Guihua Piao

Radio Resource Management for Integrated Services in Multi-radio Access Networks

UNI KASSEL
VERSITÄT

kassel
university
press

This work has been accepted by the faculty of electrical engineering / computer science of the University of Kassel as a thesis for acquiring the academic degree of Doktor der Ingenieurwissenschaften (Dr.-Ing.).

Supervisor: Prof. Dr.-Ing. Klaus David, University of Kassel
Co-Supervisor: Prof. Dr. Holger Karl, University of Paderborn

Defense day: 19th December 2006

Bibliographic information published by Deutsche Nationalbibliothek
Die Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data is available in the Internet at http://dnb.ddb.de

Zugl.: Kassel, Univ., Diss. 2006
ISBN: 978-3-89958-269-7
URN: urn:nbn:de:0002-2697

www.upress.uni-kassel.de

Printed by: Unidruckerei, University of Kassel
Printed in Germany

Acknowledgement

I would like to gratefully and sincerely thank Professor Dr.-Ing. Klaus David for his support, mentor and encouragement throughout my graduate study. I am deeply grateful to him for leading me into the research field of mobile communication networks by admitting me as a student researcher, advising me the master thesis and finally this doctoral work. I have learned from him how to perform scientific thinking and academic work. Without his encouragement and support throughout the graduate study, this doctoral work could not have finished.

I would like to thank my co-advisor Professor Dr. Holger Karl for spending his valuable time to discuss this dissertation with me and for providing invaluable comments to improve it. Many thanks go also to Professor Dr. sc. techn. Dirk Dahlhaus and Professor i. R. Dr.-Ing. Siegbert Hentschke, for willing to be the rest of my thesis committee, reading this dissertation and asking insightful questions.

The colleagues in ComTec deserve also my appreciation for technical or personal helps. I would like to especially thank Matthias Hildebrand for his assistance and guidance in getting my research work started on the right foot. Although his early leaving from ComTec after his graduation was a great distress to me, I have benefited a lot from his wisdom and software development knowledge. I also would like to thank Stephan Sigg, Niklas Klein, Rico Kusber, Alexander Flach, Andreas Pirali, Thomas Hohmann, Sian Lun lau and Nermin Bruglja for their reviews of parts of this thesis.

I would also like to thank Ingo Karla and Rolf Sigle in Alcatel SEL AG, with whom I have explored the requirements of the simulator, the algorithms and the analysis of the simulation results.

Finally, I would like to thank my family. My parents receive my deepest gratitude for their faith in me and for taking care of my son. My son, Yifan, who was born just after I started my graduate career, brings me a lot of happiness and motivates me to become a person whom he will be proud of. Most importantly, I would like to thank my husband, Yonghua, for his encouragement and patience throughout the days I studied and worked in Germany. Without the inspiration of my family, I could not have completed my graduate study.

Abstract

Several disparate types of radio access networks are currently active in the wireless communication environment. For example, GSM has achieved great success via global coverage and more than two billion subscribers worldwide since June 2006; UMTS is on the way of worldwide deployment and competes with other technologies to fulfil the increasing demand of high speed real-time and non-real-time services; WLAN has become successful in non-real-time service provision since several years and is currently redesigned to provide real-time services. New up-and-coming radio access technologies, e.g. WiMAX, will contribute to the wireless communication as well. However, these various radio access networks are currently separate, or only loosely coupled with each other if they belong to a same operator. The next generation wireless network is envisaged to be based on various complementary and cooperating radio access networks in a heterogeneous environment. How to utilise radio resources for multiservices in a coordinated way is a major issue in such multi-radio access networks. This thesis presents the multi-radio resource management which aims to optimise the capacity as well as the service quality via initial multi-radio access and intersystem handover.

Two of the challenges met by the multi-radio resource management (MRRM) are the development of decision algorithms and mechanisms for information transfer to these decision algorithms. This thesis firstly contributes an MRRM algorithm in a multiservice scenario, which bases the decision on comparisons of load, QoS requirements, and resource efficiency in different access networks. According to the diverse requirements of various services, the algorithm aims to optimise the user data rate of NRT services whereas as well as to minimise blocking and dropping of RT services, both with limited intersystem handovers. The main input parameter is the cell load information, which is meticulously defined for different types of services and well aligned with the formats of information elements of the legacy 3GPP intersystem signalling. This type of MRRM is then network-based, either centralised in the core network or distributed in each underlying network. This thesis secondly contributes an information gathering mechanism where the distributed architecture is concerned. In our proposal the MRRM entity stores the load information obtained via the latest intersystem signalling for later evaluation. This information gathering mechanism is expected to mitigate the MRRM delay at the expense of information inaccuracy and in turn capacity degradation.

The system performance is evaluated through simulations. The simulations include lots of details about the traffic characteristics, the user mobility, as well as the radio characteristics and the traditional radio resource management in accordance with various radio access technologies. Especially, GERAN, UTRAN, and WiMAX are modelled. The voice service in GERAN and UTRAN simulators with some basic mechanisms of radio resource management is

verified by analytical models, respectively. The proposed centralised and distributed multi-radio resource management for integrated services is evaluated in a GERAN/UTRAN heterogeneous scenario. The MRRM for non-real-time services and a newly proposed joint preference and user data rate optimised MRRM is further evaluated in a GERAN/UTRAN/WiMAX heterogeneous scenario.

Zusammenfassung

Verschiedene Funkzugangsnetze werden derzeit zur drahtlosen Kommunikation betrieben. GSM hat beispielsweise großen Erfolg bezüglich der globalen Netzabdeckung und mehr als zwei Milliarden Teilnehmer weltweit seit Juni 2006. UMTS ist auf dem Weg zur weltweiten Verbreitung und konkurriert mit anderen Technologien um die zunehmende Nachfrage nach Bandbreitennungrigen RT- und NRT- Diensten zu erfüllen; WLAN ist für NRT-Dienste seit einigen Jahren erforgreich und wird für RT-Dienste weiter entwickelt. Außerdem tragen neue, vielversprechende Funkzugangstechniken, wie z.B. WiMAX, zur drahtlosen Kommunikation bei. Jedoch sind diese verschiedenen Funkzugangsnetze derzeit voneinander getrennt, bzw. bestenfalls lose gekoppelt. Die nächste Generation drahtloser Netze wird voraussichtlich aus verschiedenen untereinander interagierenden und sich ergänzenden drahtlosen Zugangsnetzen in einer heterogenen Umgebung bestehen. Die effiziente Ausnutzung der Netzwerk-Ressourcen für Multi-Dienste ist eine der Kernaufgaben in solchen heterogenen Zugangsnetzen. Die vorliegende Dissertation behandelt das „Multi-radio Resource Management“ (MRRM). Als Schwerpunkt werden die Maximierung der Kapazität sowie die Optimierung der Dienstqualität über „Initial Multi-radio Access“ und „Intersystem Handover“ behandelt.

Zwei der Herausforderungen, für die Multiradio Resource Management (MRRM) Lösungen vorsieht, sind die Entwicklung der Entscheidungsalgorithmen und die Übermittlung der Informationen für diese Algorithmen. Die vorliegende Dissertation schlägt MRRM-Algorithmen in einem Multidienste-Szenario vor, die Entscheidungen aufgrund von Vergleichen der Last, der QoS-Anforderungen und der Ressourceneffizienz in den unterschiedlichen Netzen treffen. Gemäß den jeweiligen Anforderungen der unterschiedlichen Dienste, zielen die Algorithmen darauf ab, die Übertragungsgeschwindkeit der NRT-Dienste zu optimieren und das Blockieren bzw. den Abbruch der RT-Dienste zu verringern. In beiden Fällen ist die Anzahl der Intersystem-Handover zu minimieren. Die Algorithmen berücksichtigen vornehmlich die Lastinformationen in den einzelnen Zellen. Die Definitionen für die Zellenlastinformationen unterscheiden sich gemäß der Art des Dienstes. Die Formaten der Informationen richten sich nach den Formaten der Informationselemente, die die 3GPP Intersystemsignalisierung vorsieht. MRRM ist, durch das Netz kontrolliert, entweder zentralisiert im Kernnetz angesiedelt oder verteilt auf die einzelnen Netze. In Bezug auf die verteilte Architektur schlagen wir vor, dass MRRM die durch die letzte Intersystem-Signalisierung erlangten und gespeicherten Lastinformationen nutzt. Es ist zu erwarten, dass dieser Mechanismus zur Informationsgewinnung die Verzögerungen durch MRRM verkleinert, jedoch aufgrund ungenauerer Informationen zu einer Verringerung der Kapazität führt.

Die Systemleistung wird in Simulationen getestet. Diese Simulationen berücksichtigen Charakteristika des Daten- und Sprachaufkommens, die Benutzermobilität, spezifische Eigenschaften der Zugangstechnologien sowie das traditionelle Radio Resource Management. Besonderer Fokus wird auf die Modellierung von GERAN, UTRAN und WiMAX gelegt. Die GERAN und UTRAN Simulatoren, die Sprachverkehr mit einigen grundlegenden Mechanismen des Radio Resource Management modellieren, werden jeweils auch in analytischen Modellen verifiziert. Das vorgeschlagene zentralisierte und verteilte MRRM für integrierte Dienste wird in einem heterogenen GERAN/UTRAN Szenario ausgewertet. Das MRRM für NRT Dienste und ein neu vorgeschlagenes auf Präferenz- und Nutzer-Datenrate optimiertes MRRM werden außerdem in einem heterogenen GERAN/UTRAN/WiMAX-Szenario ausgewertet.

Content

Chapter 1

Introduction

1.1 Multi-radio Resource Management

In addition to conventional voice telephony, various data services like web-browsing or Email are becoming more and more popular in wired networks. As a result, wireless networks also should support such a mix of services, also called multi-services. The 3G radio access technologies (RAT), e.g. UMTS and CDMA-2000, are designed to offer high quality broad band streaming and data services. The 2G systems, e.g. GSM, have been deployed world wide and achieved great success. Some enhanced technologies are therefore introduced, aiming to provide high data rate that are considered realistic to further facilitate the current 2G infrastructure. GPRS and EDGE are among these enhanced technologies.

Based on a single carrier modulation scheme, the higher the data rate is, the shorter the data symbol duration would be. Multipath spreading in the received data symbol may cause severe inter-symbol interference and consequently severe bit error rate. In an Orthogonal Frequency Division Multiplexing (OFDM) system, a user data can be split into N pieces and be transmitted on N orthogonal subcarriers in parallel. The split data can be recombined at the receiver side. This technique has increased the immunity to multipath spreading since the transmitted bit rate per subcarrier is actually low. Thus OFDM is a promising technique for high data rate wireless services. The WiMAX forum [1] published the system profiles that employ OFDM technology based on IEEE 802.16 standard [1]-[4].

Various access networks are characterised by the air interface protocols. The access network for GSM is called GERAN while the counterpart for UMTS is called UTRAN. WiMAX is based on a set of air interface protocols standardised by IEEE 802.16. These radio access networks have heterogeneous physical frequency bands, modulation and channel coding schemes, and medium access protocols. Due to the diversity of access technologies, the channel conditions and the resource consumptions on different air interfaces are different even for a same call request.

It is envisaged that the future wireless network environment will be naturally heterogeneous [5]-[6]. Multi-radio access technologies will coexist, compete and cooperate. Consequently, multi-radio resource management (MRRM) will be a salient feature of heterogeneous wireless networks in the future. Multi-radio resource management (MRRM) aims at dynamically allocating multiservices into the most suitable network to win the capacity gain of the heterogeneous networks as well as to provide users high quality of services. The allocation at call setup is called initial RAT selection, whereas for a mobile terminal already being in connected mode, the allocation is called intersystem handover (also called vertical handover). A survey of MRRM functionalities and design criteria can be found in [7].

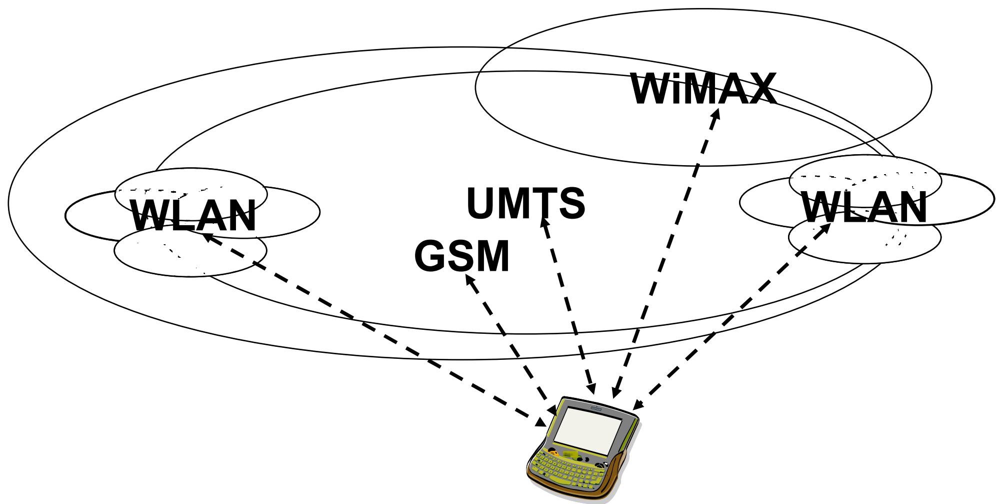

Figure 1-1: A scenario for heterogeneous radio access networks

Optimal selection of a radio access technology and seamless intersystem handover are more complex than in a homogeneous wireless network, due to the diversity of access technologies, wireless networks providers, and service requirements. As compared to handovers in a homogeneous wireless network, which normally take the signal strength as the main metric, intersystem handovers need more metrics. Multiple events may trigger an intersystem handover. These events are such as coverage, packet loss, overload in the source system, load balancing, or user preferences, etc. A decision of an intersystem handover may be based on a number of criteria or policies, such as handover statistics, network resource optimisation, service related requirements, etc. The identification and selection of appropriate comparison methods for signal strength, load, resource efficiency, and cost among disparate access networks are also challenges for intersystem handovers and initial RAT selection.

On the current product market, a GSM/UMTS dual mode mobile station, e.g. Nokia E61 [8], is able to selectively access either of the networks according to the user's preference. Some GPRS/UMTS/WLAN tri-band PCMCIA cards have realised the initial RAT selection for data services also, e.g. tri-band PCMCIA cards offered by T-Mobile and Vodafone in Germany since the beginning of 2005 [9]. Although the physical ability of multi-radio network access is ready, the support for automatic MRRM without manual configurations during active period or even at call setup phase is not commercially available at present.

In the 3GPP standard, the GSM system and the UMTS system are designed to be integrated [10]-[11]. The circuit switched calls are therefore possible to make blind handovers between GERAN and UTRAN. Blind handovers are corresponding to intersystem handovers or direct retries for initial RAT selection triggered only when the channel quality in the source RAT is not satisfied. The target RRM allocates the call handed over to the cell with the best signal strength in the conventional way. This strategy relies on signal strength only, but no load information. If the target cell selected has not sufficient resources either, the intersystem handover will be failed.

The multi-radio access with a tri-band PCMCIA card mentioned above puts in fact the WLAN with the first privilege. Such a card will first associate to the WLAN if it detects a WLAN access point since currently the WLAN is considered to have broader bandwidth than the other two networks. Strictly speaking, a tri-band PCMCIA card realises only multi-radio access at the call setup stage, i.e. on a per-session basis. The radio condition of a moving terminal normally varies dynamically. Even of a stationary user terminal, the surrounding objects can possibly be non-stationary and consequently change the radio condition of the user terminal. More over, the load status may also be varying in a shared medium system. An optimal resource allocation should adapt dynamically to the system status during a session.

A blind handover deals normally with services that have specific requirements for quality, e.g. real-time circuit switched (CS) services, whereas a tri-band PCMCIA card deals normally with non-real-time packet switched (PS) data services. It is well known that CS and PS services have different quality of service requirements. CS services are delay sensitive and error or short interruption tolerant. Thus the resource allocation strategy for a CS connection is by allocating constant bandwidth throughout the call duration; otherwise the CS connection will be blocked. An ongoing CS connection may be dropped once the allocated resources have to be reduced and then the quality of service cannot be guaranteed. In contrast, PS services with bursty nature usually do not impose stringent delay constraint. Erroneous received data, if detected, will be recovered by retransmission control, either at the link layer on a hop-by-hop basis or at the transport layer on an end-to-end basis. Thus the resource allocation for a PS connection allocates resources whenever necessary, i.e. only during the active duration. Besides, it is desired to reduce the delay and to increase the user data rate.

1.2 Problems

A comprehensive multi-radio resource management (MRRM) that deals with integrated CS and PS services is currently not practically realistic in heterogeneous wireless networks. Concerning the service dependent quality requirements and the technology dependent network capabilities, a comprehensive multi-radio access control over multi-radio access networks is non-trivial. Such a comprehensive multiservice MRRM should optimise the integrated network capacity and the QoS of both CS and PS services under the constraints of network resources.

Network-based or mobile-based MRRM is a question. Network operators are willing to control the load distribution to pursue high revenue while end users are willing to have the initiative. Assume there is a user controlled algorithm. The user terminal (a multimode mobile station) are usually able to measure the link quality of various radio technologies; A mobile station can either obtain the knowledge about network load situation and the possible amount of resources to be assigned via additional signalling, or it most likely completely lacks this type of knowledge. Thus a user controlled multi-radio access aligns more with the individual user's satisfaction than the comprehensive overall users' satisfaction and the network satisfaction. To evaluate the benefit of user-based and network-based control is out of the scope of this thesis. This thesis concentrates only on the network based multi-radio resource management.

Moreover, the realisation of an optimisation algorithm for the multiservice MRRM is also non-trivial. What is the interplay of multiple services? Should the management of intersystem handovers take place in IP layer or in radio resource management layer (layer 3 in cellular networks)? Should an end to end QoS be considered or only air interface QoS? What is the impact of underlying networks on the overall system performance?

Another challenge is the implementation of MRRM. Should this functionality be centralised or distributed? What is the necessary input information for the MRRM decision? How is the information collected?

With respect to a centralised MRRM, whatever information about the source system and the candidate systems needed for the MRRM decision should be gathered via additional signalling or encapsulated in the header of data packets. The signalling procedure should be fast enough in order not to bring too much latency. The size of the control header should be small so that it will not waste too much scarce bandwidth. An intrinsic drawback of the control header aiding MRRM decision is the ping-pong transmission of the data packet when the intersystem handover request is failed to be determined by the MRRM algorithm.

As for a distributed MRRM, at least the information about the source system is valid at the MRRM entity. The relevant information transfer can be avoided. Nevertheless the delay-related and bandwidth-related issues remain same as in the case of centralised MRRM.

1.3 Related Work

The essence of multi-radio resource management is the network selection among multiple available radio access networks. If the network selection happens before a connection establishment, it is called initial RAT selection. If an established connection is detected not any more the optimal one, a network selection decision may lead to an intersystem handover, also called vertical handover widely. In some literatures, initial RAT selection and vertical handover are sometimes not distinguished and vertical handover is used to represent the result of network selection in general. Some researchers, aiming to study the mobility management for a heterogeneous wireless access environment, explore also the issues relevant to vertical handover.

1.3.1 Modified Mobile IP for Mobility Management

Mobile IP [12] is considered to be able to resolve the mobility management issues for a heterogeneous wireless environment. Since some wireless access networks have their own addressing scheme inside the access network domain, the traditional mobile IP need some modifications in order to manage the mobility in a heterogeneous wireless environment.

An edge router is proposed in [13] (reported by EU project IST SUITED) to connect each of the gateways of the underlying access networks. Unlike the mobile IP in a fixed network which routes data packets directly to the mobile node, data packets will be routed to the edge router which corresponds to the selected access network. Data transport in every access network domain remains unchanged except the mobile terminals. The mobile terminal includes an additional functional unit – terminal interworking unit – which handles vertical handover initialization and decision using fuzzy logic. Although the user perceived QoS against a threshold is taken as one of the input parameter for the fuzzy logic deployed, the comparison of QoS concerned by various access networks is not available.

Several recent vertical management techniques are also based on modified Mobile IP, summarised in [14]. The modifications arise from the consideration of reducing the communication delay between the mobile node, the core network, the home agent and the foreign agent involved in the Mobile IP protocol, or from the necessary information exchange between the source access router and the target access router. The heterogeneous wireless access networks are usually considered to be WLAN integrated to cellular networks.

1.3.2 Policy-based Automatic Mobility Management

A policy-based framework for admission control is proposed by IETF [15]. The framework proposed contains two architectural elements: a policy decision point (PDP), where policy based decisions are made, and a policy enforcement point (PEP), where policy decisions are enforced actually. Policy is the combination of rules and services where rules define the criteria for resource access and usage. The original motivation is to provide QoS in Internet through the admission control.

Some researchers adopt the policy-based framework for mobility support in heterogeneous networks [16]-[18]. To implement the policy-based framework, one should first determine where to locate the two architectural elements. The authors of [16] accept both scenarios, either in the network or in a mobile terminal. Nevertheless they propose to locate the PDP and PEP in a same network node for both scenarios. The vertical handover is either network-based or mobile-based with respect to the PDP and PEP locations. The performance evaluations of the two scenarios are not given. The vertical handover in [17] is in principle mobile-based. The PEP is proposed to be located in mobile terminals. However, the functions of PDP are distributed at the network side and the host side, since the complexity of the raw policy set might be beyond the ability of the maintenance of a mobile device. In [18] (reported by EU project IST EVEREST), the policy decision function and PEP are located at the network side.

From the perspective of the locations of the PDP and PEP alone, it has encountered so many disparate solutions. The handover decision algorithm developed is even more disparate. A vertical handover decision algorithm is commonly believed to be more complex than a horizontal (intrasystem) handover. In conjunction with the signal strength (the single metric in a horizontal handover), service type, network conditions, mobile node conditions, operator preference, user preference, user profile and energy status should also be considered as input parameters. However, depending on the location of the decision entity, the availability of parameters and the levels of available parameters are different. The network bandwidth parameter takes a static network-dependent value in [16]. Both network side and host side information can be better gathered, with the ability of policy evaluation at the network side in [17]. However the authors focus on the architecture and the semantic implementation of policy rules and policy evaluation. Considering the protection of the security of end users, the variation of user profile in various network domains, and the users' possible unwillingness to reveal the preference to network, the authors of [19] (from EU project IST MobyDick) recommend the mobile-controlled handover decision. To optimise the decision, some network capability information need to be disclosed to mobile terminals. A layer 2 broadcast signalling is proposed to deliver the dynamic load of a network. In summary, the handover decision will encounter the challenges of the sufficiency and the efficiency of parameter gathering and evaluation.

1.3.3 Methodology

The methodology of the handover decision algorithm can be summarized into: cost function ([16], [32]), fuzzy logic ([13], [34]-[35]), and policy model ([17], [29]). The cost function is normally useful when the specified parameters (both network and mobile sides) can be quantified. The fuzzy logic is suitable when the input parameters are not precise and only a few levels of output, e.g. low, medium, and high, are expected. These two methodologies are widely utilized in decision algorithms. The policy model is initially used in the management of distributed systems. It is implemented for the mobility support among overlay networks by [17]. The policy model is comprised of policy specification and policy evaluation. It can be simply understood as that if such events occur under such conditions, a specific action should take place. This policy model is well suitable for automatic configuration of mobile terminals. A policy editor enables even policy enhancement dynamically with a high-level declarative language. Since the architecture of the policy model as well as the group of definitions associated with the policy translation and evaluation in [17] are complex, the downside of this policy model may lie on the difficulty of the popularisation and the standardisation. However, a policy model can become realistic if fewer input parameters are involved, or even better, it uses only static parameters. For example, case studies of the common radio resource management in [29] have only the service type and an indicator for indoor or outdoor as the input parameters.

1.3.4 Network-based Common Radio Resource Management

The studies [12]-[17] mentioned above have not taken the possible cooperation of the access networks into account. This network cooperation should enable some levels of system information sharing and consequently benefit the vertical handover decision. 3GPP has published a standard related to the integration of WLAN and UMTS [20]. With regard to the level of integration, the interworking between WLAN and UMTS is categorised into: open coupling, loose coupling, tight coupling and very tight coupling. The tight coupling requires an emulator of RNC in the WLAN network, namely the interworking unit (IWU). This emulator has an interface connecting to the common UMTS core network, i.e. an interface to SGSN. The very tight coupling integration regards a WLAN access point as a counter part to a UMTS Node B, which has also a connection to the RNC. The deep level of integration in the tight coupling and the very tight coupling provides the common core network between the access networks and external networks. Hence the cooperation between a WLAN and a UTRAN becomes available. The complexity of implementation increases with the level of interworking from the open coupling solution (the simplest one) up to very tight interworking. A higher degree of interworking involves a higher impact on UMTS network elements, which have to provide the necessary capacity with respect to processing power and interface capabilities.

The future heterogeneous access networks can be envisaged to be deeply integrated. Especially the tight coupling approach is very promising since it provides the opportunity to separate the local radio resource management (to keep in the pair of RNC/IWU and the mobile terminal) and the common radio resource management (a centralised or distributed management maintaining interfaces to the RNC/IWU in the underlying networks). The EU project IST WINNER designs the architecture for cooperative RRM which resides in the future network and coordinates the specific RRM entities in the legacy networks [21]. The cooperative RRM logical entity is required to have two components: layered scheduling and intersystem QoS management controller. The former one couples the radio interfacing variants to the cooperative RRM entity that includes intelligent schedulers at node Bs and access points and a multi-RAN (radio access network) scheduler at the cooperative RRM entity. The intersystem QoS management controller aggregates various vertical handover parameters. The IST EVEREST project has put the common radio resource management (CRRM) function into a Wireless Quality of Service Broker which has interfaces to the underlying local RRMs and a Bandwidth Broker for the QoS provision in the IP core network [22]. The degree of the interaction between the common radio resource management and the local radio resource management is envisaged from low to intermediate and ultimately to high level [23]. In the low interaction degree, the CRRM entity transfers the local RRM measurement to the alternative RATs. In the intermediate interaction degree, the initial RAT selection and the vertical handover decision are made in the CRRM entity based on the measurements gathered from the underlying RATs. The high interaction degree is envisaged to be the long term solution of the common radio resource management, which is linked to the perspective with the introduction of truly reconfigurable terminals and networks. The high degree of interaction is also foreseen to supply the feasibility of the scenario where a terminal supports a service with several simultaneous connections on various air interfaces. The current research about CRRM, also referred as MRRM (multi-radio resource management), mostly focuses on the intermediate solution.

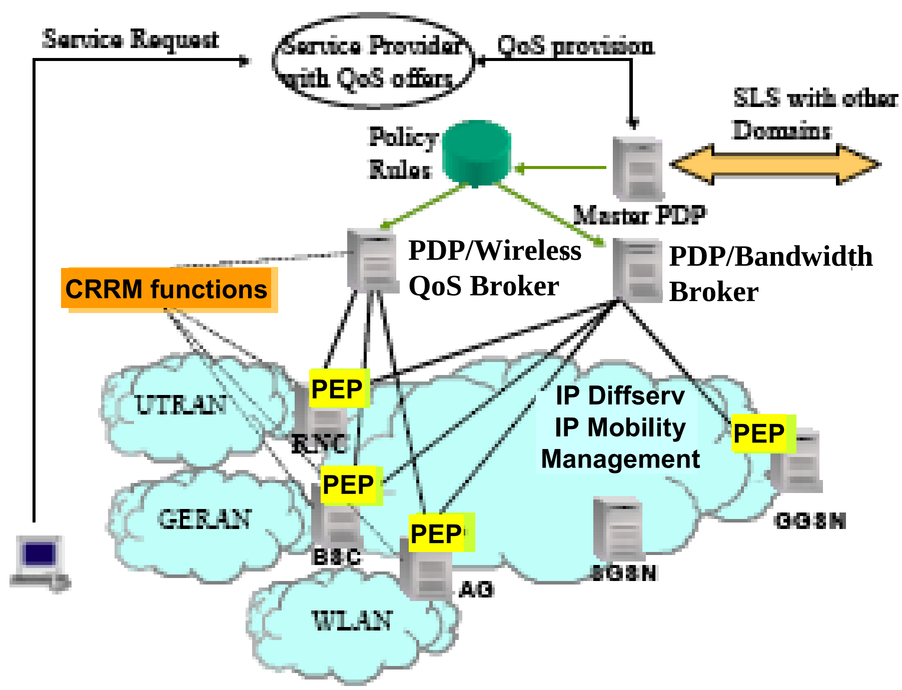

Figure 1-2: Network-based CRRM and QoS architecture [22]

1.3.5 MRRM Strategies

In 3GPP standardization [10]-[11] intersystem handover (IS-HO) protocols between GERAN and UTRAN were given. The intersystem handover corresponds to blind handover, which is requested by the source RAN. Via the signalling through the core network, the necessary information about the handover call and the system information about the source RAN are transferred to the target RAN. Most of the information is transparent to the core network. Upon receiving the request, the target network decides to accept or reject the intersystem handover. The target network treats the intersystem handover request the same as a normal call request originated in the own network. A consequent successful or erroneous intersystem handover decision is replied to the source RAN. This signalling protocol can be directly adopted for any advanced MRRM that might have intelligent algorithms to trigger IS-HOs, i.e. which call at what time, as well as to whether to make the IS-HO decision and from where to where. Nevertheless, the MRRM algorithm is undefined in the 3GPP specifications and is left to the manufactures and operators to decide.

