estimated instantaneous throughput of user i at sub-band k on stream s at TTI n. $T_i(n)$ is the corresponding average delivered throughput to the UE i during the recent past [16],[17]. $P_{k,s}$, in turn, is the transmission power for sub-band k on stream s and P_{max} is the maximum transmission power for any sub-band. Notice that the power ratio $P_{k,s} / P_{max}$ has a direct impact on the ranking of the users, since the deployed sub-band power levels affect the corresponding estimated and delivered throughput quantities. The scheduler in (1) is only used as a reference in evaluating the performance of the actual proposed scheduler to be described below.

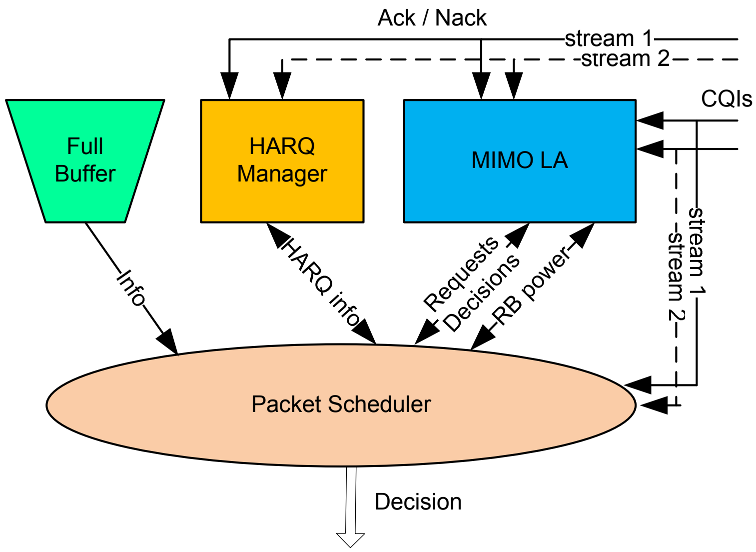


Figure 2: RRM functionalities and scheduling process.

3.2.2 Modified Power-Aware Multi-stream Proportional Fair (MPMPF)

Starting from the earlier work in [16], the following modified power-aware multi-stream PF metric is proposed in this paper:

$$\bar{\gamma}_{i,k,s} = \arg\max_{i} \left\{ \frac{P_{k,s}}{P_{\max}} \left(\frac{CQI_{i,k,s}(n)}{CQI_i^{avg}(n)} \right)^{\alpha_1} \left(\frac{T_i(n)}{T_{tot}(n)} \right)^{-\alpha_2} \right\} \tag{2}$$

In the above metric, α_1 and α_2 are scheduler optimization parameters ranging basically from 0 to infinity. $CQI_{i,k,s}$ is the CQI of user i at PRB k and stream s, and $CQI^{avg}{}_i$ is the average CQI of user i. $T_{tot}(n)$ is the average delivered throughput (during the recent past) to all users ranked in TD stage served by the BS.

The proposed scheduling metric in (2) is essentially composed of three elements affecting the overall scheduling decisions. The first ratio takes into account the transmit power fluctuations in BS for each PRB due to applied SFR scheme. The second ratio is the relative instantaneous quality of the individual user's radio channels over their own average channel qualities in a stream-wise manner. The third ratio is related to measuring the achievable throughput of individual UE's against the corresponding average throughput of scheduled users. The power coefficients α_1 and α_2 are additional adjustable parameters that can be tuned and controlled to obtain a desired balance between throughput and fairness. This will be illustrated in Section 5.

The basic idea of incorporating the sub-band power ratio $P_{k,s} / P_{max}$ into the scheduling metric (2) is that the sub-band power levels obviously affect the link adaptation and thereon the estimated supportable throughput as well as the actual delivered throughput. Thus by taking the power fluctuations and the transmit power levels in BS for each PRB due to SFR schemes into account, we seek higher fairness in the scheduling between truly realized UE throughputs, and thereon better cell-edge coverage. This will be demonstrated in Section 5. Also since the power is understood in a stream-wise manner, this gives somewhat higher priority to single-stream transmissions which also helps in increasing the coverage.