Load Balancing

It is shown that there is a trunking gain of capacity by sharing the resource pool from several overlapping layers/networks, via simulations assuming identical capacity for every overlapped layer/network [24]-[25]. Network-controlled common radio resource management is investigated. Having the knowledge about the load status of the complete overlay networks, CRRM is able to allocate a call to the least loaded cell, i.e. load balancing. The CRRM procedure for RT services is given in Figure 1-3. The CRRM procedure for NRT services is given in Figure 1-4.

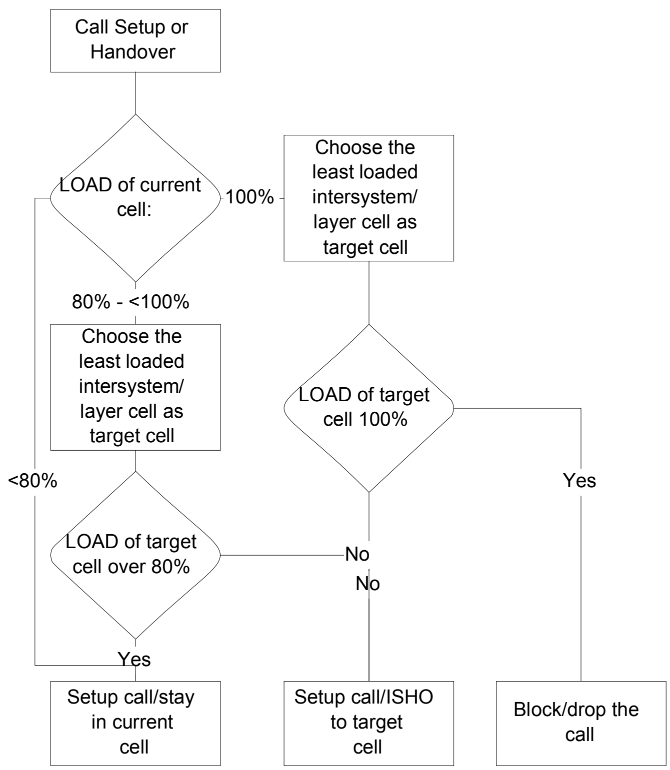

Figure 1-3: Call setup and traffic reason HO procedure with CRRM [24]

The capacity gain of pure RT service and pure NRT service are reported separately. For RT service, a load threshold (80%), where load is in terms of bandwidth, is used to trigger the RAT reselection by the source cell. With the knowledge of the target cells load level in several layers/networks, CRRM selects the least loaded target cell. The higher the RT service bit rate required, the more the trunking gain can be obtained using this method as compared to the reference case where the networks/layers are isolated and no CRRM are applied at all. Additionally, the trunking gain is increased with the increasing layers. For NRT service, a new call is directed to the least loaded cell, where the load is in terms of delay for interactive services or in terms of average bandwidth for download services. The additional intersystem network controlled cell reselection is found to be able to achieve further capacity gain.

However, for simplification, radio network specific features are neglected. They assume that a RT call requests a fixed bandwidth whereas a NRT call is assigned the total bandwidth but queues in a first in first out buffer. The diverse radio conditions of multiple users are thereby ignored. The networks in different layers do not differentiate themselves. Hence the resource consumption in different networks/layers is completely same. If the variant networks features are considered layer by layer, a call perceives bad quality of service in this network may obtain much better quality in an alternative network simply because of the multi-technology diversity. This is not investigated in [24]-[25].

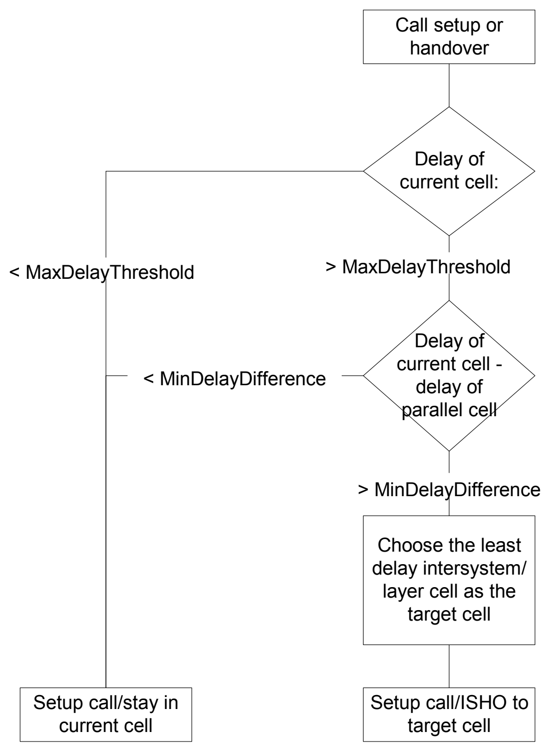

Figure 1-4: Call setup and delay reason HO procedure with CRRM

IST EVEREST project has investigated also the suitability of load balancing [26]. Nevertheless the investigation is based on realistic GERAN and UTRAN simulator modelling the physical variation of the two RATs. The load is then defined as the total resource consumption varying in the types of resources in the two networks. In GERAN, the load factor corresponds to the percentage of total time slots consumption; in UTRAN, the load factor corresponds to the percentage of total power consumption. Thus, unlike the load factor in [24]-[25], the load factor in this investigation is not service type dependent. Since only two radio access networks are considered, a call is allocated to the less loaded networks. Only initial RAT selection is investigated through their load balancing algorithm. The selected network allocates a call to a target cell via the regular call set up procedure in the individual network.

The performance of NRT services depends on a complex set including the RT users, NRT users, and the interference etc. According to the load definition there, a less loaded cell does not guarantee a better performance of the NRT services. Due to the non-linear relation between their load factor and the performance of NRT services, data services cannot benefit from this algorithm. Their simulation results have also shown that the overall throughput gain is at the expense of interactive traffic performance. Hence it is necessary to define a parameter which can better represent the load level of PS services.

Resource Efficiency

In [27] the optimum service mix (the allocation of multiservice in multi-radio access networks) is derived using mathematical formulae to optimise the capacity of multi-radio access networks. The optimisation is performed under the given capacity region of the underlying subsystems. The desired service allocation is found characterised by the fact that the relative efficiency of supporting services is equal in subsystems. For example, given same number of users of service 2, service 1 should be allocated to the system that supports more users of service 1. The time dynamic of radio channel and traffic is neglected.

Following the principle of that allocating the multiservice to the most resource efficient network yields the maximum capacity of multi-radio access networks, the authors of [28] have studied the capacity gain in a realistic scenario of multi-radio access network compromising GSM/EDGE and WCDMA. The optimisation algorithm proposed is to allocate the service to the minimum resource consumption normalized by the total resource of a subsystem at a call arrival. The radio consumption is randomly generated using a statistical log-normal distribution. By allocating each call to the most resource efficient RAT, the optimum capacity of multi-radio access networks can be reached. One difficulty with this approach is the accuracy of resource estimation, e.g. an accurate estimation for a UMTS call is currently not realistic. Even with an accurate estimation, the resource consumptions in various networks are different, hence making the evaluation of resource efficiency difficult. Moreover, the capacity gain is unclear if the stochastic characteristics of traffic and mobility are involved.

Here we take an example result from [28], see Figure 1-5, in order to further illustrate the service allocation principle. Given the capacity limits of GSM/EDGE and WCDMA, WCDMA is more efficient in supporting the data service whereas GSM/EDGE is more efficient in supporting the voice service. Hence the data service is as much as possible allocated to WCDMA whereas the voice service is as much as possible allocated to GSM/EDGE. The Combined Best draws the capacity limit of this service allocation. The Combined Worst is the result from the opposite allocation principle. Equal service mix in the two subsystems results in the capacity in between.

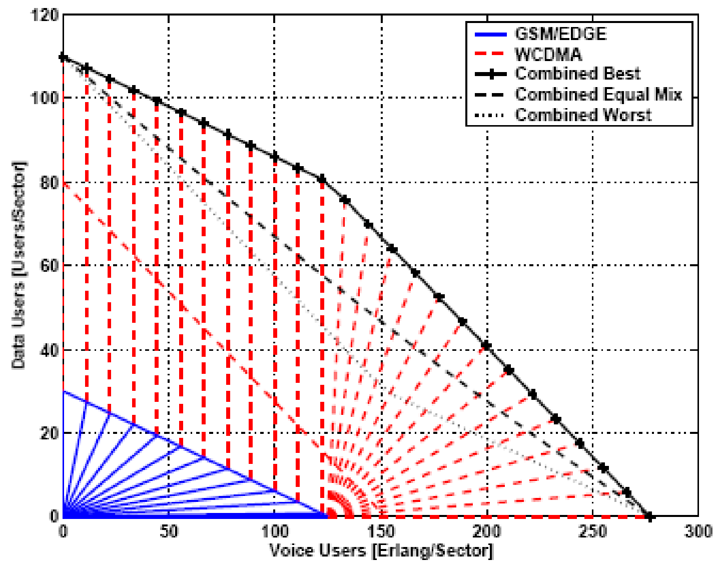

Figure 1-5: Combined capacity region of various service mixes [28]

Service-based

Feasible policy-based RAT initialization strategies are studied in literatures. A common radio resource management function selects an initial RAT to allocate a call request based on a set of input parameters and a set of predefined rules. The service type as an input parameter is especially studied in [29] for a GERAN/UTRAN multi-radio access scenario. The initial RAT selection based on types of services is called service-based policy. For a certain type of service, a prioritised list of RAT orders the available common resources. This prioritised list is used to guide the overflow direction when there are insufficient resources to support a session setup in the high priority network. The prioritised lists for different services are predefined and stored in the rule database as illustrated in Figure 1-6. The RAT selection algorithms in [30]-[31] are essentially also service-based. Performance evaluations are made by comparing the QoS gain of the presence of overflows to the separate systems. Zhang proposes to allow the overflow only of data services [30]. The argument is that data services accept elastic bandwidth allocation which will achieve much better resource utilisation. Andrisano et al. in [31] propose to allocate data services into the WLAN hotspot whenever inside the coverage, whereas the voice calls that cannot be served by UMTS are converted to VoIP and redirected to the WLAN. Again only initial RAT selection is studied in [31]. The capacity gain of integrated services is not clearly discussed in these papers.

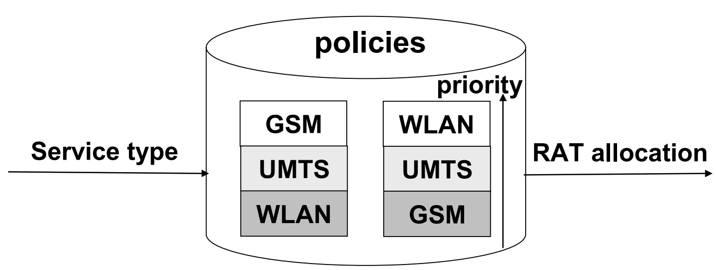

Figure 1-6: Illustration of service-based initial RAT selection

Cost Function and Fuzzy Logic

The MRRM strategies addressed in [26]-[31] are all about the initial RAT selection of call arrivals, i.e. they all deal with the RAT selection at the per-session level. A force-based model is proposed in [32] to evaluate additional IS-HOs, but merely for the circuit-switched traffic. The so-called forces model the cost of the cell load, the QoS, the time since the last handover of the mobile station and the handover overhead. The joint cost in turn decides the usage of a suitable RAT. As we have mentioned before in the methodology section, Zhu et al. in [16] propose to evaluate the priority of multi-radio access networks via a cost function to finally make the vertical handover decision. The cost function may aggregate the weighted costs of various parameters in order to determine a necessary vertical handover, e.g. bandwidth requirement, delay, battery power for all existing services etc. However, the numerical study therein merely evaluates the costs of static total bandwidth and static coverage in a simplified 3GPP/WLAN.

Similarly complex as the force model above, fuzzy logic algorithms involve the joint fuzzy logic of economic and technical parameters, and the network and the end user preference. Through the fuzzy logic, the possibility of the acceptance of a call request is represented by fuzzy variables such as yes, probably yes, probably no or no; weak or strong, bad, medium or good. Chan et al. in [33] use fuzzy ranking procedure to compare the performance of GPRS, UMTS and satellite system from the perspectives of signal strength, coverage, QoS perceived, and bit error rate. Giupponi et al. in [34]-[36] additionally use Reinforcement Learning to dynamically adjust the user non-satisfaction probability to a desired value. The quantified benefit will depend greatly on the network assumption and the fuzzy logic parameters. The scenario of integrated services is however not yet specifically discussed. The suitability of fuzzy logic for IS-HO is also not addressed.

1.4 Contribution

From the state of the art we know that there are some gaps in the current studies on MRRM:

- MRRM for integrated services considering the service requirements and the interplay of services
- Additional benefit from intersystem handover
- The impact of the heterogeneous characteristics of included multi-radio access networks
- MRRM implementation issue on input information and parameters gathering

This thesis contributes an MRRM algorithm for integrated RT and NRT services adapting to their heterogeneous requirements. Admitting application requirements oriented load mapping of integrated service, the algorithms decide whether to reallocate the input traffic into an appropriate radio access network. The triggering of intersystem handover adapts to service dependent requirements, and therefore satisfies the respective user preference and reduces the handover cost. The load information plays a key role in the decision algorithm. Special physical measurements are selected to map to the input load information. The final format mapped is aligned also to the information element carried by 3GPP signalling. All these considerations make the algorithm proposed practical for implementation in heterogeneous networks.

A mechanism of information transfer is also proposed to take place in the last epoch intersystem handover execution stage instead of in the current intersystem handover decision stage. The metric information is exchanged between the source and target access network. It is therefore suitable in distributed architecture for multi-radio resource management. This mechanism trades off the signalling overhead and the accuracy of the information provided to the multi-radio resource management algorithms.

An additional RAT pre-selection according to the system level resource efficiency is adopted to further improve the MRRM performance for NRT services in an EDGE/HSDPA/WiMAX heterogeneous scenario. In this thesis, WiMAX is, to our best knowledge, for the first time considered in the MRRM study.

The performance evaluation is carried out via a dynamic event-driven simulation tool. This tool Admits realistic local radio resource management which includes the local admission control of integrated services, realistic resource sharing policies between RT and NRT services, and intrasystem handover due to mobility and radio channel variation. The performance can therefore be interpreted close to the exact figures.

1.5 Outline

The remaining parts of this thesis are organised as follows:

Firstly in Chapter 2, the various mechanisms of radio resource management concerned are introduced, where the mechanisms are QoS requirements oriented. The QoS are focused on the air interface QoS. The MRRM strategies should be treated as a part of these mechanisms because the strategies lie on the radio resource management layer. Nevertheless they are integrated into the Chapter 4 where the MRRM performance via a simulator is investigated.

In addition to the local RRM, the radio channel properties are also specified for three types of networks, namely GSM/EDGE, UMTS/HSDPA and WiMAX. But the description of radio channel properties scatter in Chapter 3 and Chapter 6 when the corresponding system is first mentioned.

In Chapter 3, analytical models for pure CS services, which are built for GSM and UMTS sub systems respectively, are presented. Limited by the complexity and possibility, the analytical models allows us to have a statistical glance on the individual systems. The mobility influence on the system capacity and system performance can be initially investigated by the analytical model as well. Time varying system characteristics are impossible to be modelled dynamically. The numerical results are then compared to dynamic simulation results in order to validate the fundamentals of the dynamic simulator.

Chapter 4 presents the investigation of the MRRM performance based on a simulation model. By enabling more complicated dynamic properties, such as mixed mobility patterns, sectorised antennas with a specified antenna pattern, dynamic channel properties of every call, RRM and MRRM at call level and packet call level, the simulations compare the capacity of MRRM proposed with service-based policies and with the "Blind-data to UMTS" algorithm. The QoS of GSM and UMTS are investigated also. Then the performance of multiservices is evaluated.

A practical distributed MRRM is then proposed to minimise the possible signalling overhead for MRRM evaluation in Chapter 5. The performance of distributed MRRM is then compared to the separate systems and the ideal MRRM. The latter one supposes the necessary input system parameter is always available and therefore corresponds to the theoretical upper-bound performance.

The study presented in Chapter 6 extends the MRRM to a three-radio access network scenario which involves a hotspot WiMAX network overlapping partially the co-located GSM/UMTS network. In this 3RAT scenario, MRRM algorithms for NRT services are specifically implemented. A new MRRM algorithm enables a pre-selection of the most resource efficient RAT on top of the user datarate optimised MRRM. Performance evaluation is completed for separate systems, with the user datarate optimised MRRM, and with the newly proposed joint MRRM.

Finally, the conclusion and the outlook of future work are presented in Chapter 7.

Chapter 2

Radio Resource Management in a Single Network

The radio resource management (RRM) in a single network normally comprises admission control, power control (for UMTS only), packet scheduling, handover control and load control. Regarding the common sharing of resources of multi-radio access networks, load balancing and multi-radio access control are additional tasks for multi-radio resource management (MRRM). At first, we introduce the RRM in a single network which we will call local RRM in the scope of this thesis. The MRRM approaches will be given later in Chapter 4.

2.1 Admission Control

The admission control takes place at the call setup stage. Either a new arrival call or a hand over call needs a call setup in the cell to which the call wants to build up a wireless link. The admission control will decide whether a connection request will be accepted. Call blockings take place at the call setup stage, when, for a new arrival, sufficient quality cannot be guaranteed. Call droppings in contrast take place when sufficient quality cannot be guaranteed for ongoing calls.

The goal of call admission control is to maximise the capacity while ensuring the QoS provisions under the total resource constraints. Simply at the local network level, the call admission control is already a hot topic for scientific research. Dynamic resource allocation in typically non-uniform traffic distributions, optimum handover resource reservation and optimum resource partitioning between different services ([37]-[44]) all belong to the categories of admission control. Since the focus of this thesis is the multi-standard radio resource management, only the basic call admission control will be explained below. Furthermore the resource reservation and the value of the total resources are considered as fixed values.

2.1.1 Service Prioritisation

In principle, the call admission control depends on the quality of service (QoS) requirements of a specific call request. A Voice call needs dedicated resources so that the delay of the voice communication can be avoided or minimised. PS Calls, such as web browsing, only require resources during the download or upload transmission. The resources are in fact released during the reading time. Moreover, the resources assigned can be elastic. The transmission rate for a PS call is consequently elastic.

Taking into account the heterogonous requirement of the voice service and PS services, the admission control for integrated voice and PS services will firstly guarantee the resource requirement for voice calls, while the remaining resource will be shared by the PS calls. This is in fact the resource allocation strategy for integrated services in the cell level. It is foreseen that the remaining resource for the PS calls can become zero at the case that the active CS calls occupy exactly the complete resources. In this extreme case the active PS calls can temporary have data rates of zero. .

2.1.2 Resource Reservation for Handover Calls

Another admission control policy differentiates the treatment of new arrival calls and handover calls. Since users do not like their connectivity to be forced to terminate, some resource margin is usually reserved for higher priority handover calls. The resource reservation scheme usually mitigates the dropping rate at the expense of an enlarged blocking rate.

2.1.3 UMTS Specific Admission Control

A power interference aware admission control is especially applied to interference limited UMTS system and other systems employing CDMA technology. In an interference limited system, a newly admitted call will consume additional transmission power. For instance, in the down link, the increased total transmission power will cause the intra-cell interference to the other users in the same cell to increase. Meanwhile the inter-cell interference to the users in the surrounding cell is also increased. If the transmission power is allowed to excessively increase, the cell coverage and the quality of the existing connections will be reduced. So the admission control should ensure that the sum of the current load and the estimation of the load increment caused by the newly admitted call will not exceed the load threshold defined at the network planning.

The effect of a high cell load resulting in small cell coverage and vice versa is usually defined as cell breathing. The interplay of transmission power, external interference, nominal and real cell coverage, user bit rate requirement, and cell load are investigated in [45]-[46].

Depending on the definition of the load factor in the uplink and downlink, wideband power-based admission control strategy and throughput-based admission control strategy are usually deployed for the load estimation arising from a potential admittance as well as the admission decision [47]. Only if both the uplink and downlink admission control succeed, a bearer service can be established for the new admittance.

Due to the non-trivial access control encountered in a CDMA system, the joint admission control and power control strategies have been exploited in the past. A distributed power-control algorithm with active link protection is studied in [48]. The active link protection naturally supports the implementation of admission control. Two distributed connection admission control algorithms are proposed in [49], with which an infeasible call is rejected early in order to protect the feasible calls from dropping and the system converges to the Pareto optimal power assignment. Based on multiple time-scale interference prediction, rate adaptation and admission control are discussed in [50]. In [51], the SIR-based power control is involved for the power estimation for a new call and all existing calls. Besides that, adaptive different maximum allowable power limits are employed for new calls and handover calls. Some latest admission control schemes implement fuzzy logic to estimate the effective bandwidth requirement of a call request of a mobile and its mobility information [52].

2.2 Power Control

2.2.1 Power Control in UMTS

Power control is a very important and necessary RRM scheme for W-CDMA systems. For an interference limited W-CDMA system, on the one hand the quality of link level signal requires a least power to fulfil the target Eb/N0 (bit energy and noise density ratio), in order to maintain an acceptable bit error rate level. The higher the power, the better the signal quality is. On the other hand the allocated transmission power should not be much larger than the least power, otherwise the superfluous power introduce more interference to other wireless links. So the aim of power control is to adjust the power per transmitter of ongoing communication, so that the interference levels at receiver locations are minimised.

The power control mechanism is non-trivial since excessive reduction of the transmission power will in fact raise the interference level at the receiver partner. Aein has introduced the concept of carrier to interference (C/I) balancing in cochannel interference management in a satellite communication system [53]. The C/I balancing power control yields fairness in terms of user experienced C/I level. In a CDMA system, the adjacent channel (code) interference is however dominant. In the early development of a CDMA system, e.g. IS-95, power control algorithms normally aim to keep a constant received power level [54], [79]. Later on, also in IMT-2000 and UMTS systems, more sophisticated power control algorithms keep a common signal to interference ratio (SIR) level taking into account the adjacent channel interference as well as the cochannel interference due to multi-path fading [55]-[57], [80].

UMTS W-CDMA RAN employs two types of power control methods, namely inner loop power control and outer loop power control. The outer loop power control is carried out at the UE and RNC pair. It sets the target Eb/N0 for each radio bearer. The inner loop power control is also called fast power control. It is carried out at the UE and the node B pair. The fast power control occurs at every 10s/15. And the power control step is usually ±0.5dB or ±1dB [47]. A much finer power control step is not desired since it will cause likewise an infinite power update loop.

2.2.2 Power Control in WiMAX

The third RAN considered in this thesis is WiMAX. WiMAX is in fact a subset from the standard IEEE 802.16 for Broadband Wireless Access. In order to bring interoperability to the Broadband Wireless Access space, the WiMAX Forum is focused on establishing a unique subset of baseline features grouped in what is referred to as "System Profiles" that all compliant equipment must satisfy. These profiles and a suite of test protocols will establish a baseline interoperable protocol, allowing multiple vendors' equipment to interoperate; with the net result being system integrators and service providers will have option to purchase equipment from more than one supplier. Initially some profiles are published by the WiMAX forum, for example [58] where profiles for point to multi-point system with 256 OFDM in 3.5GHz band and 5.8GHz band are proposed.

It is well known that OFDM technology divides a broadband channel into N subcarriers and by transmitting a part of the data signal on every of the subcarriers it can significantly combat or even prevent ISI (inter symbol interference) when a non-linear filter channel imposed. In some situations, some of the subcarriers may become unreliable from time to time for some users. This frequency selective fading of channels can be exploited to realise multi-user diversity gain in terms of throughput by dynamic transmission power assignment, saying to allocate the subcarrier to only one user with the best channel gain [59]-[60]. After the subcarrier allocation, the available transmission power can be either equally distributed among subcarriers or be distributed following water-filling policy, i.e. assign more power for the subcarrier with high channel gain and assign less power for subcarrier with low channel gain under the constraint of maximum transmission power.

2.3 Packet Scheduling

Packet scheduling is a kind of RRM scheme which applies only to PS services because CS services require guaranteed bandwidth allocation and the requirement is therefore non-controllable. To efficiently and fairly share the resource among PS calls, packet scheduling plays the role.

Many packet scheduling schemes are focused on the maximisation of PS throughput. On an air interface where the adaptive modulation and coding schemes are involved, a user with a favourable link quality can achieve high data rate under resource constraints. So the packet scheduling pursuing throughput maximisation often results in a loss of fairness among the users. According to the user diversity, the users with a bad link quality may be very seldom scheduled or even be not scheduled at all.

Regarding the fairness, there are also several perspectives. The fairness may lie in resource or bandwidth. The resulting user data rates of multi-users can be different if the resource is not proportional to the bandwidth due to varying channel conditions for multi-users. An example of this non-proportionality is in UMTS. The user data rate is not proportional to the transmission power due to the non-linear effect arising from the pathloss and the interference. To the end of bandwidth fairness, the resource allocations of users are quite different due to the divers channel conditions of various users. This type of scheduling is proved resulting in lower capacity than fair resource scheduling [61]. From the capacity maximisation point of view the fair bandwidth scheduling is the worst scheduling.

In this thesis the fair resource scheduling is employed for the local packet scheduling. Integrated to different radio access networks, the packet scheduling in practice is quite disparate from RAN to RAN.

2.3.1 PS Scheduling Algorithm in EDGE

For GSM, the EDGE technology is considered applied to PS services. One user may be allocated to up to 8 time slots (i.e. on a single carrier frequency) and the data from up to 6 users may be interleaved inside a time slot.

Taking the SNR (Signal-to-Noise-Ratio) of each user helps to derive the data rate for a timeslot as a function of its respective SNR. The number of time slots a user required can be derived as $\frac{max\,user\ \ supported\ \ datrate}{max\ datarate\ per\ time\ slot}$, which may not exceed 8 in GSM.

Taking the fractional portion of the time slot j occupied by a user i as $frac_{i,j}$, the total number of fractions of N users in time slot j should give 1. The total sum of the time slots available for user data shouldn't exceed the limit of time slots for a cell defined in the network planning stage, denoted by M.

Taking the data rate for each time slot for a number of N total users and an number of M timeslots in a cell, the data rate of user i is

$$R_i = f(SNR_i)\sum_{j=1}^{M} frac_{i,j} \tag{2.1}$$

Hence maximising the data rate with respect to its variables will yield an optimisation expression Equation (2.2), where the third constraint ensures one user is allocated on a single carrier.

$$max\left\{\sum_{i=1}^{N} R_i\right\} = max\left\{\sum_{i=1}^{N}\left(f(SIR_i)\cdot\sum_{j=1}^{M} frac_{i,j}\right)\right\} \tag{2.2}$$

Subject to

$$i)\sum_{i=1}^{N} frac_{i,j} = 1, \forall j$$

$$ii)\sum_{i=1}^{N}\sum_{j=1}^{M} frac_{i,j} \le M$$

$$iii) if\ \ for\ \ user\ i,\ \ frac_{i,j} \neq 0, \forall j \in [a,b], \Rightarrow \left\lfloor \frac{b}{8} \right\rfloor = \left\lfloor \frac{a}{8} \right\rfloor$$

$$iv)\sum_{j=1}^{M} frac_{i,j} \le \frac{max.rate}{f(SNR_i)}, \forall i$$

Suppose the interference and noise to an EDGE user does not differentiate among the carrier frequencies inside a single cell. The optimisation problem can be simplified to a procedure searching for the maximum sum of fractions of continuous time slots allocating on a single carrier. The simplified PS scheduling algorithm is illustrated in Figure 2-1.

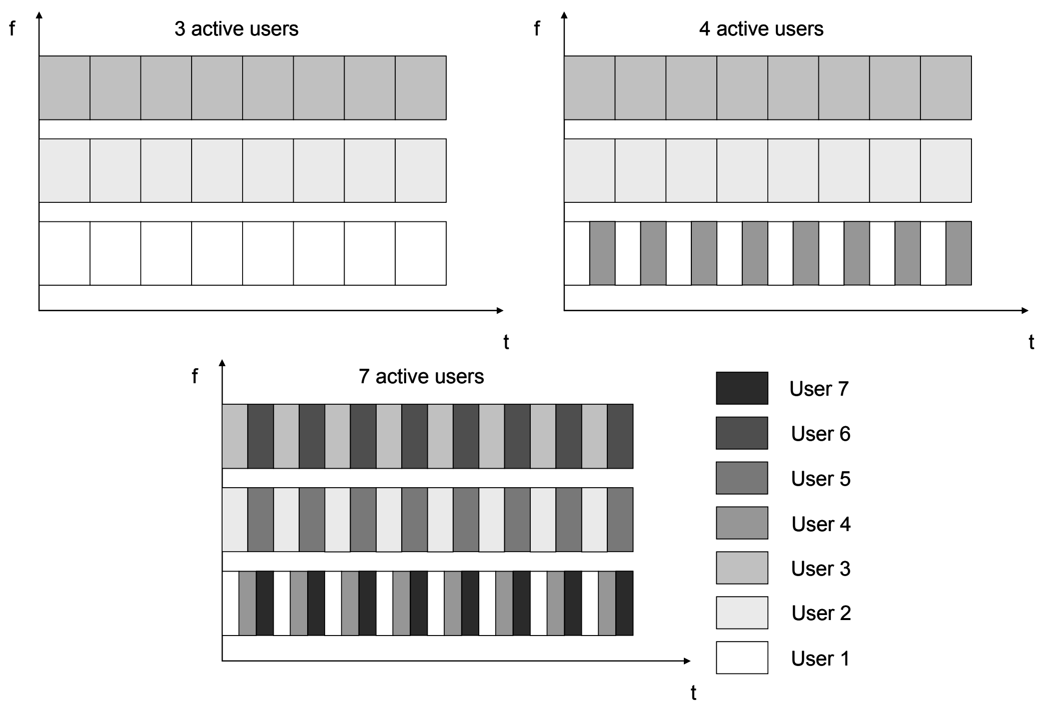

Figure 2-1: Illustration of the PS scheduling algorithm in EDGE

2.3.2 PS Scheduling Algorithm in HSDPA

For UMTS, HSDPA [62]-[63] is applied to the PS services. The scheduler is a key element of HSDPA that determines the overall behaviour of the system and, to a certain extent, its performance. For each transmission time interval (TTI), it determines which terminal (or terminals) the HS-DSCH should be transmitted to and, in conjunction with the adaptive modulation and coding schemes, at which data rate. One important change from R'99 channels is that the scheduler is located at the Node B as opposed to the RNC. In conjunction with the short TTI (2 ms) and the channel quality indicator feedback, this enables the scheduler to quickly track the UE channel condition and adapt the data rate allocation accordingly, which is normally called fast scheduling characteristics for HSDPA. Several algorithms can be used for the scheduler. Some of them are presented below: Round Robin (RR), Maximum Carrier to Interference (C/I) and Proportional Fair (PF).

RR schedules users according to a first-in first-out approach. It provides a high degree of fairness between the users, but at the expense of the overall system throughput (and therefore spectral efficiency), since some users can be served even when they are experiencing destructive fading (weak signal).

The maximum C/I scheme schedules users with the highest C/I during the current TTI. This naturally leads to the highest system throughput since the served users are the ones with the best channel. However, this scheme makes no effort to maintain any kind of fairness among users. In fact, users at the cell edge will be largely penalised by experiencing excessive service delays and significant outage.

The PF scheduler ([64]-[65]) aims to strike a compromise between the fairness of the RR scheme and the efficiency of the C/I -based scheduler by serving that flow at TTI t which maximises the ratio $R_m(t)/\tilde{R}_m(t)$, where $R_m(t)$ denotes the instantaneous gross data rate of flow m and $\tilde{R}_m(t)$ denotes its exponentially smoothed experienced gross data rate:

$$\tilde{R}_m(t) = (1-\alpha)\tilde{R}_m(t-1) + \alpha \cdot \delta \cdot R_m(t-1) \tag{2.3}$$

$\delta = 1(0)$	indicator if data flow m was (not) served in TTI $t-1$
$\alpha \in [0,1]$	the associated smoothing parameter
$\tilde{R}_m(t_0) = \alpha$	the assumed initial value at the flow's generation time t_0

It is readily verified that for $\alpha = 0$ $(\alpha = 1)$ the PF scheduler is identical to the C/I -based (RR) scheduler (loosely using the convention that $1/0 = \infty$). The usually used parameter assumes $\alpha = 0.001$.

It is important to mention that the implementation of QoS (i.e. different subscription classes) creates new constraints on the scheduler. Other parameters such as user priority level may override the above scheduling algorithms. The fairness between the users will then be dominated by the QoS requirements. The packet scheduling involved in the later investigation in this thesis is following round robin, i.e. the user data are transmitted on TTI level.

2.3.3 PS Scheduling Algorithm in WiMAX

Scheduling as an important feature is implemented for WiMAX. The scheduler is set in MAC layer in the BS of a WiMAX network, for both downlink and uplink. The uplink scheduler works on the subscriber station (SS) requests and the differentiation of services requesting resource allocation, while the downlink scheduler allocates the resource for various usages based the knowledge of its own queue status. The 802.16a MAC relies on a Grant/Request protocol for access to the medium and it supports differentiated service levels, e.g., dedicated T1/E1 for business and best effort for residential. The protocol employs TDM data streams on the downlink and TDMA on the uplink, with the hooks for a centralised scheduler to support delay-sensitive services like voice and video. The downlink TDM bursts follow a channel robustness descending order. 802.16e also provides sub-channelisation techniques to more efficiently manage the channel bandwidth among multiple end users. The base station uses sub-channelisation to optimise scheduling of multiple users having distinct spatial signatures. The various sub-channelisation schemes offered by 802.16e allow more efficient scheduling of users based on channel quality, priority, power, and bandwidth allocation.

To understand the functionality of the WiMAX service uplink scheduler, it is necessary to have the knowledge about the service classification in the standard and the corresponding QoS associated with the respective service class. There are four types of services defined in 802.16-2004, unsolicited grant service (e.g. applications like digital voice transmitted by T1/E1 digital line, and VoIP), real-time polling service (e.g. application like MPEG video stream), non-real-time polling service (e.g. application like FTP) and best effort service (e.g. application like web-browsing) [66]. The unsolicited grant service requires a maximum sustained traffic rate, a maximum latency; the variation of packet delay should not exceed the tolerated jitter. This type of service needs also request and transmission policy for resource allocation. Nevertheless, resources will be definitely assigned for an unsolicited grant service after the request sent. The distinguished requirement of the two polling services is that the real-time service requires a maximum latency whereas the non-real-time service is labelled with a traffic priority level. Both need a minimum reserved traffic rate and are polled (scheduled) by the BS on a periodic basis. Once there are too many connection requests, the BS is not able to poll the SS individually. Under this situation, the contention-based scheduling works for the uplink request and grant process.

Based on the various requirements and priorities of services and the associated request/grant mechanisms, the scheduler at BS allocates resources for uplink connections. Besides the scheduling on time, the scheduling on frequency and space should be implemented if the subchannelisation, adaptive antenna system (AAS), and space-time coding are deployed at BS and SS pair. The WiMAX scheduler is therefore potentially scheduling the air interface resources in three dimensions: time, frequency, and space.

Using a utility-based scheduling framework, optimal and suboptimal algorithms are proposed for power allocation on every subcarrier and subchannel allocation at the step of every time slot [67]. Furthermore, the TDM is allowed for multiple users sharing a subcarrier. The authors have also compared the simulation results for three subchannelisation schemes: adjacent channelisation, random channelisation, and interleaved channelisation. Using adjacent channelisation, the frequency diversity can be better exploited. Using interleaved channelisation, which is primarily used in 802.16d/e, however, reduces the variance in the channel gains across subchannels for each user.

The potential small variance of subchannels using interleaved channelisation is on the other hand also an advantage. Exploiting this feature, a user may feedback a representatively single measurement report, which further reduces the signalling overhead. Suppose the frequency diversity is negligible, the BS can schedule the downlink traffic in a fair subchannel sharing scheme. Moreover, the power allocation on each subchannel can be also fixed throughout the packet lifetime, e.g. sharing the total BS transmission power among users equally. These two resource allocation strategies make the trade-off between the implementation complexities of scheduling and maximising user throughput.

2.4 Intrasystem Handover Control

Handover is a procedure for a mobile station switching the connection from one base station to another. The base stations involved in an intrasystem handover belong to a unique radio access technology. Intrasystem handover is called also traditional handover in the contrast to intersystem handover or vertical handover between different access networks. A handover is hard when there is only one base station associated at a time. A mobile station in soft handover maintains the connections to several base stations simultaneously. An interfrequency handover is often referred to when a UMTS mobile station switches the carrier frequencies supported by the hosting cell.

Handover is an essential solution in wireless communication networks to ensure continuous connectivity for moving users who are in connected mode. In an UMTS network, handover is a family of procedures that adds or removes one or several radio links between one UE and UTRAN when an RRC connection exists and the position of the UE is known on cell level in the UTRAN. This family consists of hard handover, interfrequency handover and soft handover. In contrast, there is only hard handover valid in a GSM network. Supporting nomadic subscribers in WiMAX is however a relative new issue, Intel® Technology is the basic contributor to extend WiMAX to mobility [68]-[69].

A handover decision is made by the radio resource management component in the network based on the measurement report from the UE. An intrasystem handover decision is mainly based on the metric of signal quality. To this end,

- the UE must be able to measure the signal quality in the candidates of potential hosting cells.
- the hosting cell and the handover cells should have coverage overlap

Threshold, hysteresis, and time to trigger (ΔT) can be freely combined to build the hard handover criterion. The three parameters are illustrated in Figure 2-2. A usually implemented hard handover algorithm is involving a hysteresis for the detected situation that a handover candidate cell has a better signal quality than the hosting cell for a period of ΔT. A well quantified hysteresis and time to trigger avoid efficiently some fault decisions due to potential imperfect channel measurements [70].

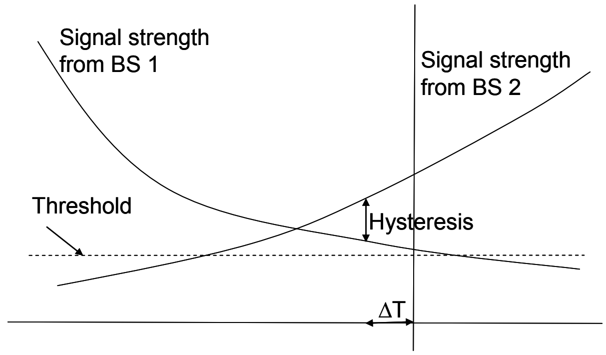

Figure 2-2: Illustration of threshold, hysteresis, and time to trigger in hard handover

Since the hard handover procedure is similar in every network, independent to the type of the wireless networks, here the hard handover in UMTS is taken as the example. Thereafter, an example algorithm for soft handover in UMTS is given.

2.4.1 Hard Handover in UMTS

Figure 2-3 illustrates a hard handover. The NW RRC determines the need for hard handover based on received measurement reports or load control algorithms.

For inter-frequency handover the measurements are assumed to be performed in slotted mode. That is that the compressed mode must be taken for the data transmission. The measurements on another frequency will take place during the gap of the data transmission.

The NW RRC first configures the NW L1 to activate the new radio links. The NW L1 begins transmission and reception on the new links immediately. The NW RRC then sends the UE RRC a PHYSICAL CHANNEL RECONFIGURATION message (several other messages e.g. RADIO BEARER RECONFIGURATION and TRANSPORT CHANNEL RECONFIGURATION can also be used to perform hard handover). The message indicates the radio resources that should be used for the new radio link. The UE RRC configures the UE L1 to terminate reception on the old radio link and begin reception on the new radio link.

After the UE L1 has achieved downlink synchronisation on the new frequency, a L2 link is established and the UE RRC sends a PHYSICAL CHANNEL RECONFIGURATION COMPLETE message to the NW RRC. After having received the L3 acknowledgement, the NW RRC configures the NW L1 to terminate reception and transmission on the old radio link.

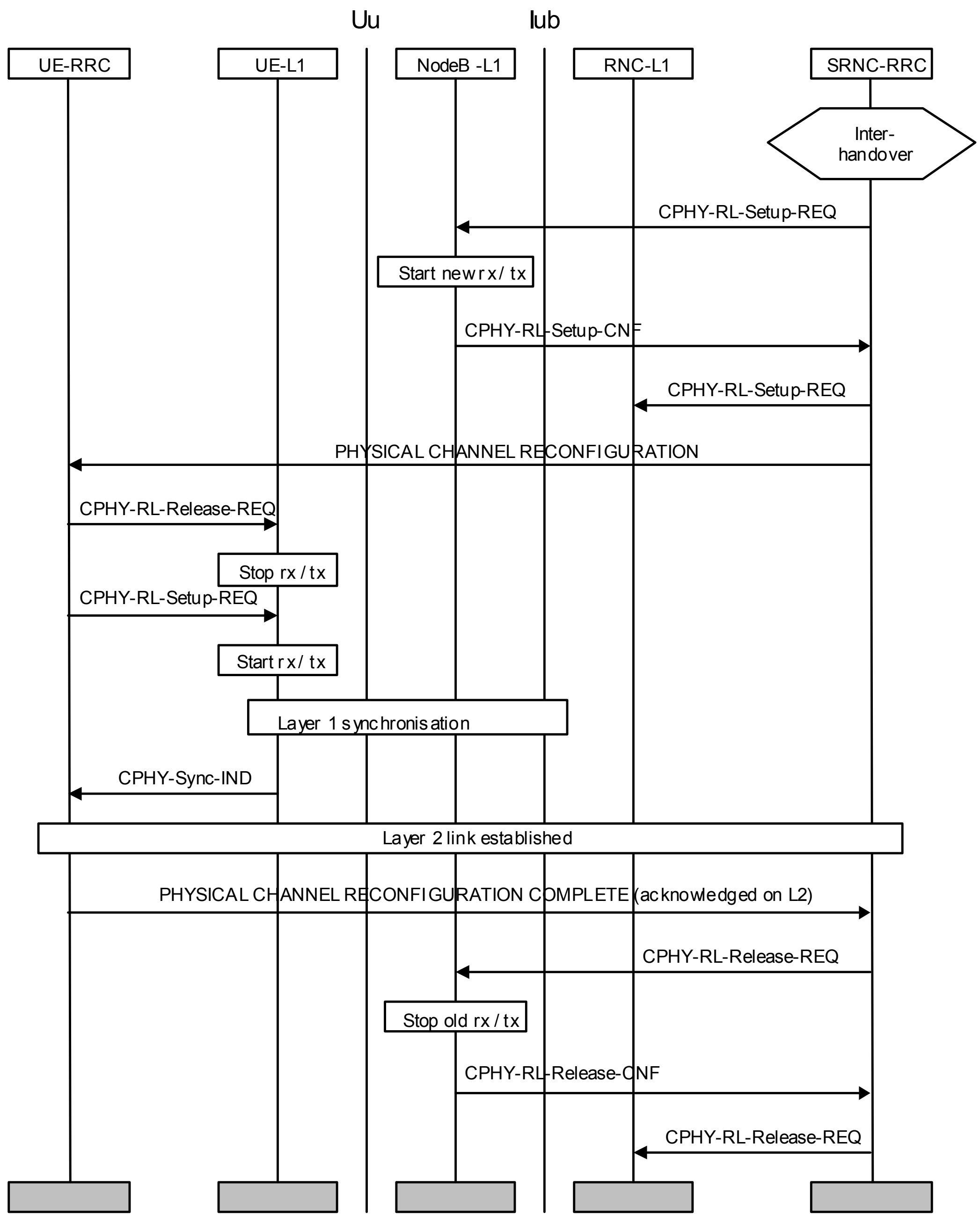

Figure 2-3: Hard handover [71]

2.4.2 Soft Handover

A describing example of a Soft Handover Algorithm [72] is represented in this section, which exploits reporting events 1A, 1B, and 1C described in [73]. It also exploits the Hysteresis mechanism and the time to trigger mechanism described in [73].

Other algorithms can be envisaged that use other reporting events described in [73]; also load control strategies can be considered for the active set update, since the soft handover algorithm is performed in the RNC.

For the description of the Soft Handover algorithm presented in this section the following parameters are needed:

- As_Th: Threshold for macro diversity (reporting range);
- As_Th_Hyst: Hysteresis for the above threshold;
- As_Rep_Hyst: Replacement Hysteresis;
- ΔT: Time to Trigger;
- As_Max_Size: Maximum size of Active Set.

Figure 2-4 describes this Soft Handover Algorithm.

As described in Figure 2-4:

- If Meas_Sign is below (Best_Ss - As_Th - As_Th_Hyst) for a period of ΔT remove Worst cell in the Active Set.
- If Meas_Sign is greater than (Best_Ss - As_Th + As_Th_Hyst) for a period of ΔT and the Active Set is not full add Best cell outside the Active Set in the Active Set.
- If Active Set is full and Best_Cand_Ss is greater than (Worst_Old_Ss + As_Rep_Hyst) for a period of ΔT add Best cell outside Active Set and Remove Worst cell in the Active Set.

Where:

- Best_Ss: the best measured cell present in the Active Set;
- Worst_Old_Ss: the worst measured cell present in the Active Set;
- Best_Cand_Set: the best measured cell present in the monitored set;
- Meas_Sign: the measured and filtered quantity.

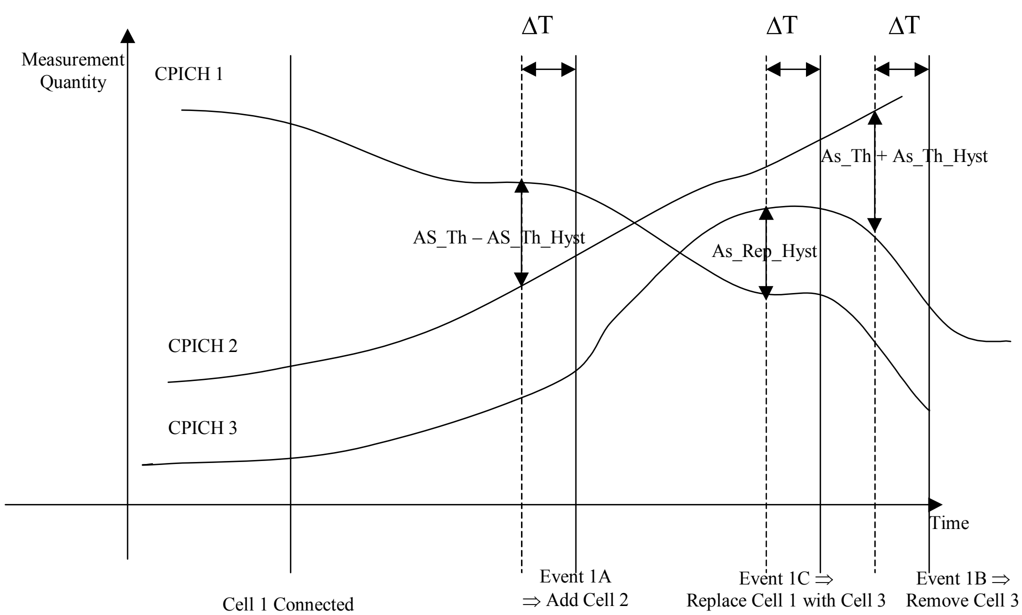

Figure 2-4: Example of Soft Handover Algorithm [72]

2.5 Load Control

An important task of the RRM functionality is to ensure that the system is not overloaded and remain stable. If the system is properly planned, and the admission control and packets scheduler work sufficiently well, overload situations should be exceptional. If overload is encountered, however, the load control functionality returns the system quickly and controllably back to the targeted load, which is defined by the radio network planning.

The often deployed load control mechanisms include: drop low priority calls, degrade bandwidth of on going calls, reject power update requests, handover to other carriers, deploy hierarchical cellular structure, and integrate multi-radio access networks etc.

A simple load control algorithm is to select the most resource consuming call to make a handover request so that the resources empted can be used for the setup/handover request which has triggered the load control. At the call setup stage, a new arrival may obtain adequate resource after the handover succeeds. At the handover stage, a dropped call may save more resources than the direct handover selection.

A relative advanced load control or load sharing algorithm is proposed in [74] for a CDMA system. The basic ideal of this algorithm is when the Eb/N0 decreases below a threshold, γ, the based station starts directing some subscribers to less loaded surrounding cells by lowering the pilot signal power, thus shrinking the cell size and increasing the total number of subscribers that the network system can handle. This algorithm is especially suitable for a network with uneven traffic distribution.

Hierarchical cellular structures are proposed to deal with mixed mobility of fast moving and slow moving users, and the trade-off of capacity and coverage [75]-[76]. The slow moving users prefer to attend to the micro cell layer, lower layer, while the fast moving users prefer to the macro cell layer, umbrella or upper layer. According to a classic fixed resource allocation, the slow moving users may overflow to the upper layer if there are not enough idle resources. The fast moving users prefer in contrast handover in the same layer when meeting inadequate resources. A handdown procedure of slow moving user gives the modified feature to the classic fixed resource allocation that confines the slow moving users not to move back to the lower layer. If the upper layer is full and there is at least one slow moving user in the upper layer, a slow moving user may make handdown back to a preferred micro cell to release resources in the upper layer. Several handdown may be required to release sufficient resources for a fast moving user. When a slow moving user terminates his connection in the lower layer, as many as possible slow moving users in the upper layer will make handdown to utilize the resources released. The hierarchical cellular structure avoids frequent handover for fast moving users and is particular useful for an urban environment with high traffic intensity.

In addition to the hierarchical cellular structure, dynamic resource allocation can further reduce the handover rate [77]. According to the dynamic resource allocation, a user is allocated to the preferred cell with the best link quality in the preferred layer if adequate idle resources are available. Otherwise it firstly tries to borrow a channel in the same hierarchical layer, if failed it tries to apply handdown procedure, if failed it ultimately borrows a channel in the upper hierarchical layer.

Chapter 3

Analytical Model and Event-Driven Simulator for Voice Service in GSM/UMTS

3.1 The State of the Art – Analytical Models

The admission control of integrated CS and PS services in a homogeneous network has been extensively studied.

A 3-D Markov model for integrated voice and data services in mobile cellular networks has been studied in [78]. The voice service is assumed on/off characteristics and has higher priority than the data service. The best effort data services share fairly the leftover bandwidth from voice services.

The shadow fading is usually modeled by log-normal distributed variable. The multiplication or division of two log-normal distributions gives also a log-normal distribution. Based on this assumption, Gilhousen et al. have firstly studied the reverse link (uplink) capacity of a cellular CDMA system using an analytical model [79]. They have also studied the capacity of a cellular CDMA system for forward links (downlinks) in the same paper, however based on a simulation model since they announce that the summation of log-normal distributions cannot be solved analytically.

At that early time, the reverse link power control in [79] is aiming to maintain the same received power level at base station. The CDMA system capacity on reverse links based on the power control maintaining the same SIR level for reverse links is extended in [80]. Unlike in [79]-[80] where only RT services are studied, the uplink performance of integrated services in CDMA system is further studied by an analytical model in [81]. The SIR based power control is applied. The contribution of this paper is that the PS service here is considered elastic traffic which bandwidth will be reduced if the total received interference at BS is beyond a maximum level by so-called congestion control. Although this thesis focuses on the downlink capacity and the performance of multiservices in multi-radio access networks, the analytical methodologies studied for the uplink capacity are still good reference.

Recent studies on the sum of log-normal distributions has approved that the log-normal distribution can be approximated to a log-normal distribution as well [82]-[87]. This offers the opportunity of using the analytical model to study the downlink capacity of a CDMA system, where the summation of downlink interference and the summation of link transmission power will be needed.

The authors of [82] have studied the downlink capacity of a CDMA cellular system via an analytical model. The propagation loss is modeled by the signal attenuation arising from the distance, the shadow fading and the multipath fading. The difference of the power requirement of data services from voice service comes from the different data rate requirement, i.e. data services require high data rate and therefore high TX power. The data services are treated in this paper requiring fixed data rate as voice service. The outage probability is finally the joint probability of a Poisson number of users and total TX power with the respective user number exceeding the maximum TX power.

To summarize, an analytical or simulation model of a system study usually contains three parts:

- The service model or models if multiple services are involved
- The specified service requirements
- The system model for resource assignment policies optionally including the approaches for admission control

The following are various models investigating integrated services in cellular communication networks.

Service/Traffic Model

- On-off model for voice service and a Poisson distribution for the number of data users [88]. User mobility is not considered in this paper.
- Poisson arrivals for both voice and data services [89]. Besides that, the call dwell time (or residual time) will be reduced by an empirical weighting factor due to users' manual termination detecting worse quality of signals. Since the mobility influence on the call dwell time is neglected, the call dwell time here is identical to the service time in the conventional Markov model.
- The voice service is modeled by on-off states and the data service considered as best effort service and is modeled by a Poisson arrival [90] with an exponential distributed data size. Since the best effort service is assigned only the remaining bandwidth from the voice service, the call duration of a data call is relevant to the current voices users, its own data size, the concurrent other data users and the total capacity of a cell. Besides the Poisson arrival and the exponential service time, the mobility influence is represented by additional Poisson handover arrivals. The final call service time by a server (a BS in a cellular network), is then the minimum of the call duration and the cell residual time before moving to another cell. This model has neglected the on-off characteristics of data sessions.

Service Requirement

- The voice users have priority over data users so that the simultaneous supportable number of data users depends on the units of remaining resource after fulfilling voice users' requirement [88]. A data user is assumed to require the same bandwidth as a voice user whereas the requirement of the bit error rate of a data user is more stringent than a voice user. The multiple-access capability (number of users) of voice users in such a CDMA cell is therefore larger than data users.

- Data services require also circuit-switched connection and need *b* times bandwidth of voice users [89]. They treatment of data services using circuit-switched connections is not suitable.
- The voice service requires a fixed amount of bandwidth for a call while the best effort service is assigned only the remaining bandwidth from the voice service [90] .