4. Radio System Simulations

4.1. Basic Features

Extensive quasi-static system simulator for LTE downlink providing traffic modeling, multiuser packet scheduling and link adaptation including HARQ is used for evaluating the system level performance of the proposed packet scheduling scheme, following 3GPP evaluation criteria [1]. The 10 MHz system bandwidth case is assumed, being composed of 1024 sub-carriers (out of which 600 are active) and divided into 50 physical resource blocks (PRB) each consisting of 12 sub-carriers with sub-carrier spacing of 15 kHz. Pilot signals are sent from base station to mobile station to determine the instantaneous channel condition. The mobile stations measure the actual channel states and the information is reported to the BS. The actual reported CQI's are based on received signal-to-interference-and-noise ratios (SINR), calculated by the UE's for each PRB. Here the UE's are assumed to use linear MMSE (LMMSE) receivers for MIMO 2x2 system simulation and MRC for SIMO case. Additionally, the UEs always report single-stream SINR as well as both single user (SU) and multi-user (MU) dual-stream SINR's at the corresponding detector output.

In a single simulation run, mobile stations are randomly distributed over standard hexagonal cellular layout with altogether 19 cells each having 3 sectors. As a concrete example, the number of active users in the cell is set to 15 and the UE velocities equal 3km/h. The path losses for individual links are directly determined based on the individual distances between the mobile and the serving base station. On the other hand, the actual fading characteristics of the radio channels are collected each TTI (1ms) and depend on the assumed mobility and power delay profile. Due to the centralized approach statistics are collected only from the central cell site while the others simply act as sources of inter-cell interference. The used MIMO scheme for performance evaluation purposes is per-antenna rate control (PARC) with two transmit antennas at the BS and two receive antennas at each UE The main simulation parameters and assumptions are summarized in Table 1.

The RRM functionalities are controlled by the packet scheduler together with link adaptation and HARQ entities. Link adaptation consists of two separate elements – the inner loop LA (ILLA) and the outer loop LA (OLLA). These are used for removing CQI imperfections, estimating supported data rates and MCS's, and stabilizing the 1st transmission Block Error Probability (BLEP) to the target range (typically 10-20%). Simple admission control scheme is used for keeping the number of UEs per cell constant. HARQ is based on SAW protocol and a maximum number of three re-transmissions is allowed. MIMO functionality requires individual HARQ entry per stream which is also implemented. Link-to-system level mapping is based on the effective SINR mapping (EESM) principle [1].

As a concrete example of unequal PRB power allocation we exploit the case of reducing the transmit power of every third RB by 1dB and the neighbouring ones yet by another 1dB (i.e. the relative power pattern is 0dB, -1dB, -2dB, 0dB, -1dB, -2dB, …) compared to maximum transmit power [18]. Notice that this is just one concrete example selected for evaluating and comparing the performance of different schedulers.

SFR scheme benefits from full bandwidth utilization for each BS and sub-band division helps in mitigating the CCI. The reuse factor used in the simulation scenarios is 3, which corresponds to sub-band division of (17, 17, 16) RB's (total of 50 RB's). Two different power patterns i.e. SFR power masks (described in section 2) define the power levels used for each RB group. In the first case the corresponding power pattern is (0dB, -4dB, -4dB) and in second case (0dB, -1dB, -4dB).

Table 1: Default simulation parameters

Parameter	Assumption
Cellular Layout	Hexagonal grid, 19 cell sites, 3 sectors per site
Inter-site distance	500 m
Carrier Frequency / Bandwidth	2000MHz / 10 MHz
Channel estimation	Ideal
PDP	ITU Typical Urban 20 paths
Minimum distance between UE and cell	>= 35 meters
Average number of UEs per cell	15
Max. number of frequency multiplexed UEs	10
UE receiver	MRC and LMMSE
Shadowing standard deviation	8 dB
UE speed	3km/h
Total BS TX power (P_{total})	46dBm - 10MHz carrier
Traffic model	Full Buffer
Fast Fading Model	Jakes Spectrum
CQI reporting time	5 TTI
CQI delay	2 TTIs
MCS rates	QPSK (1/3, 1/2, 2/3),16QAM(1/2, 2/3, 4/5),64QAM(1/2, 2/3, 4/5)
ACK/NACK delay	2ms
Number of SAW channels	6
Maximum number of retransmissions	3
HARQ model	Ideal CC
1st transmission BLER target	20%
Forgetting factor	0.002
Scheduling schemes	PF, PPF,MPMPF

The actual effective SINR calculations rely on subcarrier-wise complex channel gains (estimated using reference symbols in practice) and depend in general also on the assumed receiver (detector) topology. Here we assume that the LMMSE detector, properly tailored for the transmission mode (1-stream SU, 2-stream SU or 2-stream MU) is deployed. The detector structures and SINR modelling for different transmission modes are described in detail below.

4.2. Detectors and SINR Modeling

4.2.1. Single-Stream SU Case

In this case, only one of the two BS transmit antennas is used to transmit one stream. At individual time instant (time-index dropped here), the received spatial 2x1 signal vector of UE *i* at sub-carrier *c* is of the form

$$\mathbf{y}_{i,c} = \mathbf{h}_{i,c} x_{i,c} + \mathbf{n}_{i,c} + \mathbf{z}_{i,c} \tag{3}$$

where $x_{i,c}$, $\mathbf{h}_{i,c}$, $\mathbf{n}_{i,c}$ and $\mathbf{z}_{i,c}$ denote the transmit symbol, 2x1 channel vector, 2x1 received noise vector and 2x1 inter-cell interference vector, respectively. Then the LMMSE detector $\hat{x}_{i,c} = \mathbf{w}_{i,c}^H \mathbf{y}_{i,c}$ is given by

$$\mathbf{w}_{i,c} = \sigma_{x,i}^2 \sigma_{i,c}^2 (\mathbf{h}_{i,c}^H \sigma_{x,i}^2 \sigma_{i,c}^2 \mathbf{h}_{i,c} + \mathbf{\Sigma}_{n,i} + \mathbf{\Sigma}_{z,i})^{-1} \mathbf{h}_{i,c} \tag{4}$$

where $\sigma_{x,i}^2$, $\sigma_{i,c}^2$, $\mathbf{\Sigma}_{n,i}$ and $\mathbf{\Sigma}_{z,i}$ denote the transmit power (per the used antenna), power mask per sub-carrier, noise covariance matrix and inter-cell interference covariance matrix, respectively. Now the SINR is given by

$$\gamma_{i,c} = \frac{\left|\mathbf{w}_{i,c}^H \mathbf{h}_{i,c}\right|^2 \sigma_{x,i}^2 \sigma_{i,c}^2}{\mathbf{w}_{i,c}^H \mathbf{\Sigma}_{n,i} \mathbf{w}_{i,c} + \mathbf{w}_{i,c}^H \mathbf{\Sigma}_{z,i} \mathbf{w}_{i,c}} \tag{5}$$

The noise variables at different receiver antennas are assumed uncorrelated (diagonal $\mathbf{S}_{n,i}$) while the more detailed modeling of inter-cell interference (structure of $\mathbf{S}_{z,i}$) takes into account the distances and channels from neighboring base stations (for more details, see e.g. [17]). The MRC detector is given by

$$\mathbf{w}_{i,c} = \frac{\mathbf{h}_{i,c}}{\left\|\mathbf{h}_{i,c}\right\|^2} \tag{6}$$

4.2.2. Dual-Stream SU Case

In this case, both of the two BS transmit antennas are used for transmission, on one stream per antenna basis. At individual time instant, the received spatial 2x1 signal vector of UE *i* at sub-carrier *c* is now given by

$$\mathbf{y}_{i,c} = \mathbf{H}_{i,c}\mathbf{x}_{i,c} + \mathbf{n}_{i,c} + \mathbf{z}_{i,c} \tag{7}$$

where $\mathbf{x}_{i,c}$ and $\mathbf{H}_{i,c} = [\mathbf{h}_{i,c,1}, \mathbf{h}_{i,c,2}]$ denote the 2x1 transmit symbol vector and 2x2 channel matrix, respectively. Now the LMMSE detector $\hat{\mathbf{x}}_{i,c} = \mathbf{W}_{i,c}\mathbf{y}_{i,c}$ is given by

$$\mathbf{W}_{i,c} = \mathbf{\Sigma}_{x,i}\mathbf{H}_{i,c}^H(\mathbf{H}_{i,c}\mathbf{\Sigma}_{x,i}\mathbf{H}_{i,c}^H + \mathbf{\Sigma}_{n,i} + \mathbf{\Sigma}_{z,i})^{-1} = \begin{bmatrix} \mathbf{w}_{i,c,1}^H \\ \mathbf{w}_{i,c,2}^H \end{bmatrix} \tag{8}$$

where $\mathbf{\Sigma}_{x,i} = \sigma_{i,c}^2 diag\{\sigma_{x,i,1}^2, \sigma_{x,i,2}^2\} = \sigma_{i,c}^2 diag\{\sigma_{x,i}^2/2, \sigma_{x,i}^2/2\}$ denotes the 2x2 covariance matrix (assumed diagonal) of the transmit symbols. Note that compared to single-stream case, the overall BS transmit power is now divided between the two antennas, as indicated above. Then the SINR's for the two transmit symbols are given by