System Model

- The system behaviour for voice service and data service is modeled by two separate models in [88]. An optimal admission control for voice service proposed aims to minimise the average blocking probability of voice users per unit time. To this end, some calls are forced to be blocked in order to minimise the long-run blocking. This access policy is approved to achieve 10% to 15% less blocking probability than the direct admission policy which rejects the new arrivals. This type of admission control has effectively load control functionality. Moreover a graceful degradation model for data user reception lets the data users have a non-zero successful probability even when the number of data users is greater than the multiple-access capability of data users. This approach increases the data throughput as well as decreases the average data delay compared to the threshold model which gives hard rejection to new arrivals in case of insufficient channels. The trade off of the benefits of the graceful degradation model is the increased packet loss rate which can be compensated by retransmissions.
- The blocking and dropping probabilities are calculated based on a two dimension Markov model [89]. The authors first calculate the downlink capacity based on two levels of bit error rate. The two numbers are then set as the two boundaries of maximum number of new arrival calls and handover calls, i.e. there are guard channels for handover calls in the preposition that the quality of service may be degraded.
- A 3-D Markov Model with state index *i, j, k* where *i* represents for the total number of voice users, j represents the active voice users among i and k represents the number of data users [90] . A sophistical constraint of capacity for voice service Γ is also manually selected. The total resource sharing is then, given a capacity *C* and the number of voice users *j*, the data rate per data user (in total *k* users) is $min(\frac{C-j\gamma_v}{k}, \frac{C-\Gamma}{k})$. Their proposal for call admission control is based on six parameters $\{M_v, T_v, \beta_v, M_d, T_d, \beta_d\}$ where *M-T* is the guard bandwidth for handover calls. A new voice arrival will be accepted if the number of voice users is less than T_v, be accepted

with probability β_v if the number is exact T_v, and be rejected, otherwise. Same policy is for data users with the three parameters with subscription d. By ignoring the on-off transition of the voice service, the 3-D Markov model can be reduced to an approximate 2-D model which achieves also high accuracy. It is a very suitable one for a GSM system where the signal noise ratio related to the signal quality is less severe than the time slot resources. For interference limited systems, such as WCDMA, some other constrains need to be considered for a call blocking, e.g. the total BS transmission power constraint and the link maximum transmission power constraint. Moreover, the capacity of a WCDMA system is not a fixed value but a variable depend on the traffic and the interference conditions.

3.2 Markov Model for Voice Service in GSM

3.2.1 Markov Model

New calls arrive at a rate of λ, whereas handover calls arrive at a rate of λ_h. Suppose both call arrival process are Poisson process and they are independent to each other. Thus, the total calls arrive at a rate of $(\lambda + \lambda_h)$ and the arrival process is still a Poisson process.

Suppose the departure process is also a Poisson process with a rate μ. The corresponding service time for a call is then exponential distributed with a mean $1/\mu$. If a user is moving across a cell boundary, the residual time in a cell will be different from the service time. In the following, service time will be treated synonymously with call duration for a mobile user. A call moving out of the coverage of this cell (in practice “moving outside” is determined by the SNR below a threshold) will be handed over to another cell. The time slot occupied by an on going voice call will be released upon events either a termination of the call or a handover which the call has to make. Thus, the minimum of the call duration and the cell residual time will yield a free time slot. Suppose the residual time is also exponential distributed with a mean $1/\mu_h$. In addition, the residual time is independent on the call duration. The departure time is then the minimum of two exponential distributions with a rate $(\mu + \mu_h)$. If there are i users, the departure rate is $i\,(\mu + \mu_h)$.

We can use M/M/N Markov process to describe the voice call arrival and departure in a GSM cell. A voice call in GSM requires one time slot if the signal noise ratio (SNR) is greater than a threshold which is corresponding to the target bit error rate for the conversational communication. Suppose a GSM cell has N time slots, among which one time slot is reserved for handover traffic, new arrivals will be assigned time slots up to $(N\text{-}1)$. The state transition diagram of this model is sketched in Figure 3-1 where the states represent the number of time slots.

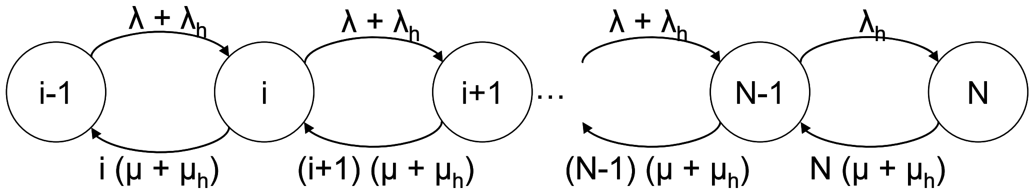

Figure 3-1: State transition diagram of the Markov process for voice calls in GSM with 1 time slot reserved for handover calls

The steady state probability of state *i* corresponds the probability of i active voice calls in a GSM cell. The steady state probability of state *i* is:

$$p(i) = \frac{\rho^i}{i!} p(0) \qquad i \le N-1$$
$$p(N) = \frac{\rho^{N-1}\lambda_h}{N!(\mu+\mu_h)} p(0) \tag{3.1}$$

$$p(0) = \frac{1}{\sum_{i=0}^{N-1}\frac{\rho^i}{i!} + \frac{\rho^{N-1}\lambda_h}{N!(\mu+\mu_h)}}$$
$$\rho = \frac{\lambda+\lambda_h}{\mu+\mu_h}$$

3.2.2 Determination of Handover Arrival Rate and Departure Rate

Assuming the overall system is homogeneous in statistical equilibrium, the mean handover arrival rate in a cell should be equal to the mean handover departure rate toward neighbouring cells. Assume one cell has *K* neighbouring cells, the probability that a handover call arrives in a neighbouring cell is *1/K*. That is:

$$\lambda_h = K \cdot \mu_h E[I] \cdot 1/K = \mu_h \sum_{i=0}^{N} i \cdot p(i) \tag{3.2}$$

As seen in Equation (3.1) and (3.2), the handover arrival rate and departure rate depend on the steady state probabilities, while the steady state probabilities are derived using the handover and departure rates. We use the iterative method introduced in [91] to determine the handover arrival rate. The iterative method is re-given as follows:

1. Set an initial value for λ_h. Let *X* and *Y* be the random variable for the cell residual time and the service time (call duration) of a call, respectively. When P_h is the probability that the time slot is released by a handover departure,

$$P_h = Pr(X < Y)$$
$$= \frac{\mu_h}{\mu_h+\mu} \tag{3.3}$$

Assuming the blocking probability and the handover failure rate is far smaller than 1, the initial value for λ_h is set to

$$\begin{aligned}\lambda_h &\approx P_h \cdot (\lambda + \lambda_h) \\ &= \frac{P_h}{1-P_h} \cdot \lambda = \frac{\mu_h}{\mu} \cdot \lambda\end{aligned} \tag{3.4}$$

2. Compute the steady state probability following Equation (3.1).
3. Compute the mean rate of handover arrivals $\lambda_{h,new}$ using Equation (3.2).
4. Let ε (>0) be a predefined small value. If $\left|1 - \frac{\lambda_{h,new}}{\lambda_h}\right| < \varepsilon$, let $\lambda_h \leftarrow \lambda_{h,new}$ and stop. Otherwise, let $\lambda_h \leftarrow \lambda_{h,new}$ and go to step 2.

Now the remaining unknown parameter need to be determined is the handover rate μ_h. We derive this based on a simple mobility model. Assume that users are uniformly located inside a cell and move with a uniform velocity keeping the initial moving direction unchanged throughout their call durations. Assuming that the cell radius is R and an MS's velocity is V, a mobile in its source cell has to travel a distance that is uniformly distributed between 0 and $2R$ before its first handover. The mean cell residual time is then:

$$\frac{1}{\mu_h} = \int_0^{2R} \frac{l}{V} f_L(l)dl, \text{ where } f_L(l) = \begin{cases} 1/2R, \text{if } 0 \le l \le 2R \\ 0, \text{otherwise} \end{cases} \tag{3.5}$$

Thus, $\frac{1}{\mu_h} = \frac{R}{V}$. It is around 960 sec for R equal to 800m and V equal to 3km/h (a pedestrian speed), and 24sec for V equal to 120km/h (a vehicular speed).

3.2.3 Blocking and Dropping in a GSM Cell

The blocking probability is the sum of the probabilities that the current number of active users is not less than N-1. The dropping probability is equal to the steady state probability of state N.

$$\begin{aligned}B &= p(N-1) + p(N) \\ D &= p(N)\end{aligned} \tag{3.6}$$

Knowing the parameters λ, λ_h, μ, μ_h and N, the blocking and dropping probability can be obtained by substituting Equation (3.1) Equation (3.6)

If there is no time slot reservation for handover calls, the state transition diagram converts then to Figure 3-2. Since the handover calls are treated same as the new arrival calls, the dropping probability is equal to the blocking probability and is given by $p(N)$.

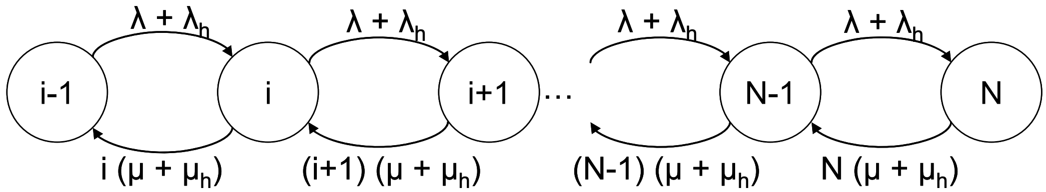

Figure 3-2: State transition diagram of the Markov process for voice calls in GSM without resource reservation

Note that if the mobile stations are stationary, there is no need to reserve some time slots for handover calls. Hence, the dropping probability for pure stationary users makes no sense. The total available resources in terms of time slot can be allocated to the traffic initiated by users who want to make a phone call. The blocking probability is thus equal to the probability that all *N* time slots are occupied. The calculation of this blocking probability is well known as the Erlang B formula that is originally used for the traffic engineering of wired networks.

Assume that there is only voice service. The mean call duration is 120 seconds. Pedestrian users are moving at 3km/h according to the simple mobility model described in 3.2.2 inside a cell with a radius of 800m. The maximum time slots available for data transfer in a GSM cell is 21. In the case that there are merely pedestrian users, either one time slot is reserved for handover calls or with no resource reservation. The resource reservation is naturally cancelled in the case that there are only stationary users. The blocking probabilities of pure pedestrian users and pure stationary users are shown in Figure 3-3. The dropping probability of pure pedestrian users is shown in Figure 3-4.

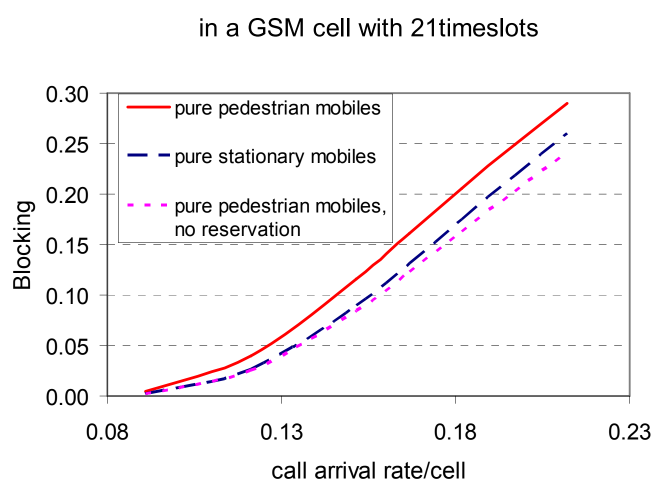

Figure 3-3: Mobility influence and resource reservation effect on voice service blocking through simple Markov model

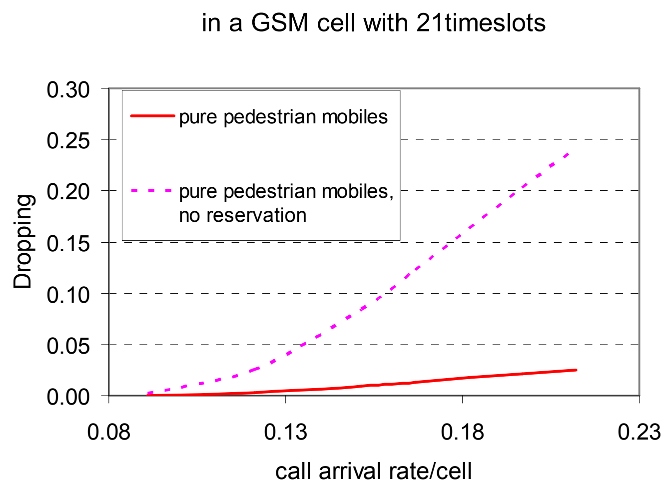

Figure 3-4: Resource reservation effect on voice service dropping through simple Markov model

The significance of resource reservation can be observed from comparing the blocking and dropping of pure pedestrian mobiles with and without resource reservation. In the case of without resource reservation, since no special treatment for handover calls is taken, the dropping probability is equal to the blocking probability of the new arrivals. The resulted high dropping probability (Figure 3-4) for on going calls is unacceptable in practice although the blocking probability is slightly reduced by more resources available (Figure 3-3).

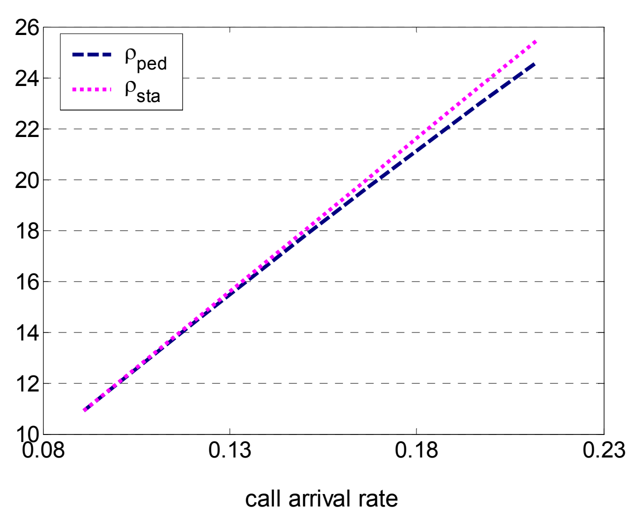

Figure 3-5: The traffic intensity in a GSM cell with pure stationary users or with pure pedestrian users assuming the identical resource in use and without resource reservation for handover calls

Regarding the mobility influence, we have to jointly consider the resource reservation scheme and the effective traffic intensity. The handover traffic created by moving users reduces effectively the traffic intensity compared to the case of with pure stationary users $\rho_{ped} = \dfrac{\lambda + \lambda_h}{\mu + \mu_h} < \rho_{sta} = \dfrac{\lambda}{\mu}$, see Figure 3-5. Therefore, the blocking of pedestrian users is even slightly lower than the blocking of stationary users when there is no resource reservation for handover calls (Figure 3-3). With one time slot less resources which is reserved for handover calls, the blocking is clearly worse than with full resources serving for call requests. Nevertheless, the resource reservation is quite necessary for mitigating the unacceptable dropping of moving users, as mentioned above. The reserved one time slot reduces significantly the dropping at the expense of blocking degradation. But the sacrifice of blocking is relative small compared with the dropping rescued. This simple, fixed resource reservation to avoid call dropping will also be adopted later in the simulation study.

3.2.4 Suitability of the Markov Model

Note that this Markov model is suitable for both uplink and downlink in GSM. By converting the interpretation of the state from one time slot to the corresponding unit resource required by streaming services, this model is also suitable for modelling RT streaming services in GSM.

3.3 Analytical Model for Downlink Voice Service in UMTS

The analytical model for voice service in UMTS presented here follows in principle the model proposed in [82]. There are some differences and extensions including:

- The activity factor of a voice call is assumed to be 1. That means a voice call needs a dedicated channel throughout the call duration.
- The multipath fading effect on the signal link is ignored, while a non-orthogonal factor is considered to involve the intra-cell interference due to Multipath fading.
- The method calculating the sum of several log-normal distributed variables follows a new one proposed in [83].
- Mobility effect on the user number is considered.
- Power reservation is taken into account.

3.3.1 Interference Model

The pathloss between a mobile station and a BS is modelled by a general accepted model, being the product of a fourth power of distance and a log-normal shadowing. Thus the pathloss is given by $L_{k,i} = d_{k,i}^{-4} \cdot \chi_k = d_{k,i}^{\ -4} \cdot 10^{\xi_k/10}$, where $d_{k,i}$ is the distance between MS i and BS k; ξ_k is a Gaussian distributed random variable with zero mean and a standard deviation σ. Fast fading on the received power of the user data is assumed not to affect the average power level.

An MS faces the interference from its own cell and the interference from the surrounding cells. With perfect orthogonality of channelisation codes, the TX power for other simultaneous downlinks and common control channels should not introduce interference to the objective link. However, the intra-cell interference is supposed arising from the product of a non-orthogonal factor and the total downlink TX power subtracting the TX power for the objective link. The extra cell interference is arising from the total TX power of the neighboring BSs. Thus the SIR of MS *i* in cell *k* is given by:

$$SIR_i = \frac{P_k \cdot \varphi \cdot \phi_{k,i} \cdot L_{k,i}}{\alpha \cdot I_{i,Intra} + I_{i,Extra} + N_0} \tag{3.7}$$

$$I_{i,Intra} = P_k (1 - \varphi \cdot \phi_{k,i}) \cdot L_{k,i}$$

$$I_{i,Extra} = \sum_{j \neq k} P_j \cdot L_{j,i}$$

φ	fraction of BS TX power assigned to traffic channels
$\phi_{k,i}$	relative TX power of cell *k* to MS *i*
$I_{i,Extra}$	extra-cell interference
$I_{i,Intra}$	intra-cell interference
P_k	total TX power of cell k
N_0	broad band thermal noise value
α	non-orthogonal factor

The bit energy noise density ratio is then

$$\begin{aligned}\frac{E_b}{N_0} &= \frac{W}{R} \frac{P_k \cdot \varphi \cdot \phi_{k,i} \cdot L_{k,i}}{\alpha \cdot I_{i,Intra} + I_{i,Extra} + N_0} \\ &= \frac{W\varphi}{R} \frac{\phi_{k,i}}{\alpha \cdot (1 - \varphi \cdot \phi_{k,i}) + \frac{I_{i,Extra}}{P_k \cdot L_{k,i}} + \frac{N_0}{P_k \cdot L_{k,i}}}\end{aligned} \tag{3.8}$$

In case that $N_0 << P_k \cdot L_{k,i}$ and $\varphi \cdot \phi_{k,i} << 1$, the function above can be simplified and derived to the function of link power.

$$\frac{E_b}{N_0} = \frac{W\varphi}{R} \frac{\phi_{k,i}}{\alpha + \frac{I_{i,Extra}}{P_k \cdot L_{k,i}}} \tag{3.9}$$

$$\phi_{k,i} = \frac{\frac{E_b}{N_0} \cdot (\alpha + \frac{I_{i,Extra}}{P_k \cdot L_{k,i}})}{\frac{W\varphi}{R}}$$

The fraction for traffic channels is assumed to be 95%. The target SIR ratio $\Gamma = \frac{E_b}{N_0} / \frac{W}{R}$ is set to -19.39933dB for voice communications. For a moving user, additional penalty of Γ, e.g. 0.6dB for pedestrian MSs, is considered imitating the imperfect power control. The non-orthogonal factor α is assumed to be 0.4. Assume the transmission power of a CS link is exactly adjusted to the level needed after power control. Substitute these parameters we get:

$$\phi_{k,i} = \frac{\frac{E_b}{N_0} \cdot (\alpha + \frac{I_{i,Extra}}{P_k \cdot L_{k,i}})}{\frac{W\varphi}{R}} = a + b \frac{I_{i,Extra}}{P_k \cdot L_{k,i}} \tag{3.10}$$

$a = 0.00484 \,; b = 0.0121 (stationary)$

$a = 0.00555; b = 0.01388 \, (pedestrian)$

$\frac{I_{i,Extra}}{P_k \cdot L_{k,i}}$ is the only variable component in this link power function, where $I_{i,Extra}$ and $L_{k,i}$ are random variables. The transmission power of a BS is a function of the on going calls served by it and thus should be modelled as a random variable. However, we assume the worst case is under investigation. The transmission power of a BS is hence the maximum power and with a unique value for every BS.

$$\frac{I_{i,Extra}}{P_k \cdot L_{k,i}} = \frac{\sum_{j \neq k} P_j \cdot L_{j,i}}{P_k \cdot L_{k,i}} = \sum_{j \neq k} \frac{{d_{j,i}}^{-4} \cdot 10^{\xi_j / 10}}{{d_{k,i}}^{-4} \cdot 10^{\xi_k / 10}} = \sum_{j \neq k} (\frac{d_{j,i}}{d_{k,i}})^{-4} \, \frac{10^{\xi_j / 10}}{10^{\xi_k / 10}} \tag{3.11}$$

To get the value of $\frac{I_{i,Extra}}{P_k \cdot L_{k,i}}$, we firstly look at a worst case shown in Figure 3-6. The antenna is assumed to be an omni antenna. A sample mobile station located at the right upper corner of BS 17 is considered to be a mobile station in the worst case. We take the same extra interference sources considered in [82]. The shadow (slow) fading in all the cells are modelled by an independent log-normal random variable with zero mean and same standard deviation σ.

$$S_i = \frac{I_{i,Extra}}{P_k \cdot L_{k,i}}$$

$$= \frac{\begin{array}{l} 10^{\xi_{j,1} / 10} + 10^{\xi_{j,2} / 10} + (2)^{-4} \cdot 10^{\xi_{j,3} / 10} + (2)^{-4} \cdot 10^{\xi_{j,4} / 10} + (2)^{-4} \cdot 10^{\xi_{j,5} / 10} \\ + (2.633)^{-4} \cdot 10^{\xi_{j,6} / 10} + (2.633)^{-4} \cdot 10^{\xi_{j,7} / 10} + (2.633)^{-4} \cdot 10^{\xi_{j,8} / 10} \\ + (2.633)^{-4} \cdot 10^{\xi_{j,9} / 10} + (2.633)^{-4} \cdot 10^{\xi_{j,10} / 10} + (2.633)^{-4} \cdot 10^{\xi_{j,11} / 10} \end{array}}{10^{\xi_k / 10}} \tag{3.12}$$

Note that in the numerator, each interference component is multiplied by a constant value arising from the distance ratio. A log-normal variable multiplied by a constant, $C \cdot 10^{\xi_j / 10}$, is still a log-normal variable, denoted by $10^{\xi m / 10}$. Assuming ξ_j is with mean, μ_j dB, and standard deviation, σ_j dB, ξ_m has then a mean value, $\mu_j + 10 \log 10(C)$ dB, and a standard deviation, σ_j dB.

The sum of log-normal variable can still be approximated to a log-normal variable. The approximation has been intensively studied in [84] - [87] since the summation of log-normal variables is very often used in engineering models. The method introduced by Schwartz [85] was used in [82]. However the Schwartz method is accurate in estimating the standard deviation only up to 8 equal log-normal variables with standard deviation less than 10dB. We take the method from [83] because it has been detected to have a good approximation in a wide range of standard deviation of the log-normal variable components. The calculation is given in Appendix A. Assume the log-normal sum of I_{extra} is 10^{ξ_s}. The mean and variance for ξ_s are μ_s *and* σ_s^2.

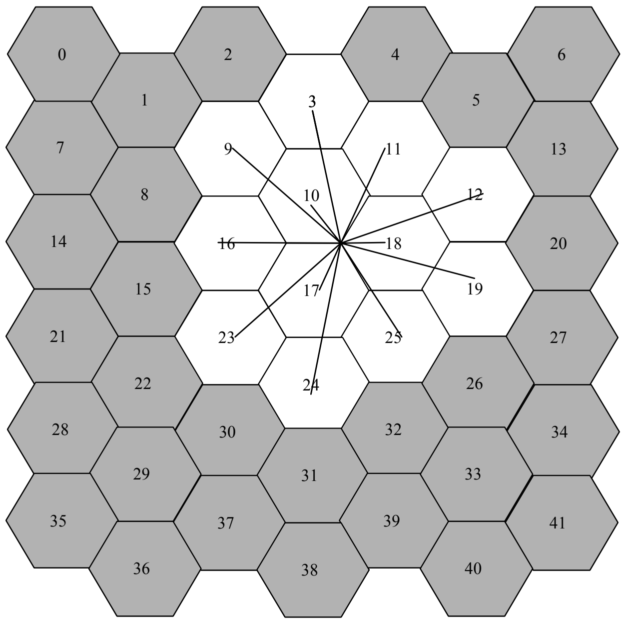

Figure 3-6: Worst case extra cell interference for a MS located at the corner of BS 17, which receives the interference from BS 10, BS16, BS 18, BS23, BS24 and BS25 in the first tier, and from BS 9, BS 3, BS 11, BS 12, BS 19 in the second tier [82].

3.3.2 Total Power

Since ξ_s *and* ξ_k are assumed to be mutually independent Gaussian distributions, the difference is still a Gaussian distribution, ξ_i and has a mean of $\mu_s - \mu_k$ and a variance of $\sqrt{\sigma_s{}^2 + \sigma_k{}^2}$. The normalized total TX power is then:

$$\sum_{i=1}^{N} \phi_i = \sum_{i=1}^{N} (a + b \cdot S_i) = Na + b \sum_{i=1}^{N} 10^{\xi_i / 10} \tag{3.13}$$

Till now the extra cell interference is aligned with the worst case. Normally uniform distributed traffic is adopted to better represent the reality of traffic. According to [82], the S_i of uniformly distributed users is $\eta = 0.4$ times the one in the worst case.

Given the number of users, N, and the reserved resources for handover calls, C_{res}, the probabilities of the total TX power exceeding the maximum power constraints for new arrivals and handover arrivals are given in Equation (3.14) and (3.15) as follows:

$$\begin{aligned} & Pr\left[\sum_{i=1}^{N} \phi_i > \varphi - C_{res}\right] \\ & = Pr\left[(Na + b\eta \sum_{i=1}^{N} 10^{\xi_i / 10}) > \varphi - C_{res} \right] \\ & = Pr\left[\sum_{i=1}^{N} 10^{\xi_i / 10} > \frac{\varphi - C_{res} - Na}{b\eta} \right] \end{aligned} \tag{3.14}$$

$$\begin{aligned} & Pr\left[\sum_{i=1}^{N} \phi_i > \varphi \right] \\ & = Pr\left[(Na + b\eta \sum_{i=1}^{N} 10^{\xi_i / 10}) > \varphi \right] \\ & = Pr\left[\sum_{i=1}^{N} 10^{\xi_i / 10} > \frac{\varphi - Na}{b\eta} \right] \end{aligned} \tag{3.15}$$

Again the sum of N independent log-normal variables $\sum_{i=1}^{N} 10^{\xi_i / 10}$ can be approximated to a log-normal distribution $\chi = 10^{\xi_t / 10}$. The mean and standard deviation of ξ_t are denoted by $\mu_t \text{ and } \sigma_t$. The probabilities, of which excessive power is required by new arrivals and handover calls, are given in Equation (3.16) and (3.17):

$$Pr\left[\sum_{i=1}^{N} \phi_i > \varphi - C_{res}\right] = Pr\left[\chi > \frac{\varphi - C_{res} - Na}{b\eta} \right] = Q\left(\frac{\dfrac{ln \dfrac{1 - C_{res} - Na}{b \cdot \eta}}{\theta} - \theta \cdot \mu_t}{\theta \cdot \sigma_t} \right) \tag{3.16}$$

$$Pr\left[\sum_{i=1}^{N} \phi_i > \varphi \right] = Pr\left[\chi > \frac{\varphi - Na}{b\eta} \right] = Q\left(\frac{\dfrac{ln \dfrac{1 - Na}{b \cdot \eta}}{\theta} - \theta \cdot \mu_t}{\theta \cdot \sigma_t} \right) \tag{3.17}$$

$$where\ Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{y^2}{2}} dy$$

$$\theta = ln\, 10 / 10$$

3.3.3 User Number

The number of users can be obtained by an M/M/∞ model. This model has intrinsic some weakness in modelling UMTS traffic. Firstly, requiring a certain level of QoS, the number of users in a UMTS cell with a maximum total transmission power constraint cannot increase to infinity. Secondly, some transmission power is reserved for handover calls. After the resources for new arrivals are completely occupied, the transit probability (rate) from the state *i-1* to *i* should be only λ_h instead of the sum of $\lambda + \lambda_h$. That is, a new arrival cannot lead to the increment of the number of users at this stage. However, given a maximum transmission power constraint, the capacity in terms of the number of supportable active users corresponding to the resources is normally a random number depending on the varying radio conditions. Just as the number of users that the shared power can support is not deterministic, the number of users that the reserved power can support is not deterministic, either. As a consequence, a fixed boundary state between the total arrivals and the handover arrivals as well as a maximum accommodation is impossible.