$$\gamma_{i,c,1} = \frac{\left|\mathbf{w}_{i,c,1}^H \mathbf{h}_{i,c,1}\right|^2 \sigma_{x,i,1}^2 \sigma_{i,c}^2}{\left|\mathbf{w}_{i,c,1}^H \mathbf{h}_{i,c,2}\right|^2 \sigma_{x,i,2}^2 \sigma_{i,c}^2 + \mathbf{w}_{i,c,1}^H \mathbf{\Sigma}_{n,i} \mathbf{w}_{i,c,1} + \mathbf{w}_{i,c,1}^H \mathbf{\Sigma}_{z,i} \mathbf{w}_{i,c,1}}$$

$$\gamma_{i,c,2} = \frac{\left|\mathbf{w}_{i,c,2}^H \mathbf{h}_{i,c,2}\right|^2 \sigma_{x,i,2}^2 \sigma_{i,c}^2}{\left|\mathbf{w}_{i,c,2}^H \mathbf{h}_{i,c,1}\right|^2 \sigma_{x,i,1}^2 \sigma_{i,c}^2 + \mathbf{w}_{i,c,2}^H \mathbf{\Sigma}_{n,i} \mathbf{w}_{i,c,2} + \mathbf{w}_{i,c,2}^H \mathbf{\Sigma}_{z,i} \mathbf{w}_{i,c,2}} \tag{9}$$

4.2.3. Dual-Stream MU Case

In this case, the transmission principle and SINR modelling are similar to subsection 2) above, but the two spatially multiplexed streams belong now to two different UE's, say *i* and *i*'. Thus the SINR's in (9) are interpreted accordingly.

Finally, for link-to-system level mapping purposes, the exponential effective SINR mapping (EESM), as described in [1-3], is deployed.

5. NUMERICAL RESULTS AND ANALYSIS

This section presents the obtained results from the radio system simulations using different PS algorithms combined with SFR schemes as described in the paper. The system-level performance is generally measured and evaluated in terms of:

- Throughput - the total number of successfully delivered bits per unit time. Usually measured either in kbps or Mbps.
- Coverage – the experienced data rate per UE at the 95% coverage probability.
- Fairness per scheduling scheme measured using Jain's fairness index [19]

Initially, we demonstrate the behaviour of the proposed MPMPF scheduler by using different power coefficients α_1 and α_2 and comparing it against other PF scheduling algorithms. Secondly, we illustrate the potential of combining the use of different SFR power masks and again tuning the MPMPF scheduler' power coefficients. To emphasize the role of power-aware and CQI based priority metric calculation in (2), we fix the value of α_2 to 1 and change the values of α_1 as $\alpha_1 = \{1,2,4\}$ [11], [12].

Figure 3 illustrates the average cell throughput and cell-edge coverage for the different schedulers in SIMO (sub-figures (a) and (b)) and MIMO (sub-figures (c) and (d)) system simulation cases. The power coefficient values are presented as index M, where M1 represents the first couple, i.e., $\alpha_1=1$, $\alpha_2=1$, M2: $\alpha_1=2$, $\alpha_2=1$ and M3: $\alpha_1=4$, $\alpha_2=1$. The obtained results with the proposed scheduler are compared with the reference PF schedulers – ordinary PF, power-aware PF described in (1) and MMPF scheduler from [13]. For the cases M1-M3, based on Figure 2, the new MPMPF scheduler achieves coverage gains in the order of 63-91% at the expense of only 15-22% throughput loss compared to ordinary PF in the SIMO scenario (sub-figures (a) and (b)). In the corresponding MIMO (sub-figures (c) and (d)) scenario, we obtain 71% constant coverage gain for the same throughput loss as in previous SIMO case. Compared to the MMPF scheduling principle [16] or to the power-aware PF in (1), similar coverage gains are obtained, as can be read from the figure.

Similarly, Figure 4 illustrates the same performance statistics as in Figure 3 for the different scheduling approaches and different power mask cases. By combining SFR and PF scheduling we achieve small throughput gain, while coverage is increased with 10%. Changing the scheduling method to MPMPF and coefficient values (M1), we achieve coverage gains in the order of 58% at the expense of 20% throughput loss for PM1 case scenario presented in (a) and (b). In PM2 case scenario illustrated on (c) and (d) we obtain nearly the same coverage gain (60%) for small throughput loss (13%).

Continuing the evaluation of the proposed method – SFR combined with MPMPF, we clearly see a trade-off between average cell throughput and coverage for different power coefficient values in PM1 and PM2 cases. Furthermore, the remaining power coefficient values can be used for tuning the overall scheduling performance. For the rest of the cases, the cell throughput loss is decreased stepwise with around 4% per index M, which on the other hand corresponds to coverage gains increase from 74% to 80%. Consequently, an obvious trade-off between average cell throughput and coverage is clearly seen. Similarly, in PM2 case scenario, we obtain coverage gains between 75% to 82% corresponding to throughput losses from 10% to 8% as illustrated by the performance statistics.

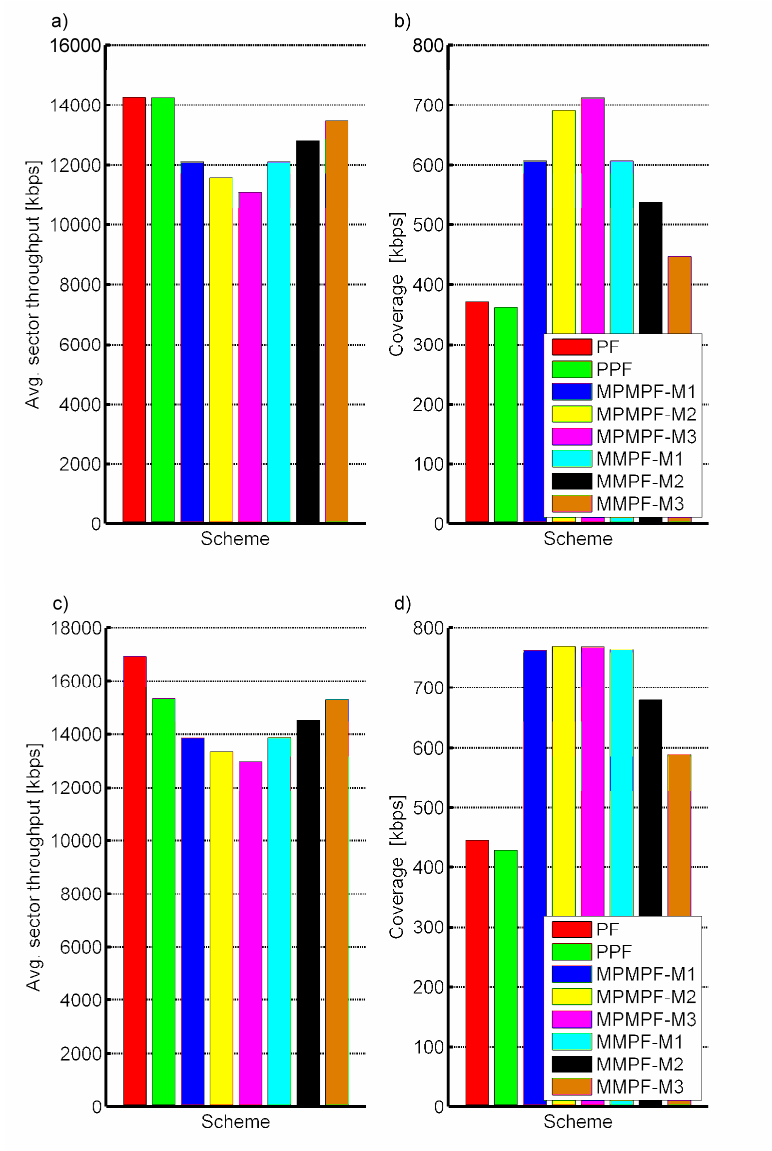


Figure 3: Average cell throughput and coverage gain over the reference PF scheduling scheme for the different simulation cases – SIMO (a, b) and MIMO (c, d).