The arguments for taking an M/M/∞ model are twofold. First, this model is relative accurate if the system has a large number of servers, each is in accordance with one user inside one same cell. This is especially helpful because blocking and dropping happens normally when the system contains a large number of users. Second, the reserved resources are normally small compare to the total resources. So the inaccurate arrival rate occurs at only few transitions. The transit diagram is given in Figure 3-7. The steady state probability of state *i* is given in Equation (3.18) where the coefficient 1.3 stands for the traffic increase by soft handovers.

$$Pr(i) = \frac{1.3\left(\frac{\lambda + \lambda_h}{\mu + \mu_h}\right)^i}{i\,!} e^{-1.3\left(\frac{\lambda+\lambda_h}{\mu+\mu_h}\right)} \tag{3.18}$$

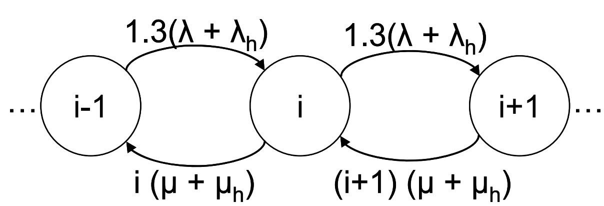

Figure 3-7: Transit diagram for an M/M/∞ chain modelling the number of active UMTS users

3.3.4 Blocking and Dropping in a UMTS Cell

Outage probability is equal to the sum of the probabilities where the interference level is exceeding the interference limit (threshold).

In contrast to the UMTS uplink, where the received multiple access interference (MAI) at the Node B defines the cell capacity, the downlink capacity is limited by the maximum transmission power. Due to the fact that interference is a very important issue in a UMTS analytical model, the blocking and dropping probability is indeed a joint probability of the interference limit and the user number, see Equation (3.19) and (3.20).

$$B = \sum_{N=1}^{\infty}(Pr\left[\sum_{i=1}^{N}\phi_i > 1 - C_{res}\right]\cdot Pr(N)) \tag{3.19}$$

$$= \sum_{N=1}^{\infty} Q(\frac{\frac{ln\frac{1-C_{res}-Na}{b\cdot\eta}}{ln10/10}-\mu_t}{\sigma_t})\cdot\frac{1.3(\frac{\lambda+\lambda_h}{\mu+\mu_h})^N}{N!}e^{-1.3\frac{\lambda+\lambda_h}{\mu+\mu_h}}$$

$$D = \sum_{N=1}^{\infty}(Pr\left[\sum_{i=1}^{N}\phi_i > 1\right]\cdot Pr(N)) \tag{3.20}$$

$$= \sum_{N=1}^{\infty} Q(\frac{\frac{ln\frac{1-Na}{b\cdot\eta}}{ln10/10}-\mu_t}{\sigma_t})\cdot\frac{1.3(\frac{\lambda+\lambda_h}{\mu+\mu_h})^N}{N!}e^{-1.3\frac{\lambda+\lambda_h}{\mu+\mu_h}}$$

The numerical results of blocking and dropping are shown in Figure 3-8 and Figure 3-9, respectively. Scenarios with pure pedestrian mobiles and pure pedestrian mobiles are taken to clarify the mobility effect. The maximum transmission power is 20W, among which 20% is reserved for handover calls if the resource reservation is enabled. A reduced resource reservation with a fraction of 0.1 is also tested and shown. The standard deviation of log-normal shadow fading is 10dB. The speed of pedestrian users is 3km/h. The handover arrival rate and departure rate are following the same calculation as in 3.2.2. The parameters *a* and *b* are given in Equation (3.10) in accordance with the mobility type.

Recall the intrinsic downsides of modeling the number of users with the M/M/∞ chain. Specifically, the benefit of power reservation for handover calls is not confirmed by the dropping obtained, see Figure 3-9.

Resource Reservation Influence

It is apparent in Figure 3-8 that the more resources reserved for handover calls, the higher the blocking will be. As mentioned previously, the exact number of new arrivals which transmission power aggregates up to the limit is impossible to be identified. The corresponding dropping calculation cannot reveal the benefit of resource reservation since the new arrivals continuously share the resources reserved for handover calls in this model indeed.

The importance of resource reservation has been proved in the GSM model. It should be involved in the UMTS local resource management as well. Firstly, due to the relative high capacity indicated from the analytical model (more than 6 times more input traffic at 5 % blocking for pure pedestrian mobiles without resource reservation compared to a GSM cell), it is necessary to reserve resource for more than one link. Secondly, considering the interference limited characteristics, the reserved resource should be able to support a mobile which is in a worst location, i.e. in a most resource-consuming situation. Suppose the maximum link power budget is 2W, resource reserved for 2 links at worst case requires 4W.

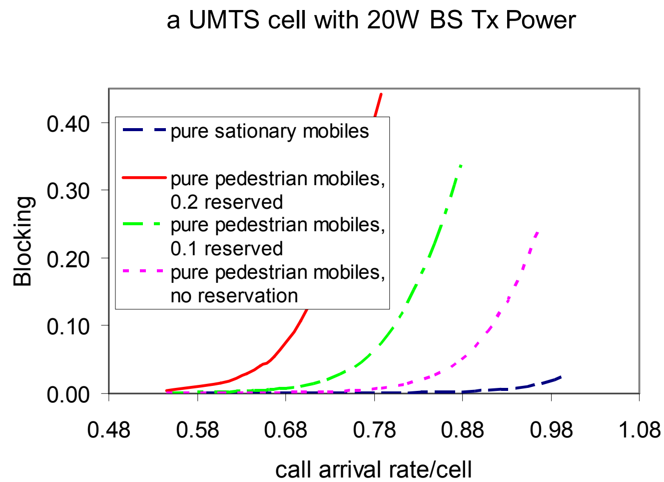

Figure 3-8: Mobility influence and resource reservation effect on voice service blocking through a joint interference and Markov model

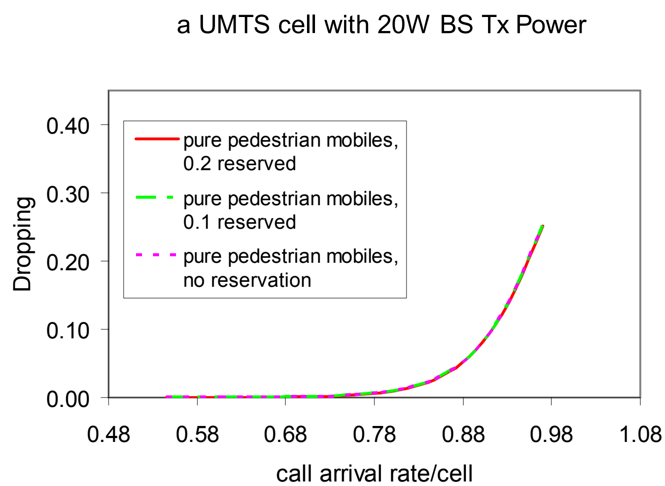

Figure 3-9: Mobility influence and resource reservation effect on voice service dropping through a joint interference and Markov model, three curves overlap themselves

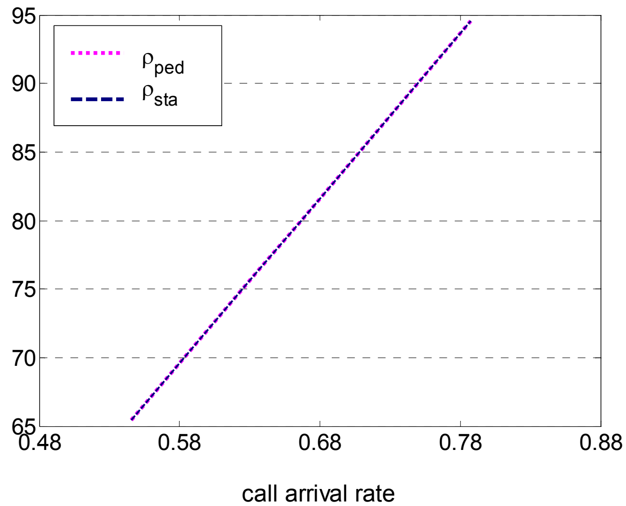

Figure 3-10: Mobility impact on traffic intensity of a UMTS cell

3.4 Comparison between Analytical and Simulation Models

3.4.1 Event-driven Simulation Model

Assume that omni antennas are located at the centre of the cells illustrated in Figure 3-11: and the GSM and UMTS systems have the same cell layout. The transmission power and the antenna gain of a GSM cell is configured in a way that the EIRP (equivalent isotropic radiated power) can exactly cover a circular area with the same radius as the hexagon cell. This strategy helps to re-present the strict grid of GSM cells in the simulation model as well. As for the UMTS system, the based station transmission power should cover an area which exceeds the cell area according to the cell grid. In this way the overlapped area between UMTS cells offers the possibility for soft handovers.

In the event-driven simulator, events and the consequent actions such as call generation, resource request, resource allocation, user movement, link measurement, handover request-decision-execution, call dropping, call termination, and power control are modelled in a time sequence. The voice traffic is created by a call generator which models a new call arrival in accordance with a Poisson process. For every new arrival, an exponentially distributed value is assigned for the call duration. A mobile station becomes active and associated with a statistical created call and a uniformly distributed location inside the rectangular area in Figure 3-11. A communication link is then established between the corresponding mobile station and the base station achieving the best link quality among the cells of the same system.

The intrasystem handover is mobile-assisted by the link measurement report of an active mobile station. Hard handovers are realised for GSM calls, where no handover hysteresis is considered. A mobile station maintains an active

BS set and a monitored BS set in order to model hard handovers and soft handovers for UMTS calls according to the mechanisms described in Chapter 2.

Power control algorithm is triggered when 1) a new call is admitted, 2) an existing call is terminated or handover to another cell, 3) the extra interference of any monitored cell is changed and 4) the distance travelled is equal to the decorrelation distance of the shadow fading [92]. Nevertheless, only when the measured SIR differs from the target value more than the power control step, a power control or equivalently a power update will be executed.

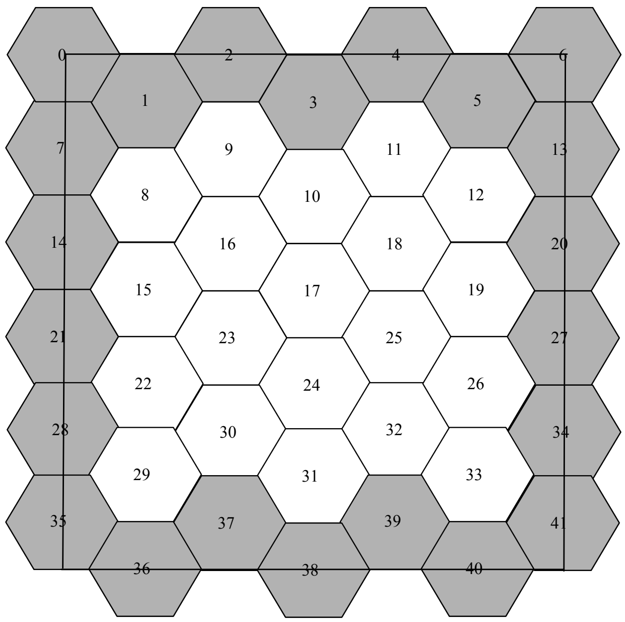

Figure 3-11: Cell layout of GSM and UMTS systems with omni antennas

UMTS Power Allocation

The voice service has absolute priority over NRT services as a typical admission control mechanism mentioned in Section 2.1.1. Hence the voice users are free to be allocated up to the maximum transmission power. A target SIR ensures the signal quality in terms of maximum bit error rate. Thus the link quality is guaranteed by allocate adequate transmission power for voice users. Due to multipath effect, the channelisation codes in a cell cannot be 100% orthogonal. Thereafter transmission power for other users in the same cell may cause the intra-cell interference, i.e. the adjacent channel interference. The SIR of a voice user is given in Equation (3.21).

$$SIR_i = \frac{S_i}{\alpha \cdot I_{i,Intra} + I_{i,Extra} + N_0} \tag{3.21}$$

$$SIR_i = \frac{P_{k,i} / L_{k,i}}{\alpha \cdot (P_k - P_{k,i}) / L_{k,i} + \sum_{j \neq k} P_j / L_{j,i} + N_0}$$

$I_{i,Intra}$ intra-cell interference [W]
$I_{i,Extra}$ extra cell interference [W]
N_0 broad band thermal noise [W]
α non-orthogonal factor

$L_{k,i}$ path loss of user i in cell k including shadow fading and antenna gain
P_k total TX power [W] of BS k
$P_{k,i}$ TX power [W] for user i in BS k

GSM/EDGE Link Quality

The adjacent channel interference can be neglected due to good separation of the channels in the frequency domain. According to [93], the typical adjacent channel protection is 18 dB. Therefore, only the co-channel interference is taken into account. Due to the high reuse factor, the distance between the mobile and the interfering cells is several times the cell radius. Therefore, the interference of the co-channel cells does not change very much if the mobile moves around within the serving cell. This has been verified by simulation [94]. Therefore, it is assumed that the co-channel interference is constant i.e. does not depend on the position of the mobile. The downlink SIR of a GSM user i is calculated following Equation (3.22), not distinguishing the service type.

$$SIR_i = \frac{P_{k,i} \,/\, L_{k,i}}{I_0} \tag{3.22}$$

I_0 constant interference and noise value [W]
$P_{k,i}$ TX power [W] for user i in BS k
$L_{k,i}$ path loss of user i in cell k including shadow fading and antenna gain

3.4.2 Comparison between Analytical and Simulation Results

Table 3-1 lists the system configuration for the GSM and the UMTS simulation models, respectively.

The analytical results are very close to the simulation results as shown in Figure 3-12 and Figure 3-13. Nevertheless analytical blocking and dropping are larger than the simulation results. The difference is caused by the fact that a real antenna will cover ideally a circular range without the effects of slow and fast fading while the strict coverage separation of the analytical model can be represented by a hexagon grid, see Figure 3-14. To guarantee the coverage of a hexagon area, the simulated GSM cell coverage is larger than the former one. The fact is then leading to overlapping between neighbour cells. The fact of overlapping of cell coverage results in some probability that a call experiencing insufficient quality in cell A may be served by cell B if it is also in the coverage of cell B. Consequently, the blocking and dropping are reduced.

GSM total time slots	21 (3TRX)
GSM link TX power	1W
GSM interference	-105dBm
Minimum GSM voice link SIR	9dB
GSM capacity reservation for handover calls	5%
UMTS carrier	1
UMTS max. BS TX power	20W
UMTS thermal noise	-100dBm
UMTS capacity reservation for handover calls	20%
Minimum UMTS voice link SIR	-19.39933dB
UMTS shadow fading: standard deviation	10dB
UMTS Soft handover additive hysteresis	2.5dB
UMTS Soft handover removal hysteresis	4.5dB
UMTS Intrasystem handover hysteresis	2.0dB
GSM/UMTS cell radius R	800m
GSM/UMTS BS separation	2400m
GSM/UMTS SIR penalty for pedestrian	0.6dB
UMTS handover hysteresis timer	0.25s

Table 3-1: System configuration for the simulation model

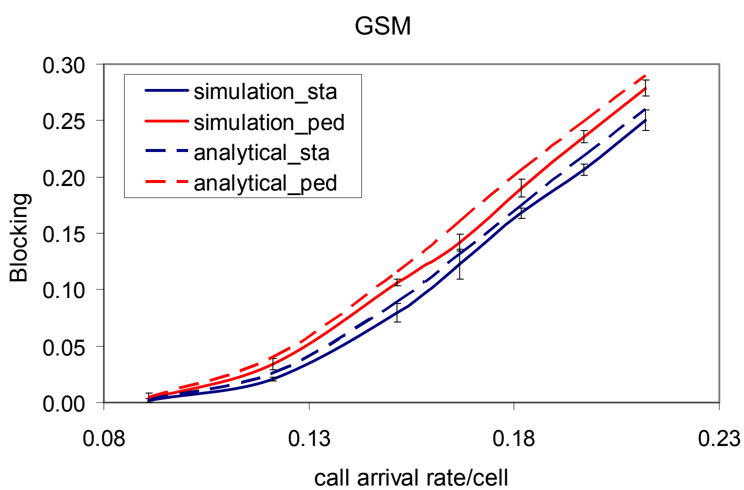

Figure 3-12: Comparison between the analytical and simulation GSM blocking where the 99% confidence intervals are combined with the simulation results

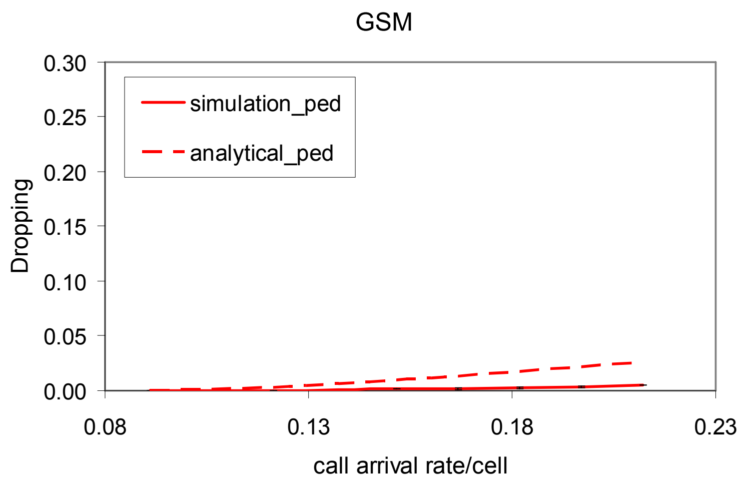

Figure 3-13: Comparison between the analytical and simulation GSM dropping where the 99% confidence intervals are combined with the simulation results

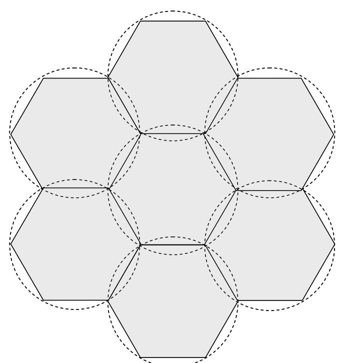

Figure 3-14: Circular coverage of a GSM BS is larger than the hexagon cell and overlapped with the coverage of the neighbouring cells

The UMTS results are given in Figure 3-15and Figure 3-16. The simulation results are slightly apart from the analytical results (labelled analytical_sta/ped_soft1.3) where the effective input traffic is assumed 1.3 times, increased by soft handovers. Especially in the case of pure pedestrian users, the inaccurate traffic assumption has larger influence on the analytical results because the target SIR value required by a pedestrian user is larger than required by a pedestrian user. The average legs of voice calls resulted from the simulator is 1.22 indeed. By substituting the traffic increase factor 1.22 into the analytical model the results (labelled analytical_sta/ped_soft1.22) are again quite close to the simulation results.

The fact of simulation results being close to the analytical results verifies the simulation results.

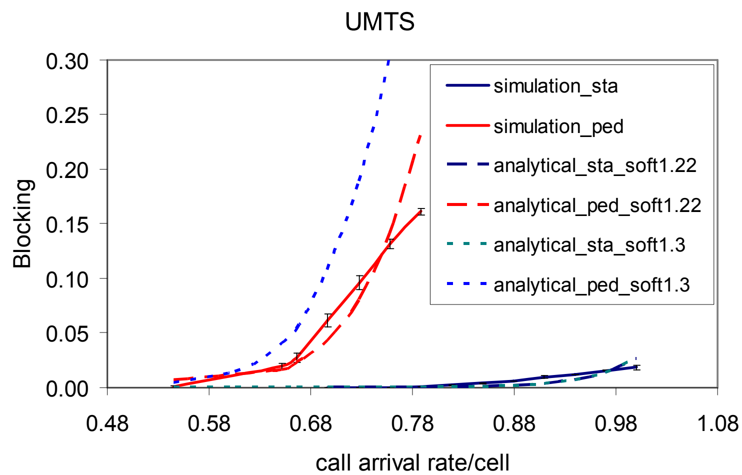

Figure 3-15: Comparison between the analytical and simulation UMTS blocking where the 99% confidence intervals are combined with the simulation results

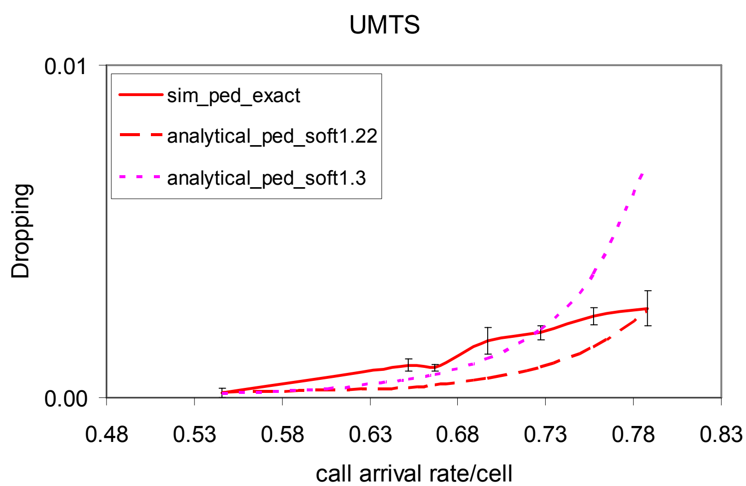

Figure 3-16: Comparison between the analytical and simulation UMTS dropping where the 99% confidence intervals are combined with the simulation results

3.4.3 More Features Modelled in the Simulation Model

Although the simulation results well agree to the analytical results, plenty of features of the respective networks are not involved into the models. The complexity level of the modelling has in fact a direct influence on the system capacity and performance. In this section, we firstly integrate a limited part of advanced features aiming to illustrate the influence of the additional features. The integrated features are listed as follows:

- Radio channel characteristics for GSM radio links including the propagation loss model and the additional shadow fading.
- Experimental propagation loss models for GSM and UMTS radio links.
- A pedestrian mobility type models the behaviour of a slow moving user who potentially changes the direction often and without preference. But the

users are moving with a constant velocity. In contrast, a straight moving trajectory is assumed before.

- A real interference calculation is carried out whenever the intra- or inter-cell interference varies in the UMTS simulation model. In the UMTS analytical model, the total interference calculation is simplified to be proportional to the worst case interference.
- The most power-consuming link in a UMTS cell is selected and recommended for dropping instead of directly dropping the handover call.
- The reserved resource is guaranteed to be used for handover calls alone. The resource (TX power) reservation is not conserved for handover calls only in the UMTS analytical model.
- The handover hysteresis and the handover retry for GSM calls are realised in the simulation models. The handover retry for UMTS calls are realised in the simulation models but not in the analytical model.

Simulation	***Analytical model***
An experimental path loss model	-
The shadow fading	-
the cell coverage determined by a SIR threshold	Geographical cell grid
a random moving trajectory but with a constant velocity	Constant velocity and constant moving direction
Handover hysteresis and handover retry	-

Table 3-2: Summary of differences between GSM analytical and simulation models

Simulation	***Analytical model***
An experimental propagation loss model	$10\log10(d^4)$
A real-time interference calculation	$0.4 \cdot$ Worst case interference
Reserved power for handover calls only	The boundary of user number in accordance with the reservation is ambiguous.
a random moving trajectory but with a constant velocity	Constant velocity and constant moving direction
Recommend the most power consuming call for dropping	Direct drop the handover call if resources are not sufficient
handover retry	-

Table 3-3: Summary of differences between UMTS analytical and simulation models

Equation (3.23) gives the channel propagation model for a GSM link, where the vehicular propagation loss is the urban COST 231-Walfish-Ikegami model proposed in [95] for GSM 900. The base station height is substituted by 10m and the mobile station height 1.5m.

$$
\begin{aligned}
&r < 0.027: \quad L = 0 \\
&r < 5.10: \quad L = L_{FS} = 31.53 + 20 \cdot lg(r) \\
&r \geq 5.10: \quad L = L_V = 18.8 + 38 \cdot lg(r)
\end{aligned}
\tag{3.23}
$$

L propagation loss [dB]
L_V "vehicular" propagation loss [dB]
L_{FS} free space loss [dB]
r distance between BS and MS [m]

Equation (3.24) gives the channel propagation model for a UMTS link, where the vehicular propagation loss is derived with the assumptions of 10m base station height, 1.5m mobile station height and 2GHz carrier frequency [96]. If the vehicular propagation loss is even less than the free space loss, the free space loss will be adopted.

$$
\begin{aligned}
&r < 0.012: \quad L = 0 \\
&r < 20.7: \quad L = L_{FS} = 38.46 + 20 \cdot lg(r) \\
&r > 20.7: \quad L = L_V = 15.3 + 37.6 \cdot lg(r)
\end{aligned}
\tag{3.24}
$$

Similarly as the analytical model, a normal distributed shadow fading is added to the propagation loss to produce the total pathloss, all in unit of dB.

Admitting the experimental propagation loss model, the propagation loss is enlarged around 13dB for a GSM link while around 8dB for a UMTS link, compared to the models with the pathloss exponent being 4, see Figure 3-17 and Figure 3-18.

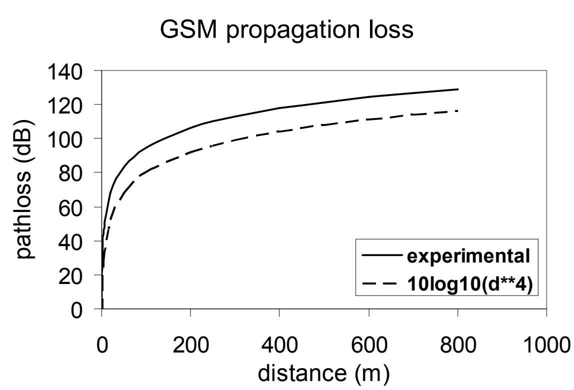

Figure 3-17: GSM experimental propagation loss compared to the propagation loss with an exponent of 4

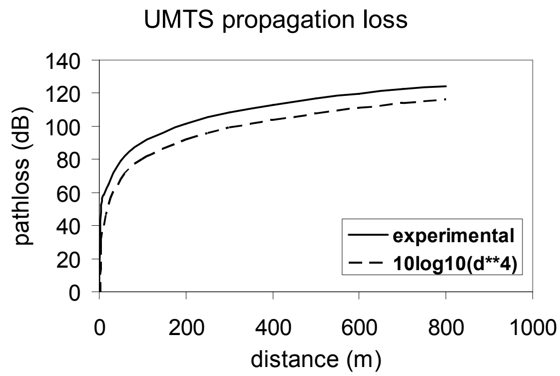

Figure 3-18: UMTS experimental propagation loss compared to the propagation loss with an exponent of 4

3.4.4 Influence of the Additional Features

The results from the simulations involving the additional features given in the last section are compared to the results from the simulations exactly modelling so many features as the analytical models have. The former ones are labelled as advanced simulation results. The later ones are labelled as exact simulation results.

The additional features reduce the blocking of GSM voice service (Figure 3-19) while increase and flatten the dropping of GSM voice service (Figure 3-20). As mentioned previously, the GSM coverage determined by the minimum SIR threshold should be a circular area centred at the BS when the shadow fading is ignored. The consideration of random shadow fading causes the cell boundary becoming irregular. To guarantee the hexagon cell coverage, the advanced simulation model has to enlarge the antenna gain and in turn the EIRP. So the overlapping area of adjacent cells in fact larger than without the consideration of shadow fading. The lower blocking is arising from the enlarged overlapping of cell coverage.

The enlarged dropping rate is mainly caused by the complex mobility model and the handover hysteresis. A dropped call corresponds to a handover call which is normally at the edge of a cell. Firstly, according to the mobility model, users are often changing the moving direction without preference. The probability of crossing cell boundaries is higher than the case of constant moving direction. Secondly, the handover hysteresis reduced the ping-pong effect of handovers which is not concerned by the analytical model. Nevertheless this handover hysteresis effectively enlarges the distance a handover call has to travel and consequently increases the handover traffic intensity. Moreover, the mitigated blocking indicates that the new arrivals share more resource than in the analytical model. Given a fixed amount of total resources, the remaining resources for handover calls are reduced. So involving the radio characteristics (pathloss, shadow fading and interference), the advanced mobility model and handover criteria, the dropping is high. However, the possibility of handover retry can prevent the dropping increase and flatten the dropping curve (Figure 3-20).