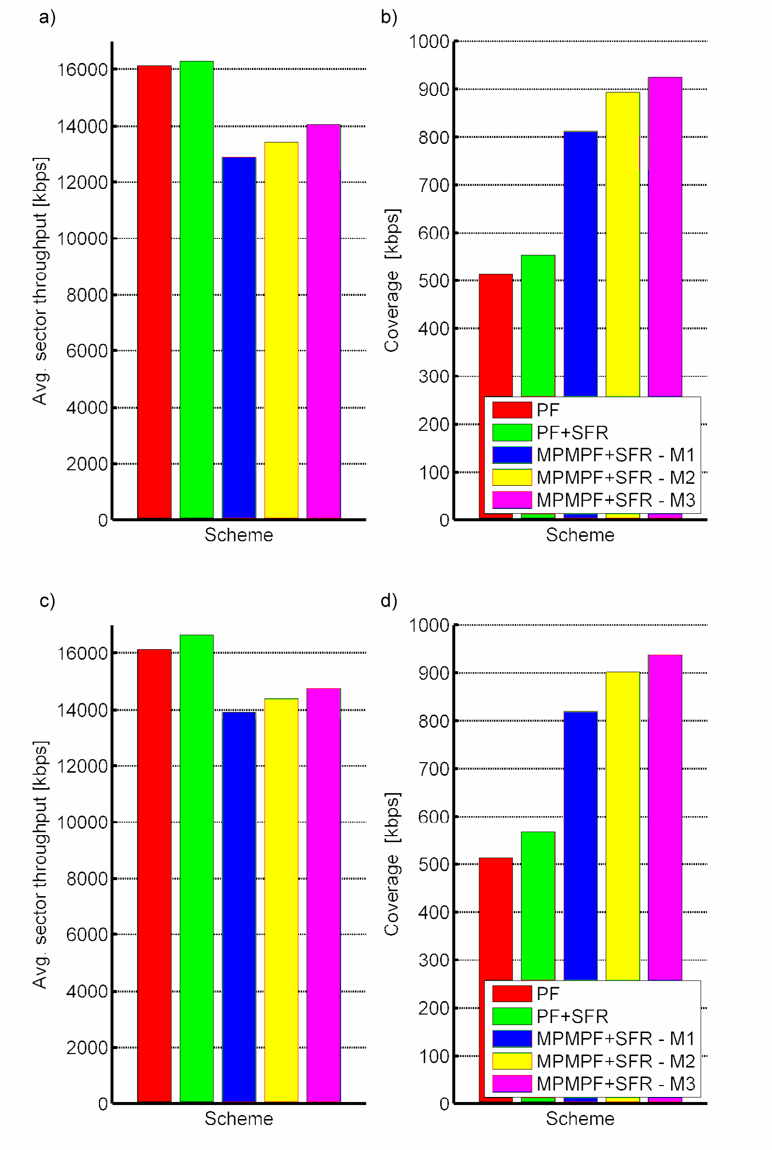


Figure 4: Average cell throughput and coverage gain over the reference PF scheduling scheme and PF +SFR scheme for the different power masks scenarios – PM1 (a, b) and PM2 (c, d). The schemes M1-M3 refer to the MPMPF scheduler with power coefficient values.

Figure 5 illustrates the Jain's fairness index [14] per scheduling scheme for SIMO and MIMO scenarios, calculated using the truly realized throughputs at each TTI for all 15 UE's and over all the simulation runs. The value on the x axis corresponds to the used scheduler type (1 refers to ordinary PF scheduler, 2 refers to PPF, etc.). Clearly, the proposed MPMPF scheduler outperforms all other scheduling algorithms and the received fairness gains are in range of 17%-20% in the SIMO case and 35%-37% in the MIMO case.

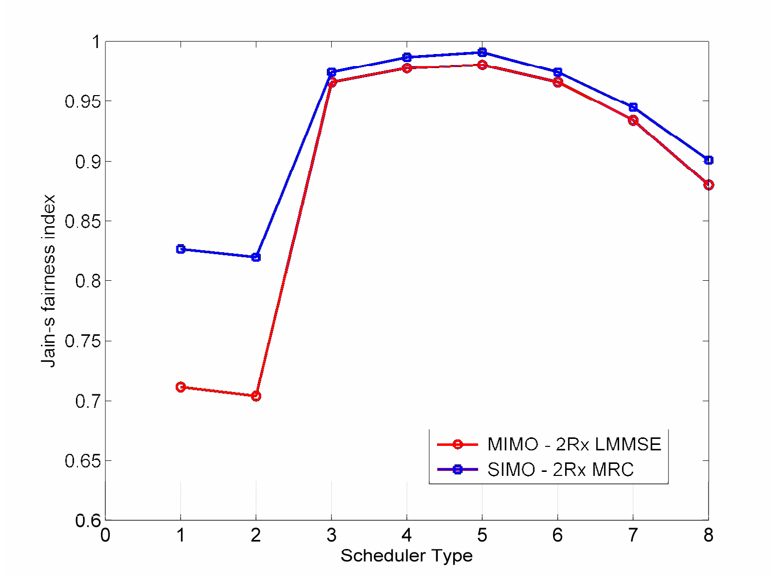


Figure 5: Jain's fairness index per scheduling scheme for SIMO and MIMO cases. Scheduler type 1 means ordinary PF, 2 means power-aware PF from (1), 3-5 mean proposed modified power-aware PF from (2) and 6-8 mean modified PF from [13].

Figure 6 illustrates the Jain's fairness index per scheduling scheme for different SFR power masks scenarios, calculated over all the I_{TOT} = 15 UE's. The value on the x axis corresponds to used scheduler type, where 1 refers to the reference PF scheduler, 2 refers to PF+SFR, 3 refers to the MPMPF+SFR with index M1, etc. Clearly, the fairness distribution with MPMPF+SFR outperforms the used reference plain PF scheduler and PF+SFR for both analyzed cases. The received fairness gains are in range of 17% to 31% in the PM1 case and 18% to 32% in the PM2 case. Compared to the PF+SFR case the corresponding gains are in the range of 15% to 17% for both cases.

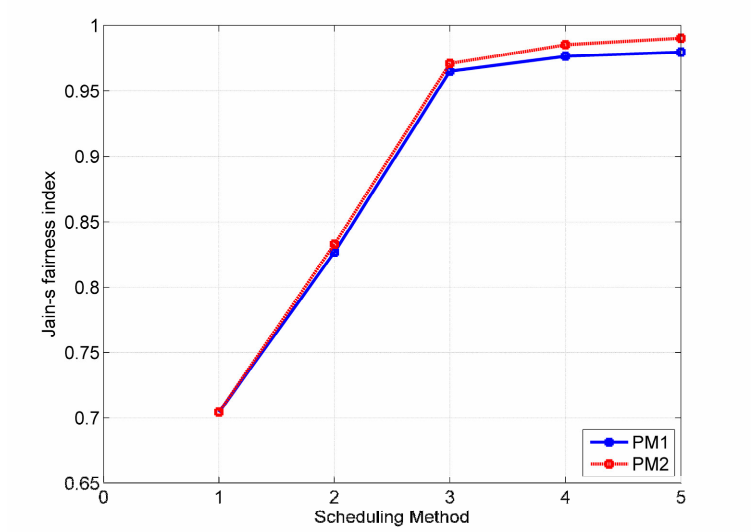


Figure 6: Jain's fairness index per scheduling scheme. Scheduler type 1 means ordinary PF, while 2 means PF +SFR, 3-5 means MPMPF +SFR with power coefficients M1,M2 and M3 correspondingly.

6. CONCLUSIONS

In this paper, a new power-aware multi-stream proportional fair scheduling metric covering time-, frequency- and spatial domains was proposed. Moreover, we have studied the potential of combining soft frequency reuse schemes with advanced power-aware multi-user packet scheduling algorithms in OFDMA type radio system context, using UTRAN long term evolution (LTE) downlink in Macro cell environment as a practical example. The proposed MPMPF metric takes into account the transmit power allocation on RB basis defined by SFR power masks, the instantaneous channel qualities (CQI's) as well as resource allocation fairness. The achievable throughput performance together with coverage and fairness distributions were analyzed and compared against the corresponding statistics of more traditional earlier-reported proportional fair scheduling techniques with SDM functionality in SIMO and MIMO cases, as well as combined with SFR functionality. In the case of fixed coverage requirements, the proposed scheduling metric calculations based on UE channel feedback offers better control over the ratio between the achievable cell/UE throughput and coverage increase, as well as increased UE fairness taking into account irregular BS transmission power. As a practical example, the fairness in resource allocation together with cell coverage can be increased significantly (more than 50%) by allowing a small decrease (in the order of only 10-15%) in the cell throughput for plain PF scheduling case and more than 15% increase over the traditional PF when power consideration are taken into account or combined with SFR.

ACKNOWLEDGEMENTS

The authors would like to thank Prof. Jukka Lempiäinen, Tampere University of Technology, Tampere, Finland, for fruitful discussions.

REFERENCES

[1] "Physical Layer Aspects for Evolved UTRA", *3GPP Technical Report TR 25.814*, ver. 7.1.0, Oct. 2006.

[2] WiMAX Forum™, http://www.wimaxforum.org, Aug. 2006.

[3] ITU global standard for international mobile telecommunications "IMT-Advanced", IMT-ADV/1-E,Mar. 2008.

[4] "Wireless World Initiative New Radio (WINNER)," IST-507581, European Information Society Technologies, Technical Report, Dec. 2006.

[5] N. D. Tripathi et al., *Radio Resource Management in Cellular Systems*, Springer, 2001.

[6] E. Dahlman et al., *3G Evolution: HSPA and LTE for Mobile Broadband*, Academic Press, 2007.