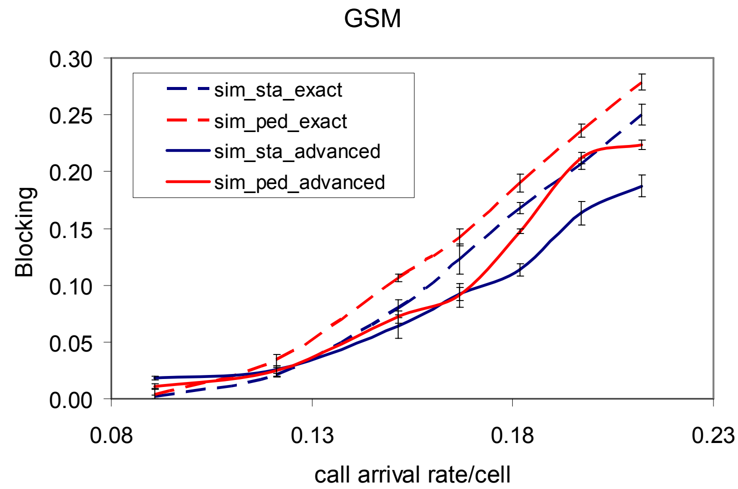

Figure 3-19: Comparison of GSM blocking between the exact and advanced simulation results where the 99% confidence intervals are combined

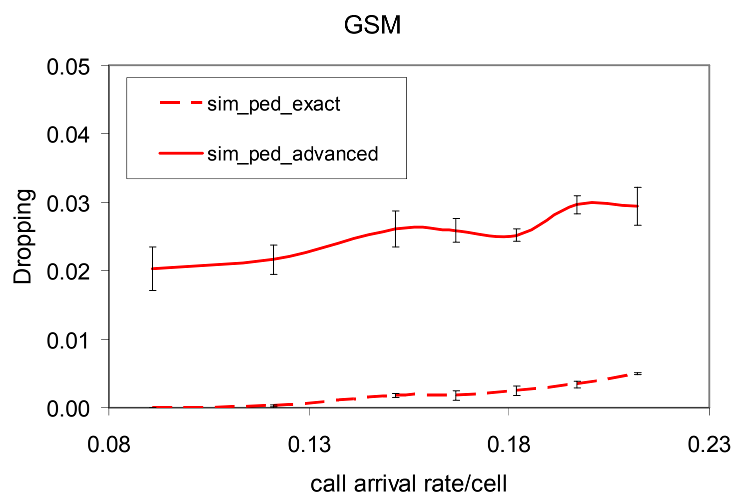

Figure 3-20: Comparison of GSM dropping between the exact and advanced simulation results where the 99% confidence intervals are combined

The exact simulation model of UMTS has employed already the shadow fading, the interference model and the handover hysteresis since these issues are considered in the analytical model. Recall that the additional UMTS features considered are the experimental propagation loss, the recommendation of maximum power-consuming call for dropping, the handover retry scheme and a complex mobility model for pedestrian users. The enlarged propagation loss increases the blocking rate, see Figure 3-21.

Despite the negative influence of often change of moving direction on the dropping rate, dropping the most expensive link and handover retry mitigates the dropping. Moreover, the handover retry flatten the dropping rate similarly as the impact in the GSM network. Note that these influences are really micro influences with respect to the dropping scale. For UMTS, a cell has reserved so much power that it can satisfy the power requirement for two extremely worst links. The extremely worst link is defined by the 2W maximum link power limit. Thus the guard resources preventing dropping is quite sufficient, 20% of the total transmission power. The resulted dropping rate shrinks to a very limited range, less than 1% in Figure 3-22.

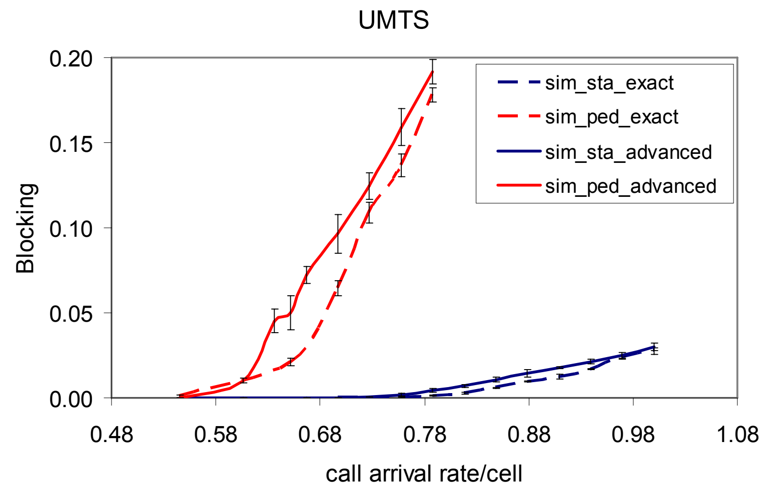

Figure 3-21: Comparison of UMTS blocking between the exact and advanced simulation results where the 99% confidence intervals are combined

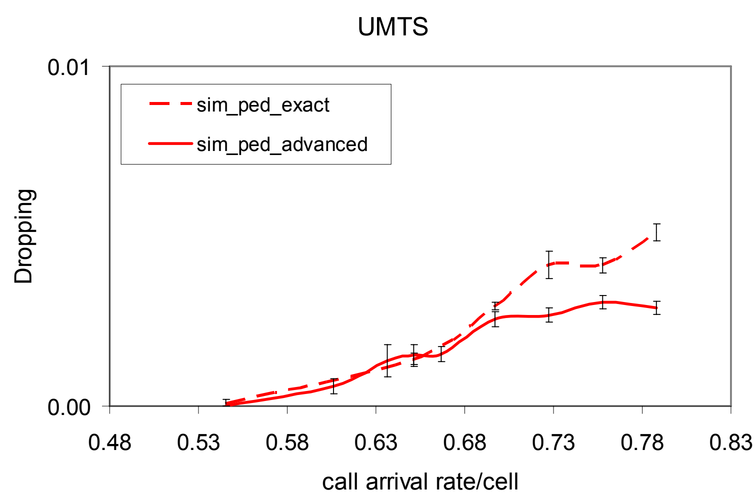

Figure 3-22: Comparison of UMTS dropping between the exact and advanced simulation results where the 99% confidence intervals are combined

3.5 Summary

Analytical models for GSM voice traffic and UMTS voice traffic are implemented to study the system capacity for voice traffic, respectively. The mobility influence and the resource reservation strategy are investigated. The well known capacity degradation due to user mobility is verified by the numerical results. A minimum resource reservation strategy is found efficiently reducing the dropping probability of moving users.

The analytical results are compared with the simulation results, where the simulations model exactly the features considered in the analytical models. The fact that simulations produce very close performance to the analytical results validates the simulation models of the GSM and UMTS networks.

The complexity of a real network makes it difficult for an analytical model to reveal any features. Some additional features involved into the simulation model have shown that the system features do vary the system performance significantly. On the other hand, many detailed system features are not involved in the analytical model or at this moment a complex analytical model considering the detailed system features are not available. Examples of these features are sectorised antenna, mixed mobility pattern (e.g. co-existing pedestrian and vehicular users), handover hysteresis and handover retry etc. Moreover, the multi-radio resource management at call level is difficult to be studied with an analytical model. In the following MRRM study, we use the computer simulation to learn the MRRM of integrated voice and data services based on detailed underlying radio access networks at call level.

Chapter 4

MRRM for Integrated Services in Multi-Networks of GETRAN and UTRAN

This thesis focuses on the study of the network-based MRRM. The network, via the traditional radio resource management functionality, has better overview about the load status and the service requirements of various users. This overview knowledge enables the network-based MRRM to achieve fairness among networks, cells and users. At the user terminal side, such overview information is either not necessary, such as the load status of a network, or not authorised, such as the service requirements of other users. If a user terminal need execute a fairness algorithm, it must require exclusive signalling for collecting the load status and the requirements of other users. Otherwise a user controlled MRRM will adjust to a single user's preference. The effect of user preference on a MRRM algorithm might be unilateral and consequently might not efficiently utilise the scarce radio resources.

We propose MRRM strategies for integrated voice and data services in a heterogeneous GSM/UMTS scenario, which manage initial RAT selections as well as IS-HOs. The underlying radio access networks also plays key roles besides the MRRM algorithms. However, some already well deployed network features are too complex to be modelled via a comprehensive mathematic/stochastic model, e.g. random mobility, link adaptation of EDGE/HSDPA, interleaving, and handover hysteresis. On the other hand these features are quite influential factors on individual system capacity and performance. Accordingly, the study, based on the more realistic network simulator, takes into consideration the relationship between the realistic characteristics of underlying networks and the performance resulted from MRRM.

In this chapter, the possible MRRM Architecture is firstly discussed. The proposed MRRM algorithm for multiservices is then presented together with the review of two existing MRRM strategies. The capacity gain and the performance evaluation with various MRRM strategies will be discussed based on simulation results. Parts of this chapter are presented in [97].

4.1 MRRM Architecture

It is apparent that the physical location of MRRM functionality in the two networks will definitely affect the availability of system information at MRRM or the complexity of the methods gathering the system information necessary for MRRM.

Since GSM and UMTS are designed to have a common core network in the 3GPP project, a straight idea of the location will be a centralised MRRM in the core network having interfaces with RNC and BSC, respectively. The protocol architecture of a centralised MRRM is illustrated in Figure 4-1. If the multi-radio access networks belong to various operators, the implementation of such interfaces may need also negotiation or contracts between operators. In addition to the necessary implementation of interfaces, all the system information about UTRAN and GERAN transferred to MRRM, required by a working MRRM strategy, will bring about signalling overhead. MRRM will be therefore motivated only when the benefit is considerable large.

Another MRRM realisation is distributed MRRM. In this scenario, MRRM will be located in each radio access network which comprises the multi-radio access networks. The MRRM functionality may be even co-located inside the same physical entity together with the local RRM, illustrated in Figure 4-1. One advantage of the distributed MRRM is that at least the system information of the local network is directly referable. At the first glance, the signalling overhead caused by MRRM will be reduced to a half compared to the centralised MRRM. Moreover, some legacy signalling protocols, e.g. IS-HO signalling, between GERAN and UTRAN can be further adapted to serve for the information collection. Utilising the legacy signalling protocols may even completely cancel the signalling overhead MRRM needed. This mechanism will be exploited in the next chapter.

The distributed MRRM architecture also provides convenience for multi-vendor or multi-operator access selection. The concept of distributed MRRM was elaborated both by 3GPP [24] and IST EVERST [23]. The distributed MRRM entities may adopt a special MRRM strategy for each, not necessary adhering to a unique MRRM strategy. As long as the MRRM entity collects adequate input parameters, the multi-resource management of multi-operator or multi-vendor can proceed. Although this type of MRRM may be suboptimal due to the lack of centralised management, it offers flexibility to enable MRRM in a multi-operator and multi-vendor environment.

In this chapter, we will discuss the centralised MRRM performance of integrated services. The distributed MRRM performance will be analysed in the next chapter.

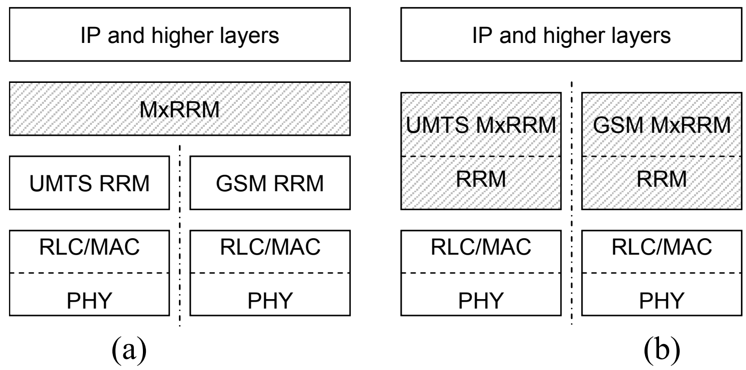

Figure 4-1: Illustration of MRRM protocol architecture (a) centralised MRRM (b) distributed MRRM

4.2 MRRM Algorithm

Two types of representative services are considered: the conventional voice telephony service, standing for real-time CS services, and the web browsing data service, standing for non-real-time PS services. QoS requirements of the two types of services are completely different. The voice service is delay sensitive and requires a guaranteed transmission rate throughout the call duration. The NRT services, defined as best effort services, have no stringent requirements for delay and bit rate on the 3GPP air interfaces [98]. The web browsing service requires, however, low bit error rate. Retransmission is usually the most popular mechanisms to improve the bit error rate of PS services. Hence the anticipation for a large PS transmission rate precedes the anticipation of reducing the delay caused by IS-HO. Regarding the interaction between services, CS services are assumed to have absolute priority over PS services in order to reduce the CS blocking and dropping.

In this section, we firstly discuss the challenges of intersystem handover, and the challenges of comparing load of various systems. Then review two existing MRRM strategies. Then a new approach based on the load status is given.

4.2.1 Challenges of Intersystem Handover

As mentioned in the introduction chapter, current studies about MRRM algorithms and policies are mostly concentrated on the initial RAT selection stage ([26]-[31], [33]-[35]). It is commonly agreed that IS-HO is able to further improve the efficiency of resource utilisation, i.e. IS-HO is anticipated to increase the capacity of integrated multi-radio access networks. However, only few researchers have explored the MRRM algorithms enabling IS-HO ([25], [32]). The capacity gain studied in [25] and [32] is focused on single type of service but not integrated services.

The IS-HO criteria are possibly involving the issues of resource utilisation, user preference, operator preference, energy limitation of user terminal, and economic benefits etc. Simply from the resource utilisation point of view, the possible IS-HO criteria are listed as follows:

- No sufficient resources in the source system
- Load the in source system is greater than a threshold while the load in the target system not.
- The target system is less loaded than the source system by a hysteresis.
- Load the in source system is greater than a threshold while the load in the target system is not. In addition, the target system is less loaded than the source system by a hysteresis.

Specifically, the occurrence of IS-HO should not be very often for CS services to minimise the signalling overhead and the signal delay. The load in various systems should be comparable to make a suitable IS-HO decision. The capacity should be optimised and the user QoS should be optimised.

Regarding the IS-HO of NRT services, the packet loss of the IP layer due to IS-HO should be as small as possible so that it can be compensated by the TCP. If we consider a web browsing type service, a user will download a web page, read, and then download the next page and so forth. Referring to the web browsing model in [96], a web page (packet call denoted there) contains several IP datagram which sum up to 96kbits data in average. It takes 1 second to download with a 96kbps data rate. Too often IS-HO should be hindered, thus the starting point of each web page becomes a good trigger for selecting a suitable RAT.

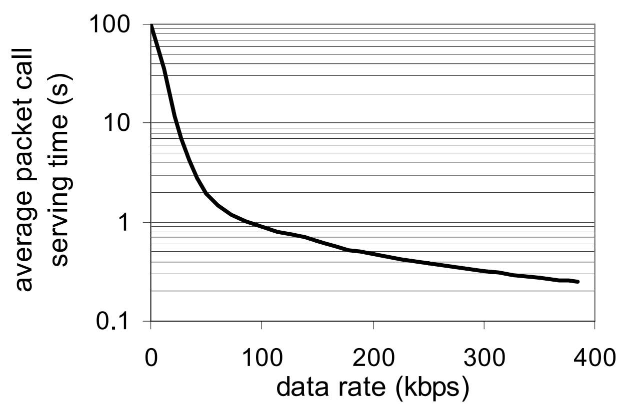

Figure 4-2: Serving time of packet call containing 96kbits with varying transmission rate available

4.2.2 Challenges of Load Comparison

The definition of load factor makes it possible to compare the load level in various systems. The load factor of various RANs must be well formulated to make the load factors reflect the most representative load situation as well as to make them comparable.

The load balancing algorithm in [26] makes the initial RAT selection based on the load factors which represent the total load status in a cell. However, due to the non-proportional relation between the total load factor they defined and the data throughput, data services cannot benefit from this algorithm. Their simulation results have also shown that the overall throughput gain compared to the one with a service based policy is at the expense of interactive traffic performance. Hence it is necessary to define a parameter which can better represent the load level of data services.

Therefore various load factors are necessary for representing different categories of services.

4.2.3 ” Blind” for Voice and Data to UMTS (Blind_DU)

The commercially available blind IS-HO between GERAN and UTRAN deals with the voice traffic, according to 3GPP R99. The data service is allocated to UTRAN whenever within its coverage.

A detection of bad channel quality imperatively triggers Blind IS-HO, both at call setup stage and during call connection stage. It means voice calls are to stay in the default RAT as long as possible. If the source network has no sufficient resources to admit a new arrival or becomes unable to maintain an ongoing CS call, this call will be overflowed/handed over to the cell with the best carrier to interference ratio (C/I) in the alternative network. This strategy still concerns the sole metric of signal strength, but no load information. Hence the target cell selected may eventually have insufficient resources either. In that case, the Blind IS-HO results in a failure.

The object call to be handed over to the alternative system is selected differently in GERAN and UTRAN. In GERAN the call experiencing insufficient channel quality is selected, while in UTRAN the most expensive call which consumes the most transmission power is selected. Nevertheless, the local radio resource management carries out this selection.

The data service is allocated to UTRAN if within the coverage based on the prerequisite condition that a UTRAN cell has more data capacity than a GERAN cell. Overflow of data service from UTRAN to GERAN is not enabled. The IS-HO function will not be given, either. In this way, the multi-radio access function of current multi-mode GPRS/UMTS PCMCIA cards is modelled.

4.2.4 Service-based

The IST EVEREST project has proposed using policy-based [99] initial RAT selection algorithms for heterogeneous networks [29]. A common radio resource management function selects an initial RAT to allocate a call request based on a set of input parameters and a set of predefined rules. The service type as an input parameter is especially studied in [29]. The initial RAT selection based on the type of service is called service-based policy. The RAT selection algorithms in [30], [31] are essentially also service-based.

Given a heterogeneous environment of GSM and UMTS, there are two possible service-based policies. One is to allocate the voice service to GERAN and the data service to UTRAN (VGDU); the other is to allocate the voice service to UTRAN, the data service to GERAN (VUDG). To avoid that a call is eventually blocked due to the lack of sufficient resources in the first choice, an alternative allocation using the other RAT is possible to utilise the common resources. So both policies enable the prioritised RAT selection guiding the overflow at call setup stage if needed. According to the simulation results of [29], VGDU turns in higher throughput than VUDG. Note that these simulations implement dedicated channels for data services but not high speed shared channels. Actually, the performance of a respective policy is heavily dependent on the underlying system capability.

4.2.5 Load Information Aiding MRRM (LIAM)

We propose an MRRM decision algorithm concerning both Initial RAT selection and IS-HO.

Since CS users are unwilling to make IS-HO because of the combined undesirable delay, imperative IS-HO is more realistic for CS services than a load balancing algorithm [26][101] that potentially results in high IS-HO rate perusing similar load level in the two networks.

The triggering of initial RAT selection and IS-HO of LIAM for voice calls is same as of Blind. However the decisions of the two events are made different from Blind, i.e. the algorithms themselves are different. For voice service, LIAM selects the cell with the best C/I among the cells which have sufficient resources for a new arrival. Hence the cell load values based on the assumption of a new admittance in the target candidate cells are the necessary input information for LIAM.

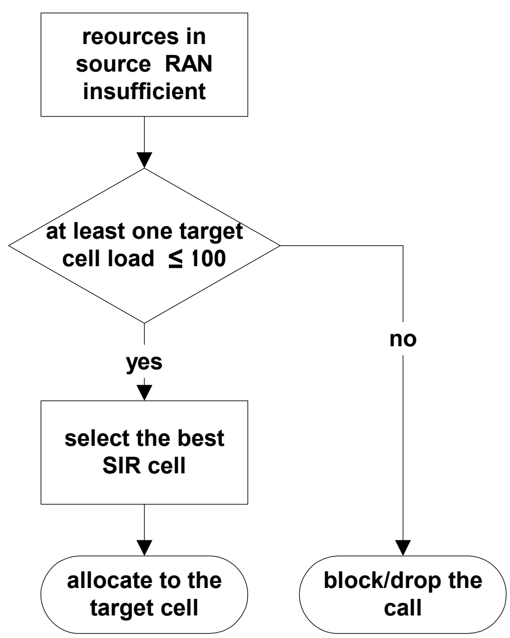

Figure 4-3: LIAM for RT services

The starting point of each packet call is triggering the RAT reselection for data services as we analysed in Section 4.2.1. In a real environment, a packet arrival in downlink is always associated with two con-current events. Event one:

there is always an uplink request before a web page download. Event two: in the downlink, the RRC state has been CELL_FACH and the RLC buffer has been empty for an inactivity timer, which should correspond to the reading time. Whether the two events are sufficient and how to use the events to determine the exact arrival time, need further study.

LIAM firstly evaluates the NRT load information in the two cells with the best C/I in the two networks; then allocates each packet call to the cell with the lower NRT load information.

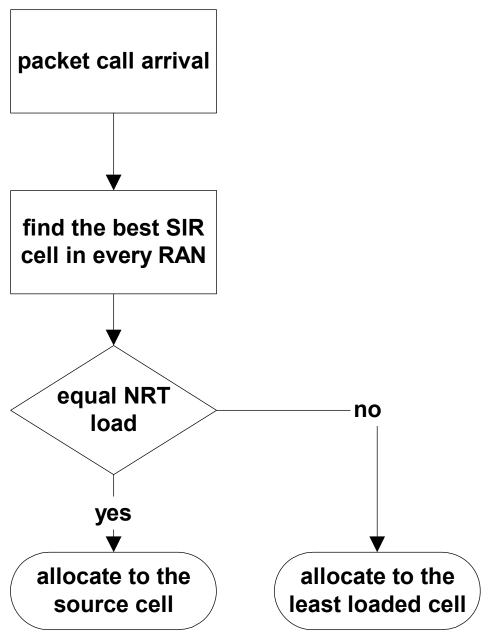

Figure 4-4: LIAM for NRT services

Some formulae map the physical cell load status into the load information. The later one is reported to the MRRM entity. This mapping is called load mapping. The load mapping includes a cell load mapping and an NRT load mapping.

The absolute priority of the voice service over the data service requires LIAM to take merely CS load relevant cell load into account with regard to the resource allocation for a voice call. Thus, LIAM also ensures the CS priority at the MRRM level. The cell load of a GSM cell is equal to the time slots occupied by CS users including the new arrival, TS_{CS}, divided by the maximum number of time slots available in a cell for user data transmission TS_{max}:

$$Cell\ Load_{GSM} = \frac{TS_{CS}}{TS_{max}} \tag{4.1}$$

The cell load of a UMTS cell in the downlink is equal to the power assigned for the CS users including the new arrival, P_{CS}, divided by the maximum BS transmission power P_{max}:

$$Cell\ Load_{UMTS} = \frac{P_{CS}}{P_{max}} \tag{4.2}$$

NRT load information is independent on the RAT types. The average data rate is mapped into four NRT load levels labelled as low, medium, high, and overload. Figure 4-5 illustrates the mapping scheme. The mapping curve can be a linear, a convex or a concave curve. A concave curve, which gives more dense resolution in the high data rate range, is suitable for low load situation. A convex curve is appropriate for high load situation.

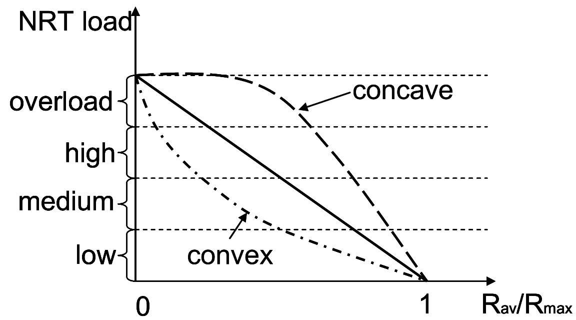

Figure 4-5: NRT load mapping as a function of the average user bit rate R_{av} is normalised to the maximum user bit rate R_{max}

From the implementation point of view, the NRT load mapping is realised with the aid of a dynamic updating of three thresholds. The MRRM entity tracks the history of the average data rates of each completed packet calls. Three thresholds are updated every 10000 packet calls, where the thresholds are defined to be the average data rates corresponding to three selective cumulative distribution probabilities, e.g. 25%, 50% and 75%.

$$R_{low} = CDF^{-1}(25\%)$$

$$R_{medium} = CDF^{-1}(50\%)$$

$$R_{high} = CDF^{-1}(75\%)$$

$$\begin{aligned}
&NRT\ Load\ Information = low, \forall\ R < R_{low} \\
&NRT\ Load\ Information = medium, \forall\ R_{medium} > R \geq R_{low} \\
&NRT\ Load\ Information = high, \forall\ R_{high} > R \geq R_{medium} \\
&NRT\ Load\ Information = overload, \forall\ R \geq R_{high}
\end{aligned} \tag{4.3}$$

The formats of the load information above are completely aligned with the ones given in 3GPP TS48.008 [11].

In summary, overflow and IS-HO are the two main schemes that enable the resource sharing of various RATs. Regarding the MRRM evaluation, a service-based strategy can benefit from the diversity of resource efficiency in various networks, while LIAM aims to balance the load as well as to improve the system performance. LIAM involves the most schemes of MRRM, which can be clearly learned from the overview given in Table 4-1.

	Blind_DU	*Service-based*	*LIAM*
CS overflow at call setup	Yes	Yes	Yes
PS overflow at call setup	No	Yes	Yes
CS IS-HO	Yes	No	Yes
PS IS-HO	No	No	Yes
Input parameters			
Service type	Yes	Yes	Yes
Channel quality	Yes	Yes	Yes
Cell load information	No	No	Yes

Table 4-1: Overview of MRRM algorithms

4.3 Simulation Model

A dynamic event-driven simulator [45]-[46], [100]-[101] has been implemented within the OPNET [102] simulation environment using C++. Figure 4-6 illustrates the simulator structure and the functionality entities. The simulator comprises RAT independent and dependent modules. The RAT independent modules include the traffic source and the MRRM entity. The local radio resource management (RRM) entities, the base stations (BSs), the multi-mode mobile stations (MSs), and the path loss models are the RAT dependent modules. The information and commands flows between the entities are given in Figure 4-7.

4.3.1 Radio Access Technology Independent Modules

Two types of traffic are generated: voice and web-browsing. Both services are generated by Poisson processes. The duration of a voice call is exponentially distributed with a mean of 120s. The duration of a packet session of web-browsing depends on the real transmission rate, the data amount generated and the reading time between packet calls. A packet session is illustrated in Figure 4-8 following the traffic model proposed in [96] (3GPP TR 101 112). As can be see from Figure 4-8, a packet session consists of time periods where data is transmitted, each being defined here as a packet call, and time periods without any transmission, each being defined here as reading time. As detailed in [96], a packet call is corresponding to a web page download containing several datagrams. So IP layer packet is modelled. The number of packet calls, the number of datagrams inside a packet call and the reading time are generated from geometric distributions with a mean of 5, 25 and 12s, respectively. The datagram size is modelled by a Pareto distribution with the location of 81.5bytes and the shape of 1.1.

The MRRM entity is where the MRRM strategies operate. It maintains interfaces with the local RRM entities. The MRRM entity gathers the necessary knowledge concerning the local networks from the local RRM entities and commands the local RRM entities about decisions of resource allocation.

As the downlink is the potential bottleneck of multi-services communication, only the downlink communication is considered in this thesis. The traffic source generates CS/PS calls and delivers them to the MRRM module. Based on the decision of the corresponding MRRM algorithm, a call is further delivered to the appropriate access network.

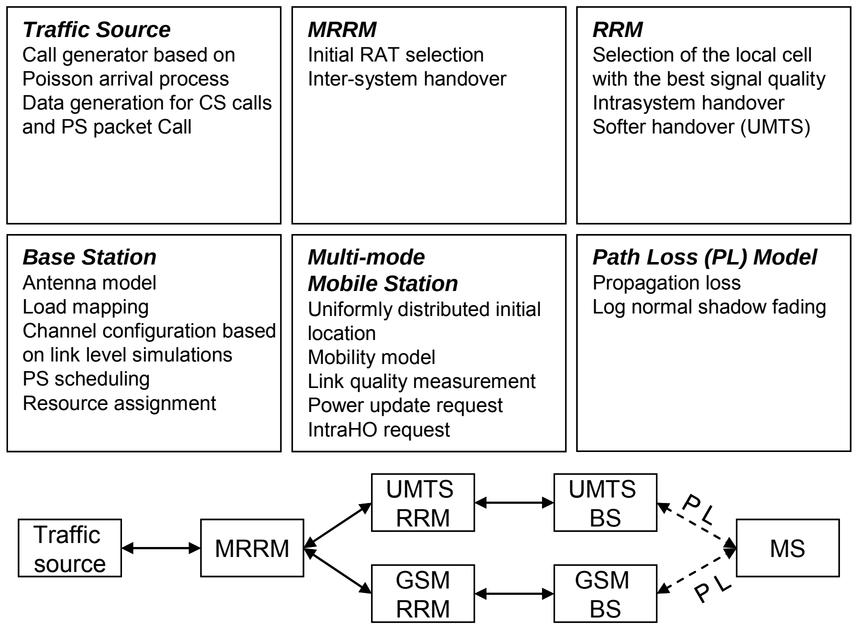

Figure 4-6: Functionality and architecture of the simulation

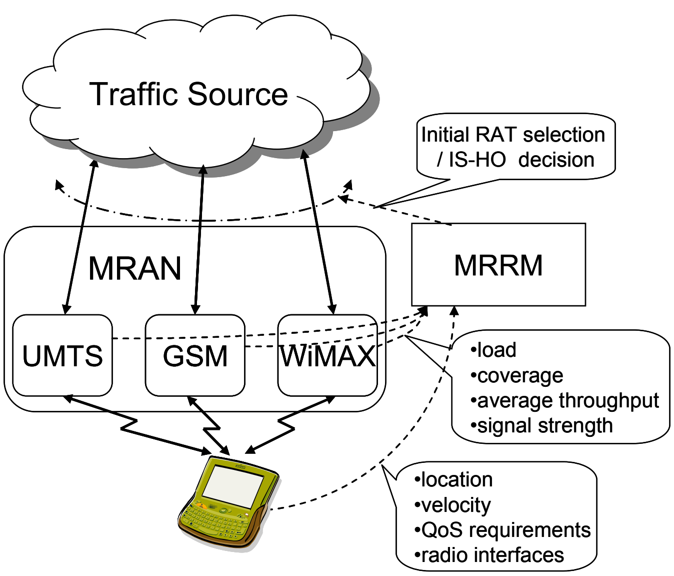

Figure 4-7: Information and commands flows between the entities

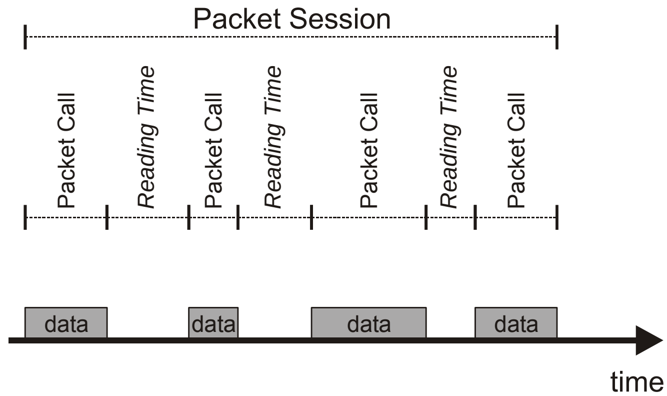

Figure 4-8: Illustration of a PS session

4.3.2 Radio Access Technology Dependent Entities

Once a call is delivered to GSM/UMTS by the MRRM entity, the RRM will allocate the call to the most suitable cell in accordance with the MRRM command. The BS assigns the required resources upon a CS call arrival or the available resources upon a PS call arrival.

PS Scheduling

It is assumed that EDGE and HSDPA are deployed in GERAN and UTRAN in order to provide high data rate for data services. HSDPA adopts the round robin scheduling for PS connections whereas EDGE adopts a scheme of fair time slots sharing as presented in Section 2.3.2 and Section 2.3.1. An HSDPA connection may be assigned 15, 10 or 5 multiple SF16 codes adaptive to the traffic conditions. In conjunction with the assignable spreading codes, modulation and channel coding varieties adapted to the channel conditions determine the data throughput of a user. The user throughput ranges from 0 to 7088kbps. In order to let the EDGE PS data rate to be comparable with the HSDPA data rate, a multimode MS is assumed supporting up to 8 time slots adaptively using MCS-1 to MCS-8, which gives maximum user data rate up to 384kbps. Upon a user request, GSM BS scans every continuous maximum supported time slots and allocates the user into the group of timeslots where it can achieve the highest data rate. Data from up to 6 users may be interleaved inside a time slot.

HSDPA SIR

A HSDPA user will be exclusively served if it is scheduled. Hence the active connection can utilized up to the maximum TX power for HSDPA, which is equal to the maximum BS TX power subtracting the TX power for common control channels. Different spreading codes cause intra-cell interference in between although they may be assigned to a same user. For simplicity and not losing the generality, we take the total TX power as the intra-cell interference source. The downlink SIR of an active HSDPA user *i* is calculated according to Equation (4.4) which is different from the SIR of a UMTS CS user given by Equation (3.21).

$$SIR_i = \frac{P_{k,i} / L_{k,i}}{\alpha \cdot P_k / L_{k,i} + \sum_{j \neq k} P_j / L_{j,i} + N_0} \tag{4.4}$$

P_k total TX power [W] of BS k

$P_{k,i}$ TX power [W] for user i in BS k

N_0 broad band thermal noise [W]

α non-orthogonal factor

$L_{k,i}$ path loss of user i in cell k including shadow fading and antenna gain

Call Admission Control

With respect to a voice call, one time slot is allocated in GERAN. One dedicated channel with one spreading factor 128 code is allocated in UTRAN. The transmission power should guarantee the signal interference ratio (SIR) required by a 12.2kbps constant rate. Due to the different technologies adopted for data and voice services, the system capacity is service mix dependent.

CS service has absolute priority of PS service. Hence the CS calls may share up to the entire resources allocated for a cell, while the PS calls can only share the remaining resources in a cell after fulfilling the QoS requirement of the concurrent CS calls. Once a BS detects there are insufficient resources to accommodate a CS call request, it will first try internally to degrade the resource sharing (therefore the bandwidth) of active PS calls hosted. If the resources saved by the internal PS bandwidth degradation are still insufficient, an urgent intrasystem hard handover will proceed.

A fractional bandwidth is reserved for handover CS calls in GSM and UMTS, respectively.

Sectorised Antenna

The popular three-sector antenna is implemented at GSM and UMTS BSs. The antenna pattern (Appendix B in [96]) is re-given in Figure 4-9. The direction-dependent antenna gain is then added to the propagation loss (Section 3.4.3) and a random shadow fading. This type of pathloss consideration is also applied to GSM. Rayleigh fading was not included since it is considered to be averaged out.

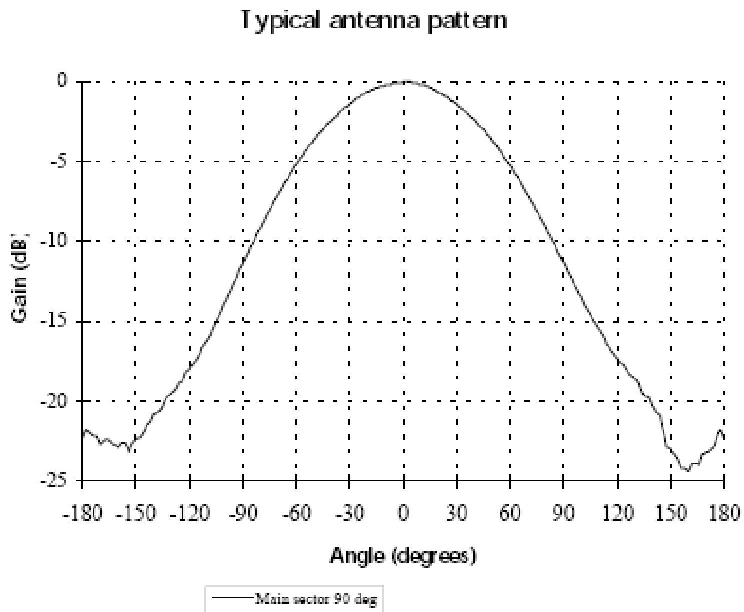

Figure 4-9: Typical horizontal antenna pattern [TR 101 112]

Mobility Model

A MS moves around only when there is a call associated with it. Once a MS becomes active, i.e. once a new call arrives, a uniformly distributed random initial position is assigned. From this initial position on, the mobile starts moving with a random speed and a random direction, where a slowly-moving user alters his moving direction more often than a fast-moving user. A pedestrian mobility type models the behaviour of a slow moving user who potentially changes the direction often and without preference. The speed of a pedestrian user is set to 3 km/h constantly. A vehicular mobility type models the behaviour of a fast moving user who may have small changes to the current direction but keeps to an overall main direction. The speed of a vehicular user is set to 120 km/h constantly.

Power Control

Power update is operated at ± 1dB for a UMTS MS to maintain the SIR against the varying radio condition. The radio condition may vary by the several possible reasons: First, the path loss variation due to the mobility described above; Second, the intra-cell interference variation due to an acceptance of a new call, a termination of an ongoing call, or power update for other calls in the same cell; Third, the inter-cell interference variation due to power update in neighbouring cells. Among these variations of radio condition, the propagation loss variation is being continuously checked.

In practice, the path loss is normally measured by physical device, e.g. by mobile stations for the downlink. Concerning to the characteristics of shadow fading in the radio environment, a decorrelation distance of the values of shadow fading at two separate positions exists ([92]). In the computer simulation, however, it is calculated in the path loss model and is reported to the corresponding mobile station at stepwise when the mobile has moved a straight distance of a decorrelation distance from the position of the last report. This stepwise pathloss measurement is followed by a SIR update which may trigger a power update event.

Intrasystem Handover

Intrasystem/soft handovers belongs to the type of mobile-assisted HO. MSs report the measurement of signal strength of various BSs. The RRM command the corresponding BSs and the MS to carry out the (soft) handover. The soft handover follows the algorithm in Section 2.4.2, where the configurations of hysteresis and time to trigger are listed in Table 4-2. Note that only hard handovers are possible for a GSM call or a UMTS PS call. A handover retry timer confines the handover rate as well as the handover retry interval.

Datarate Determination

A link level simulation, modelling the physical channel coding, modulation, link adaptation and retransmission, provides tabular information of channel capacity in a format of SIR and data rate pairs for the GSM voice, the UMTS voice, the EDGE data, and the HSDPA data, respectively. For the detailed channel capacity of each system, please refer to Appendix B. With the aid of the pathloss model, a channel SIR is calculated, and consequently the available data rate is mapped. Inversely, a necessary TX power can be derived to match a target SIR value

Due to the resource sharing among active PS users, the effective data rate of a user must be further calculated from the mapped data rate for which the exclusive channel usage is assumed. For an HSDPA user who is concurrent with n-1 other active users, the effective data rate is consequently 1/n of the mapped data rate. For an EDGE user, the channel capacity of a time slot must be shared by the users whose data are interleaved inside the common time slot.

4.4 Capacity Region

The system capacity of a certain service type can be measured as the total throughput under a QoS criterion. One can understand this QoS criterion as a minimum level a type of service satisfying. Multiservices usually have different quality requirements. Conventionally, the system capacity of voice service is confined to the carried traffic with a maximum blocking rate, while the system capacity of data service is confined to the throughput with a minimum data rate. We set the voice blocking rate equal to 1% as the criterion for the voice capacity. However, the simulator is designed to admit unlimited data calls without a minimum data rate constraint, for the sake of presentation for best effort data services. We consider the capacity of data service to be violated when the average throughput per user in a cell is less than 10kbps during 10min. The capacity limit of the heterogeneous networks is defined by a matching of any of the criteria in any underlying networks.

The PS throughput versus CS users is presented to stand for the capacity limit. The curve itself is called capacity limit. The area below the capacity limit is called capacity region. Load inside the capacity region is anticipated to result in better QoS than at the capacity limit.

4.4.1 Simulation Configuration

The simulation tested a heterogeneous network with 42 pairs of collocated GSM and UMTS cell cites, illustrated in Figure 4-10. The system configuration given in Table 3-1 is still valid for the integrated services scenario except two parameters are cancelled: GSM antenna gain for pure stationary mobiles, and GSM antenna gain for pure pedestrian mobiles. The maximum user data rate for a NRT connection is not limited by user terminals. That means NRT service users will be assign as much as possible bandwidth. The complementary system configuration on top of Table 3-1 is given in Table 4-2. System performance in the white areas of Figure 4-10 is investigated for avoiding the boundary effect.

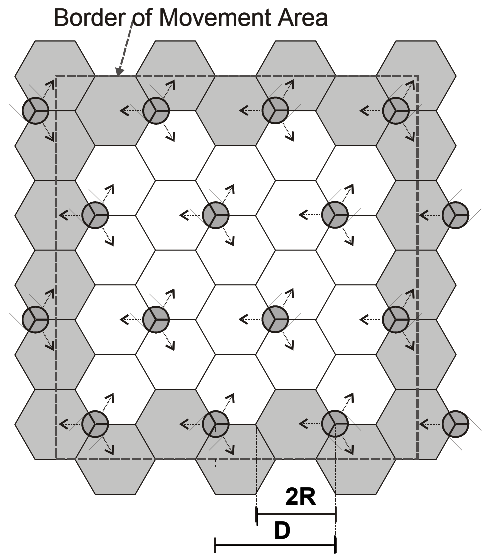

Figure 4-10: System scenario with collocated GSM and UMTS base stations

The resource planning confines the carried traffic to the network capacity. Therefore, it deserves careful consideration in the simulator. The resource configuration of GSM and UMTS cells are derived from the GSM/UMTS licenses [103] owned by German digital cellular network operators, such as T-Mobile Deutschland GmbH (D1) and Vodafone D2 GmbH. Both have UMTS licenses consisting of a pair of 10 MHz FDD frequency bands in the ranges 1920-1980 MHz and 2110-2170 MHz, and one 5 MHz TDD frequency band in the range 1900-1920 MHz. Concerning the GSM licenses for them, both have a pair of 12.4 MHz frequency bands in the range 890-915/935-960 MHz (D-GSM-frequency), and a pair of 5 MHz frequency bands in the range 880-890/925-935 MHz (E-GSM-frequency). The resource planning of the simulated networks are based on the following assumptions:

- Homogeneous operator for the UMTS network or for the GSM network, i.e. no multi-operator for a network with a common RAT
- FDD for the UMTS network

- A single carrier for the UMTS network, i.e. a half of the licensed band
- D-GSM-frequency band for the GSM network, also a half of the licensed band
- Frequency re-use factor is 9 for the GSM network

Thus the UMTS network is configured with a single carrier (5 MHz bandwidth). The counterpart of resource of the GSM network is with 6.2 MHz bandwidth, and equivalent to 3 frequency carriers in a cell under frequency re-use factor 9. Assuming that the control channels and the broadcast channels in a GSM cell occupy 3 timeslots in average, a GSM cell has 21 timeslots available for data transfer.

UTRAN carrier	1
UTRAN max. BS TX power	20W
UTRAN max. HSDPA TX power	19W
UTRAN thermal noise	-100dBm
GERAN total time slots for data	21 (3TRX)
GERAN link TX power	1W
GERAN interference	-105dBm
GERAN / UTRAN cell radius R	800m
GERAN / UTRAN BS separation D	2400m
Number of GSM/ UTRAN cells	42
Number of investigated GERAN / UTRAN cells	20
Soft handover addition hysteresis	2.5dB
Soft handover removal hysteresis	4.5dB
Intra handover replacement hysteresis	2.0dB
Handover hysteresis triggering timer	0.25s
Handover retry timer	0.5s
Pedestrian MS	87.5%
Vehicular MS	12.5%

Table 4-2: System Configuration

4.4.2 Capacity Region of Separate Systems

The two dashed curves in Figure 4-11 indicate that the two systems have much high PS capacity in terms of cell throughput than CS capacity. In addition, the capacity limits in both networks are nonlinear. This nonlinearity is basically resulted from the adaptive EDGE and HSDPA technology. Based on our system configuration, the CS capacity in a UTRAN cell is more than triple of the CS capacity in a GERAN cell; the GERAN resource is more efficient for serving data service than voice service. The reason for the very high voice capacity of a UTRAN cell is because the 800m cell radius is relatively small with respect to the 20W transmission power of Node B and the 17dBi antenna gain at the bore sight.

4.4.3 Capacity Region of Overall Systems with MRRM Enabled

The solid curves in Figure 4-11 confine the capacity region of various MRRM strategies.

One of the service-based strategies, namely VGDU, obviously has the smallest capacity region. The other service-based one, VUDG, provides the maximum capacity limit since the resource efficiency of the two networks is utilised to allocate the traffic. So from the perspective of system throughput under a certain number of voice users, VUDG is better than VGDU, which is opposite from [29]. Note that the system throughput obtained in [29] is from fixed numbers of voice and www users. Therefore the larger system throughput represents in fact the larger user throughput. But the system throughput under a fixed user number is not synonymous with the system capacity defined by us. From the performance of data users' point of view, VGDU is better than VUDG in [29]. Nevertheless from the system capacity's point of view, it is not clear which is better based on the results presented there.

The proposed algorithm, LIAM, outperforms Blind_DU and VGDU in the whole service mix range. LIAM outperforms VUDG in the case of a service mix of quite high voice traffic component (more than 72 voice users) and in the case of pure voice traffic.

No algorithm shows gain for the capacity of pure data service because the elastic bandwidth of web browsing service can fully utilise radio resources [30]. As long as the packet calls are allowed to share the common resource pool, the overall PS capacity limit is determined by the sum of the PS capacity limits of the two systems. Nevertheless, the user data rate with different underlying MRRM strategies may vary in size.

MRRM algorithms enabling IS-HO achieve the improvement of voice capacity. The sum of the voice capacity limits of the two systems gives the voice capacity with the service-based strategies. LIAM contributes 8% voice capacity gain compared to the capacity sum. Blind_DU that also enables IS-HO achieves the voice capacity similar to LIAM. This gain of voice capacity arises from the trunking gain of the common resource pool which is enhanced by IS-HOs and the various traffic filling schemes. By default, the traffic is uniformly generated in the two systems with same arrival rates. When using LIAM or Blind_DU, the traffic is filling into the two systems with equal probability; overflow and IS-HO is possible whenever necessary. When using the service-based strategies, the traffic is firstly filling up one system then redirected to the other. The total capacity is equal to the sum of the two systems and thus the trunking gain of voice capacity is eliminated when using the service-based strategies.

The similarity of the voice capacity gain with Blind_DU and LIAM can be explained from the perspectives of CS MRRM evaluation. General speaking, because the voice capacity of GERAN is less than 1/3 of UTRAN, IS-HOs often occur by transferring voice load from GERAN to UTRAN. In UTRAN, a voice channel in the cell with the best C/I consumes most likely the lowest transmission power. So despite the knowledge of load situation in the target system, the advantage of LIAM over Blind_DU for the CS service is trivial if the UTRAN has much more CS capacity than GERAN. It might be significant if the two networks have similar capacity and the load level is therefore a crucial issue for IS-HO.

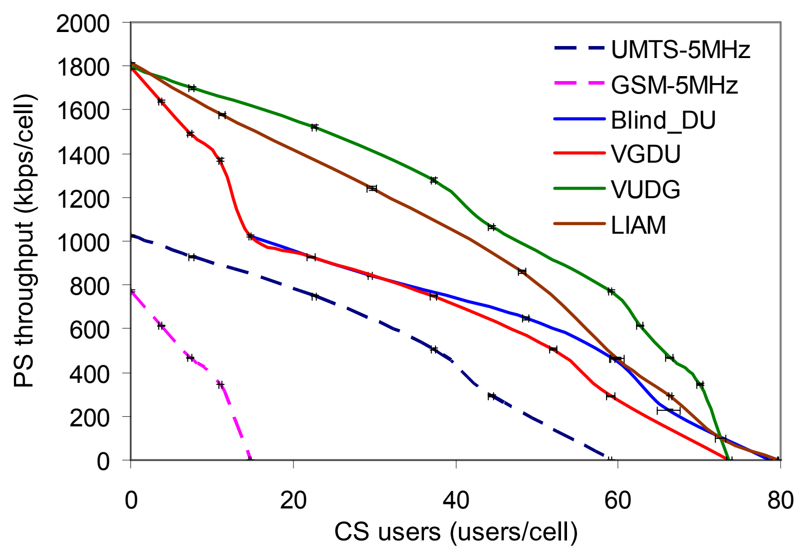

Figure 4-11: Capacity region of separate systems, Blind_DU, service-based and LIAM, where 99% confidence intervals are shown for both variables

4.5 Performance Evaluation of VUDG and LIAM

Although VUDG achieves larger capacity than LIAM in most of the service mix range, it is also worth having an insight into the system performance in terms of PS data rate, the CS blocking and CS dropping with varying service mix. For the sake of simplicity and not losing generality, we investigate the PS performance under varying input PS traffic plus a fixed amount of coexisting voice traffic; additionally, we investigate the voice blocking and dropping under varying input voice traffic plus a fixed amount of coexisting PS traffic.

The service mix in Figure 4-12 is below the capacity limit. From Figure 4-12 we have seen that the PS data rates with VUDG are limited below 300kbps with respect to the given range of input PS traffic. The limitation mainly comes from the maximum bit rate available for one user in EDGE, 384kbps in contrast to 7088kbps in HSDPA. Recall that GERAN is prioritised to accommodate the PS traffic with VUDG, while the PS traffic is allocated to the system which can assign higher bandwidth with LIAM. Moreover, VUDG operates per packet session basis while LIAM operates per packet call basis. The resulting gain of PS data rate with LIAM sharply increases with the decreasing input traffic.

Additionally, Figure 4-12 indicates why LIAM results in suboptimal capacity. The data traffic is allocated mainly to UMTS/HSDPA to fulfil the design objective of LIAM, namely optimized user data rate. This allocation is against the resource efficiency principle which VUDG follows. The capacity with LIAM is therefore suboptimal in comparison to VUDG, see Figure 4-11.

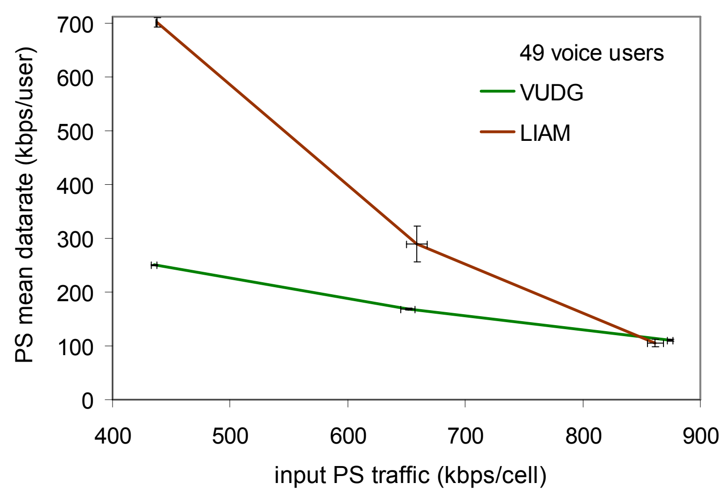

Figure 4-12: Comparison of PS performance with VUDG and LIAM, where 99% confidence intervals are shown for both variables

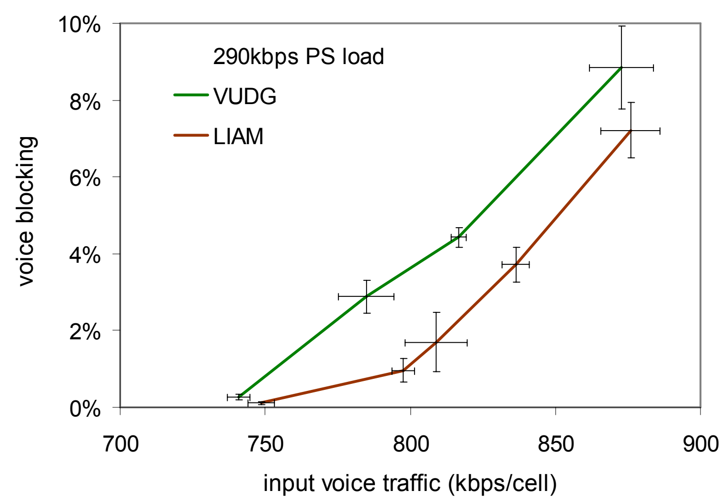

Figure 4-13: Comparison of voice blocking with VUDG and LIAM, where 99% confidence intervals are shown for both variables

As for the voice performance, the service mix in Figure 4-13 is below and also beyond the capacity limit where the blocking is confined to not greater than 1%. VUDG also worsens the voice blocking, see Figure 4-13. Under the input traffic giving 1% blocking with LIAM, the blocking with VUDG is enlarged to 3.5%. The capacity gain with the mixed traffic is anyway less than the 8% capacity gain with pure voice traffic. The outstanding blocking with LIAM is due to the various traffic filling schemes we mentioned in Section 4.4.3. Moreover, the additional IS-HOs enabled enlarge the capacity as well. Note that the voice dropping with both algorithms is maintaining 0% in correspondence with the investigated traffic range.

4.6 Summary

Several multi-standard radio resource management (MRRM) strategies for integrated voice and data services are implemented to heterogeneous GSM/UMTS networks via computer simulations. The networks adopt EDGE and HSDPA technologies to provide high bandwidth for data services. The capacity limits of the GSM and UMTS networks are consequently nonlinear.

The capacity limit varies when different MRRM strategies are implemented. The currently commercial available MRRM functionalities for the voice service and data services have quite low competence compared to the service-based strategy and the load information aiding MRRM (LIAM). A service-based MRRM strategy that allocates a type of service to the most resource efficient network (in our case VUDG) achieves already the maximum multiservices capacity. However, the service-based MRRM does not optimise the performance of each specific service. In contrast, LIAM improves the performance of each type of service via IS-HOs based on load information. This performance gain is the trade-off of the capacity maximisation. However, the function of intersystem handovers enlarges the system capacity of RT services, e.g. 8% in the case of pure voice traffic.

In practice, LIAM is suitable for heterogeneous networks with 1) pure RT services, 2) integrated services with RT traffic more than the sum of the RT capacity of each underlying network, 3) integrated services where the total load is inside the capacity range, 4) pure NRT services. The service-based MRRM strategy that allocates a type of service to the most resource efficient network (in our case VUDG) is suitable for integrated services for which total load is close to the capacity limit.

Chapter 5

Distributed MRRM – An Approach for MRRM Implementation

Algorithm based multi-radio access normally needs the cell level load information of the subsystems, or even link level quality information of connections [25] - [32]. This requires the necessity of signalling to transfer the load information from the cells to the MRRM controller. The way to collect the load information through signalling varies both accuracy and freshness of the information. The signalling can be periodic, or can be reactive whenever acquired, or can be triggered at every change of the local load information. Due to the cost of signalling overhead, the signalling should be used carefully. Moreover, the signalling procedure should not take too much time to reduce the handover process delay. On the other hand, out-of-date load information introduces a risk of bad or even fatal decisions of intersystem handover. So far, the papers investigating MRRM algorithms usually do not consider the signalling effect on system performances.