[7] Bohge Matias, et. al., "Optimal Power Masking in Soft Frequency Reuse based OFDMA Networks," in Proc. of the European Wireless Conference 2009 (EW'09), Aalborg, Denmark,May 2009

[8] 3GPP; Huawei, "Soft frequency reuse scheme for UTRAN LTE," R1- 050507, May 2005.

[9] K. Doppler, X. He, C. Witjing, and A. Sorri, "Adaptive soft reuse for relay enhanced cells," in Proc. of the 65th IEEE Vehicular Technology Conference (VTC-Spring 07), Apr. 2007, pp. 758–762.

[10] Y. Sun et al., "Multi-user scheduling for OFDMA downlink with limited feedback for evolved UTRA", in *Proc. IEEE VTC-2006,* Montreal, Canada, September 2006.

[11] N. Wei, et al., "Performance of MIMO with frequency domain packet scheduling", in *Proc. IEEE VTC-2007*, Dublin, Ireland, May 2007.

[12] C. Wengerter, J. Ohlhorst, and A.G.E Von Elbwert, "Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA", in *Proc. IEEE VTC-2005,* Stockholm, Sweden, May 2005.

[13] P. Svedman, S.K.Wilson, L.J. Cimini, and B. Ottersten, "A simplified opportunistic feedback and scheduling scheme for OFDMA", in *Proc. IEEE VTC-2004*, Los Angeles, USA, pp. 1878-1882, May 2004.

[14] S. Y. Kyoung et al., "Power efficient frequency domain packet scheduling for OFDMA systems", in *Proc. Int. Conf. Convergence Information Technology, 2007,* Dhaka, Bangladesh, pp. 1906-1911.

[15] S. Nonchev, J. Venäläinen and M. Valkama, "New frequency domain packet scheduling schemes for UTRAN LTE downlink", in *Proc. ICT Mobile Summit*, Stockholm, Sweden, June 2008.

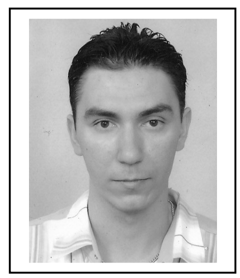
[16] S. Nonchev and M. Valkama, "Efficient packet scheduling schemes for multiantenna packet radio downlink", in *Proc. Fifth Advanced Int. Conf. Telecommunications (AICT-2009)*, Venice, Italy, May 2009.

[17] S. Nonchev and M. Valkama, "A new fairness-oriented packet scheduling scheme with reduced channel feedback for OFDMA packet radio systems", Int. Journal of Communications, Network and System Sciences (IJCNS), 2009.

[18] S. Nonchev and M. Valkama, "Efficient power-aware packet scheduling schemes for multiantenna packet radio systems", in The First International Workshop on Wireless & Mobile Networks (WiMoNe - 2009), December 2009, Chennai (Madras), India

[19] D. Chui and R. Jain, "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks", Computer Networks and ISDN Systems, 1989.

Authors



Stanislav Nonchev was born in Burgas, Bulgaria, on January 14, 1981. He received the M.Sc. Degree in telecommunications engineering (TE) from Tampere University of Technology (TUT), Finland, in 2006. Currently he is a Ph.D student at Tampere University of Technology, Department of Communications Engineering. His research spans radio resource management for ad-hoc and mobile networks, radio network planning, and network protocols.

Mikko Valkama was born in Pirkkala, Finland, on November 27, 1975. He received the M.Sc. and Ph.D. Degrees (both with honors) in electrical engineering (EE) from Tampere University of Technology (TUT), Finland, in 2000 and 2001, respectively. In 2002 he received the Best Ph.D. Thesis -award by the Finnish Academy of Science and Letters for his thesis entitled "Advanced I/Q signal processing for wideband receivers: Models and algorithms". In 2003, he was working as a visiting researcher with the Communications Systems and Signal Processing Institute at SDSU, San Diego, CA. Currently, he is a Full Professor at the Department of Communications Engineering at TUT, Finland. He has been involved in organizing conferences, like the IEEE SPAWC'07 (Publications Chair) held in Helsinki, Finland. His general research interests include communications signal processing, estimation and detection techniques, signal processing algorithms for software defined flexible radios, digital transmission techniques such as different variants of multicarrier modulation methods and OFDM, and radio resource management for ad-hoc and mobile networks.

Tampereen teknillinen yliopisto
PL 527
33101 Tampere

Tampere University of Technology
P.O.B. 527
FI-33101 Tampere, Finland

ISBN 978-952-15-3080-7
ISSN 1459-2045