In this thesis, an approach of distributed MRRM entities in UTRAN and GERAN is presented, where the intersystem information exchange is carried out according to the 3GPP signalling capabilities. The distributed MRRM entities then base their RAT selection decisions on their system knowledge, for which the amount, age and accuracy of the available intersystem information are determined by the signalling capabilities. The influence of these signalling effects on the system performance of MRRM managed heterogeneous networks is assessed through simulation studies. This proposal and the performance evaluation are also published in one of our paper [104].

5.1 Load Information Exchange Mechanisms in 3GPP R5

Intersystem handover (IS-HO) signalling capability between GERAN and UTRAN was standardised in 3GPP Release 5 [10] - [11]. Cell load information can optionally be attached to the signalling protocols during the resource reservation phase of the IS-HO procedure, see Figure 5-1 sketching the message flows. The cell load information element conveyed by these signalling messages contains the following 4 cell load values [11]:

- Cell Capacity Class Value (0 ...100)

* (Total) Load Value (0 ... 100 %)
* RT Load Value (0 ...100 %)
* NRT Load Information Value (0 (low), 1 (medium), 2 (high), 3 (overload))

The cell load information can thus be exchanged between two cells involved in a successful or failed IS-HO. The precision of the load information is limited by the allowed load values above.

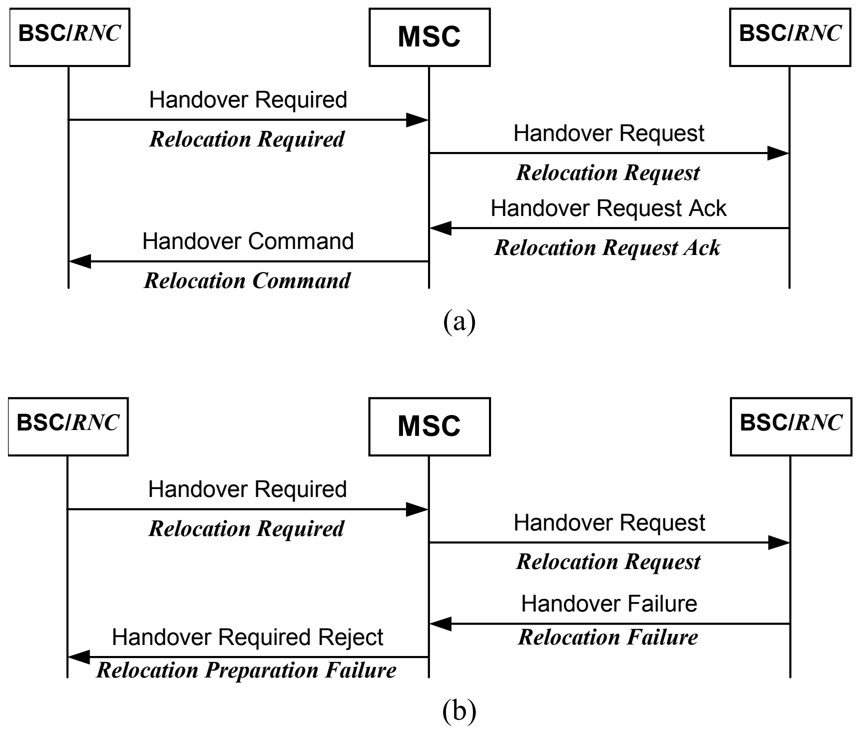

Figure 5-1**:** Signalling flow charts illustrating the load information exchange during a handover/relocation procedure [24], (a) successful IS-HO, (b) failed IS-HO

IS-HO events are specifically relevant for circuit switched (CS) services. For packet switched (PS) services, we consider that the mobile stations are in RRC connected mode. In this status, the network can initiate an intersystem cell change order (IS-CCO). However, no seamless IS-CCO to GERAN was defined in Release 5 and therefore no load information can be included in the IS-CCO message. Thus all the knowledge about the NRT load state in the cell of the other RAT can only be obtained from previous IS-HOs of CS connections or by the method of dummy-handover, explained in the next section.

5.2 Distributed MRRM Based on Realistic Signalling

5.2.1 Load Information Updated by IS-HO

The 3GPP intersystem signalling capabilities are modelled in the present work by a distributed MRRM architecture, one MRRM entity in each RAT, physically located in the RNC/BSC. The MRRM entities evaluate and decide themselves, whether to accept or keep a certain connection, or whether to direct it to the other RAT. Attached to each MRRM entity is a cell load database, which stores the known information about the cell load status in the other RAT. These database entries are updated for both involved cells whenever an IS-HO takes place, and are then stored for future use.

Whereas the MRRM entity always has an accurate and up to date load information about the local cells, it has to use those stored load values from the database for cell load information about the other system.

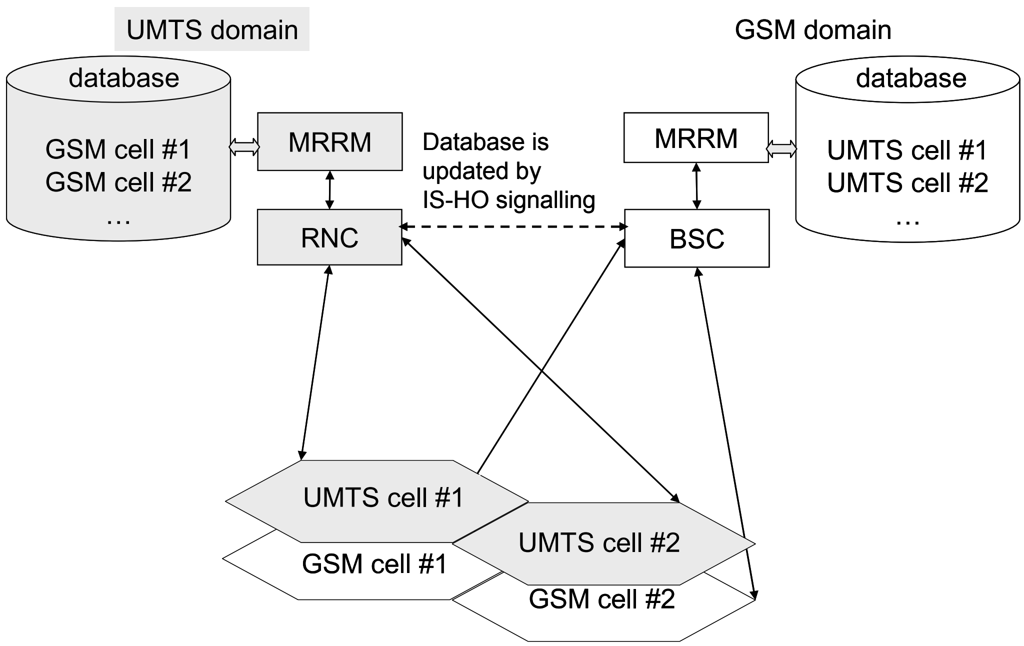

Figure 5-2: Load information maintenance of target cells in alternative radio access networks

In the event of an RT call which cannot or can no longer be hosted in the source system, the load information aiding MRRM algorithm selects a target cell in the alternative RAT. Based on the stored cell load values, firstly a list of suitable candidate target cells with sufficient free resources is created. Inaccurate cell load information may cause one of the following: Either a well suitable target cell is not considered, or an already fully loaded target cell is wrongly assumed to still be able to host this call. The candidate target cell which has the best C/I signal strength, is finally chosen.

For NRT traffic, the load information aiding MRRM algorithm evaluates at the beginning of each packet call the possibility to perform an IS-CCO. In a first step, the both candidate cells in the source and target RAT, which have the best C/I signal strength, are selected. Thereafter, the MRRM algorithm compares the NRT load values and chooses the less loaded RAT. As outlined in [100], NRT traffic is always accepted in the system without any NRT blocking or dropping.

5.2.2 Dedicated Signalling – Dummy Handover

A mechanism with an erroneous IS-HO is proposed in [24], with the sole purpose to exchange the source/target cell load information. This erroneous IS-HO, resulting in an IS-HO failure, is labelled as 'Dummy Handover'. Thus, the mechanism of dummy handover provides cell load information exchange at the expense of increased signalling load. Here we describe how to efficiently trigger dummy handovers to improve the MRRM performance based on the realistic signalling.

For efficient usage of dummy handover, we propose to introduce several triggers to tune the dummy handover overhead. The distributed MRRM entity of each RAT keeps track of the load information reported. Once it detects that a current load value in one of its cells differs by more than a configurable amount (Δtotal-load or ΔNRT-load) from the previously transmitted one, it triggers a dummy handover to inform the other RAT about the changed load status of a cell. The closest cell in the other RAT is selected to be the target cell.

In the present study, the following four configurable triggers are employed:

- The current total load is larger than or equal to a threshold (InitialLoadThreshold) and simultaneously, any one of the following three conditions is fulfilled.
- The time span from last report (Load information age) is larger than or equal to a timer (DummyHandoverTimer).
- The current total load value differs by more than or equal to 'Δtotal-load' from the most recently reported value.
- The current NRT load value differs by more than or by equal to 'ΔNRT-load' from the most recently reported value

These dummy handover triggers ensure that any entry of the cell load database neither exceeds a certain time span limited by DummyHandoverTimer, nor differs by more than a configurable amount from the real load values. The database entries of both cells involved are updated whenever a dummy handover takes place. Most practically, the protocol of dummy handover is essentially same as of intersystem handover. So the additional querying for load information via dummy handovers does not require new signalling protocols at all.

5.3 Simulation Results

5.3.1 Simulation Scenarios

The distributed MRRM is implemented on top of the simulation model introduced in Section 4.3. We investigate here a unique MRRM strategy – load information aiding MRRM - implemented in both sub systems although different MRRM strategies are also possible due to the distributed architecture. A GERAN cell has less than 1/3 voice capacity than a UMTS cell according to the 5MHz bandwidth configuration used in the last chapter. In order to let a GSM cell has comparable voice capacity as a UMTS cell. A GSM cell is configured with 60 time slots, whereas the resource configuration of a UMTS cell remains the maximum transmission power 43dBm (20W) and a single carrier. Moreover, users move completely randomly. That is the 87.5% pedestrian and 12.5% vehicular users distribution is cancelled. A user moves with a random velocity in a random direction. The moving velocity and direction also dynamically change during the call duration. The presented simulation results show the average values over all investigated cells.

The heterogeneous system performance is studied for scenarios with three different MRRM strategies:

1. *"Separate Systems"*: Separated systems without any MRRM and intersystem handover functionality
2. *"Realistic MRRM*": MRRM with 3GPP signalling, where intersystem cell load information is exchanged by IS-HOs protocols only
3. *"Ideal MRRM"*: MRRM with always up-to-date intersystem cell load information (this scenario characterises the theoretical achievable MRRM system performance)

5.3.2 System Performance for Pure RT Traffic

This section presents results for pure RT traffic, which is offered symmetrically to both RANs. In the signalling scenario, the intersystem cell load information is exchanged by regular IS-HOs only, without the use of any dummy handovers.

As a result, this non-up to date and potentially inaccurate cell load information may result in wrong MRRM decisions which cause unnecessarily blocked calls, and thus wasting system capacity.

As can be seen in Figure 5-3, at low load, the offered traffic can usually completely be accepted into the system. At traffic loads starting from about 1220 kbps (100 voice calls), the signalling effects then lead to an increased blocking rate compared to the theoretical case of always accurate cell load information knowledge. However, the performance of the MRRM managed system - including the realistic signalling- is superior to the case of separated systems, where blocking already occurs starting from about 1050 kbps (86 voice calls).

Upon an IS-HO request, the potentially inaccurate cell load information can deteriorate MRRM capabilities to operate according to its traffic distribution strategy. However, the MRRM entity - even with potentially non-accurate information - can prevent most calls from being dropped, because one cell maintains free resources to be able to host at least two additional handover calls. As long as the inaccuracy of the stored information does not exceed the size of these reserved free resources, a handover call can still be kept in the target system.

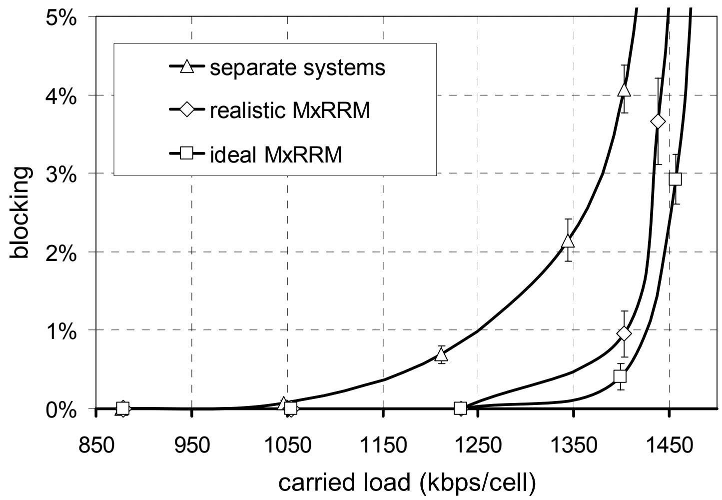

Figure 5-3: Blocking probability for pure RT traffic; shown are 99% confidence intervals

As shown in Figure 5-4, these potentially inaccurate information do therefore barely deteriorate the MRRM dropping prevention capabilities, - except for quite high traffic load starting from here at about 1400 kbps (114 voice calls) corresponding to 1% blocking, where it becomes especially important to select correctly a cell with still available capacity. The scenario with separate systems, in contrast to the other two scenarios, presents an inferior dropping performance due to the absence of the possibility to transfer any connections to the alternative RAT in situations of local congestion in one cell.

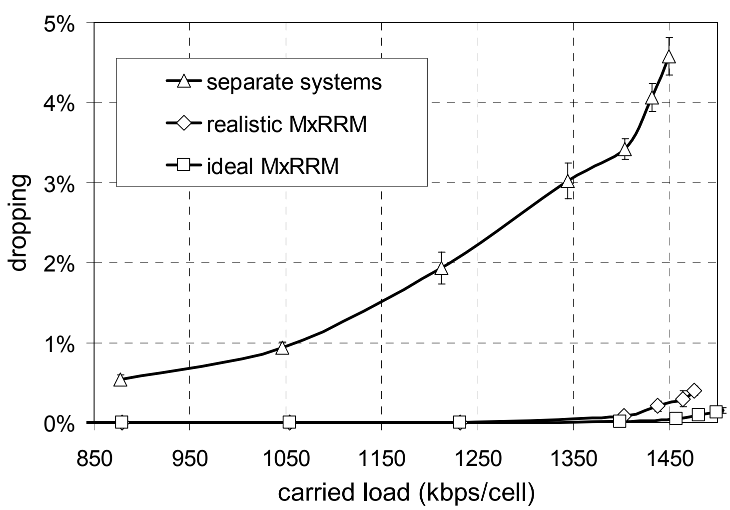

Figure 5-4: Dropping probability for pure RT traffic; shown are 99% confidence intervals

The carried traffic at a certain quality of service (QoS) level, labelled here as system capacity, is for the scenario with realistic MRRM only a few percent below the theoretical upper limit with ideal MRRM, while being considerably superior compared to separate systems. The actual numerical values can be read from the figures above and are indicated exemplarily in Table 5-1 for two different QoS levels in form of a combined dropping and blocking conditions.

	Separate Systems	***Realistic MRRM***	***Ideal MRRM***
Blocking <1% && Dropping < 0.5 %	880 kbps	1400 kbps	1430 kbps
Blocking <5% && Dropping < 1%	1060 kbps	1448 kbps	1472 kbps

Table 5-1: Cell capacity at different QoS levels

5.3.3 System Performance for Integrated Traffic

In this section, studies with integrated RT+NRT traffic are presented, for which 1240 kbps/cell prioritised RT traffic is offered in conjunction with 410 kbps/cell NRT web-browsing traffic. As presented in the previous Figure 5-3 and Figure 5-4 for symmetrical UMTS-GSM traffic, this RT traffic alone corresponds to a traffic load at which an MRRM managed system can handle the traffic completely without any blocking and dropping at all. The additional non-priority NRT traffic is sufficiently low, so that in addition to the simultaneously present RT traffic, all NRT traffic can be handled and served with quite large data rate. Results for symmetrically distributed UMTS-GSM input traffic are presented as well as the ones for asymmetrical scenarios, where the RT traffic is offered asymmetrically, 60% to UMTS and 40 % to GSM and vice-versa; the background PS traffic remains symmetrical in all scenarios.

As shown in Figure 5-5, the blocking probability is larger in those integrated RT+NRT traffic scenarios compared to the one with the same amount of pure RT traffic alone (Figure 5-3); Table 5-2 compares the values for symmetrical UMTS-GSM traffic at that given relatively low traffic load.

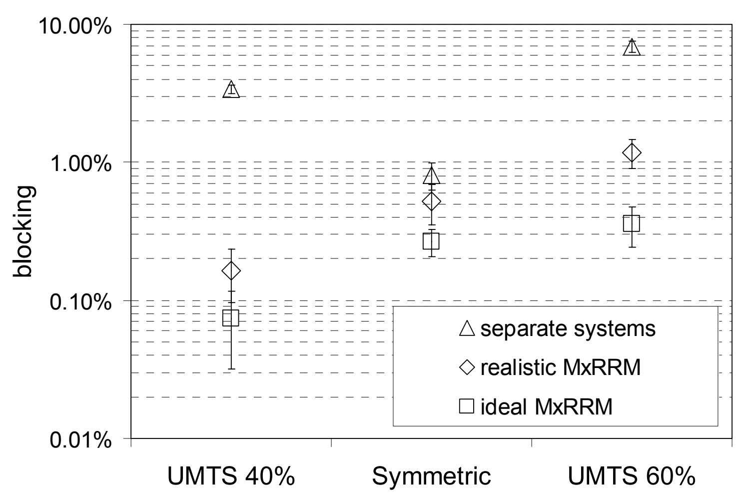

Figure 5-5: Blocking probability for integrated RT+NRT traffic. Shown are three scenarios with a symmetric or asymmetric traffic distribution between UMTS and GSM and 99% confidence intervals.

	Separate Systems	*Realistic MRRM*	*Ideal MRRM*
Pure RT traffic	0.69 ±0.12%	0.00 ±0.00 %	0.00 ± 0.00%
Integrated RT+ NRT traffic	0.81±0.18%	0.52 ±0.17%	0.27 ±0.06%

Table 5-2: Comparison of the blocking probability ± 99% confidence interval of pure RT traffic (1240 kbps/cell) with symmetrical integrated RT (1240 kbps/cell) + NRT (410 kbps/cell) traffic.

Those increased dropping and blocking rates with integrated traffic originate from two effects:

First, the bursty characteristics of the additional packet switched NRT traffic causes more frequent fluctuations of the cell load values. Thus, a previously reported cell load value tends to differs more quickly by a significant amount of inaccuracy compared to the case with pure circuit switched RT traffic which has more slowly varying cell load values. These faster fluctuating cell load values lead then to an increased effect of realistic signalling: The degradation of the system performance becomes more pronounced in those integrated traffic scenarios compared to corresponding ones with RT traffic alone.

Second, the co-existing NRT users cause some additional interference in UTRAN, which in turn then requires larger transmission power to guarantee the QoS requirements of the UMTS RT users. As a consequence, the system operates at a higher load level which then can lead more easily to blocked or dropped calls.

Both of these effects have a stronger influence on the UMTS blocking probability than on the GSM one, due to the faster fluctuation of the UMTS total cell load values which are based on the fluctuating transmission power. Consequently, the blocking probability is increased in those asymmetric traffic scenarios with more offered UMTS traffic (60%) and is reduced in the opposite traffic scenario as can be seen above in Figure 5-5.

Obviously, in the absence of any MRRM load balancing capabilities, the blocking rate increases for separate systems because one of the networks receives quite a high traffic load. This illustrates once again one conclusion in [100]-[101] that the system performance benefits especially from the MRRM capabilities in the case of asymmetric input traffic.

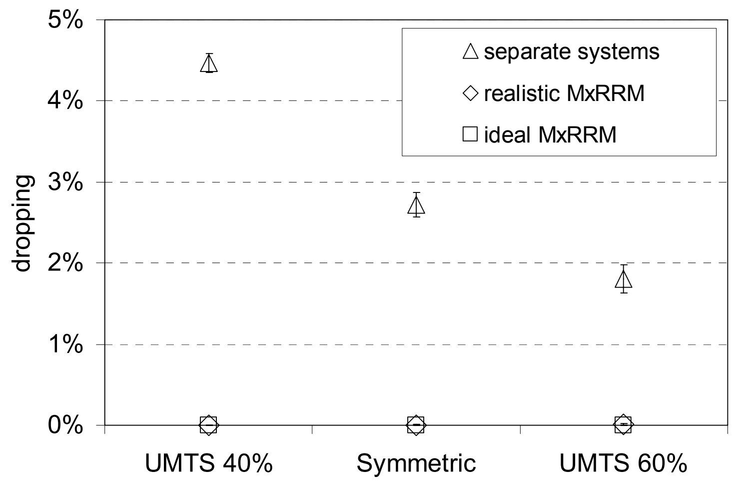

Figure 5-6: Dropping probability for integrated RT+NRT traffic. Shown are three scenarios with symmetric and asymmetric traffic distribution between UMTS and GSM and 99% confidence intervals.

Dropping is not yet present in any of these MRRM managed integrated RT+NRT traffic scenarios (Figure 5-6), because of the strong dropping preventing capability as discussed earlier in Section 0. For separated systems, some characteristics of the air interfaces become visible. The UMTS subsystem is here less sensitive to dropping, due to the UMTS soft handovers capabilities and since the amount of reserved resources for via handover incoming calls is configured in such a way, that it could be larger for UMTS than for GSM. In GSM, a cell reserves free resources for exactly two additional calls, while UMTS reserves transmission power for two calls given the worst possible interference situation, but usually the reserved UMTS power is sufficient to support more than two additional calls. This behaviour is reflected in the dropping probability of separate systems in Figure 5-6, which is larger in the case of more GSM traffic and smaller for more UMTS one.

As the total traffic load is not high, all offered NRT traffic can be well served with a typically quite high average data rate. Due to the lower priority of the NRT traffic, its behaviour is quite coupled with other effects and can also reveal even opposite performance measures. For example when the system performance gets worse for the RT traffic, represented by an increased RT blocking and RT dropping rate, then more radio resources remain for all the NRT traffic, leading to an improved average data rate for the NRT users.

5.3.4 Influence of Dummy Handover on the System Performance

This section analyses the system behaviour under the influence of additional dummy handovers to allow more frequent cell load information exchange. One of the scenarios from the previous section has been taken, integrated RT+NRT traffic with asymmetrically distributed traffic, 60% to UMTS and 40% to GSM. Additional dummy handovers are used to improve the accuracy of the load information, triggered by different thresholds as specified in the scenarios presented in Table 5-3.

	Total Blocking (%)	***Total Dropping (%)***	***Number of IS-HO / (min*cell)***
Separate Systems	6.90±0.65%	1.81±0.17%	-
Realistic MRRM	1.18±0.28%	0.01±0.01%	2.4±0.4
DHO by Δtotal-load=30	0.98±0.25%	0.00±0.00%	2.6±0.3
DHO by ΔNRT-load=3	0.85±0.35%	0.00±0.00%	2.8±0.5
DHO by Δtotal-load =5	0.69±0.15%	0.00±0.00%	2.9±0.4
Ideal MRRM	0.36±0.12%	0.00±0.00%	2.8±0.5
	Number of Dummy HOs / (min*cell)	***Average cell load information age (s)***	***Average inaccuracy of total load value***
Separate Systems	-	-	-
Realistic MRRM	-	28.0±2.7	8.7±1.1
DHO by Δtotal-load=30	9.8±0.5	6.4±0.3	6.1±0.4
DHO by ΔNRT-load=3	110.2±8.6	1.3±0.2	2.4±0.1
DHO by Δtotal-load =5	316.0±15.0	0.7±0.1	0.9±0.1
Ideal MRRM	-	-	-

Table 5-3: Influence of dummy handovers on the system performance ± 99% confidence intervals at the selected load

As can be seen in Table 5-3, the blocking probability with realistic MRRM gets improved by the use of additional dummy handover, at expense of an additional intersystem signalling load. By adapting the dummy handover trigger thresholds, the dummy handover strategy and activity can be tuned leading to a total system performance in the range between scenarios with realistic MRRM (without any dummy handovers), and the one with ideal MRRM.

It has been found that the moderate use of quite few dummy handovers can already significantly improve the system performance and can be recommended, while the signalling overhead by a large dummy handover rate is not expected to justify the further MRRM performance gain.

In order to be able to study the direct and fast system behaviour, and to obtain pure and non-levelled results, no kind of cell load averaging of the cell load values was used in the present study. As a consequence, the absolute values of the dummy handover triggers and the dummy handover activity are larger than the ones in a real system, which averages the cell load values before further processing them.

Please note that in the case of pure NRT traffic, the MRRM entity requires the use of dummy handovers, as these are the only means of intersystem information exchange.

5.4 Summary

The system performance of distributed MRRM entities was investigated where the 3GPP signalling mechanisms for intersystem load information exchanged were taken.

Depending on the QoS requirements, the capacity of the system with realistic signalling is a few percent below the theoretical upper limit of a system with always up to date and accurate global information knowledge. However, its system capacity is significantly superior compared to the case of separate systems without any MRRM capabilities. The effect of signalling on the dropping probability depends strongly on the amount of reserved resources for via handover incoming calls.

The optional use of dummy handovers improves the intersystem knowledge of the MRRM entities allowing the latter to perform more exactly its RAT-selection strategy. The dummy handover activity is highly configurable and the system can be tuned according to a desired information knowledge basis. The moderate use of some dummy handovers can be recommended, while in the case of a high dummy handover frequency its involved signalling effort cannot be justified by the further additional system performance gain.

Chapter 6

WiMAX Integrated to GSM and UMTS

We studied the RAT reselection between GSM and UMTS triggered by the arrival of a new web page based on maximum utilisation, bit rate fairness and minimum service algorithms in [100]. The maximum utilisation algorithm, which is "load balancing" indeed, was shown by simulation results being incompetent to deal with heavy load. The other two are QoS optimised algorithms and are more PS service oriented MRRM algorithms. Here we further study MRRM in a heterogeneous networks comprising GSM, UMTS and a third air interface – WiMAX. We assume that WiMAX acts the additional resource to cover the data service demand in a heavy loaded hot spot area. The Bit Rate Fairness algorithm in our previous work [100] is extended, enabling initial RAT selection as well as IS-HO for non-real-time (NRT) services among three radio access networks. Furthermore, the mapping of an NRT load level is designed to be adaptive to the NRT traffic load. Our study focuses on the MRRM effect on the load redistribution and the performance of heterogeneous networks.

The remainder of this chapter is organised as follows: the WiMAX simulation model is firstly introduced. Then the integration of the three radio access networks and the MRRM algorithms applied are explained. Finally, the capability of individual networks and the MRRM effect on the system performance are studied based on dynamic simulation results. Parts of this chapter are presented in [105].

6.1 WiMAX Simulation Model

6.1.1 WiMAX Path Loss

Equation (6.1) gives the channel propagation model of WiMAX, where the vehicular propagation loss is an extended Hata-Okumara model [106]. The frequency band is supposed to be allocated in the 3.5GHz band.

$$\begin{aligned} &r < 0.0068: \quad L = 0 \\ &r < 2.62: \quad L = L_{FS} = 43.32 + 20 \cdot lg(r) \\ &r > 2.62: \quad L = L_V = 36.96 + 35.22 \cdot lg(r) \end{aligned} \tag{6.1}$$

L propagation loss [dB]
L_V "vehicular" propagation loss [dB]
L_{FS} free space loss [dB]
r distance between BS and MS [m]

6.1.2 WiMAX Power and Subchannel Allocation

In the simulation model, a simple static power allocation and a static subchannel allocation are adopted. Although the subchannelisation is implemented, without dynamic subchannel allocation, the static subchannel allocation is effectively same as the function of TDMA. Adaptive modulation and coding effect are involved into a link level simulation which provides tabular information of channel capacity in format of SIR and data rate pairs for the system level simulation.

Multiple users share fairly the total subchannels. The total subcarriers of a WiMAX cell are divided into M subchannels, each of which is a subset of subcarriers comprising of an equal number of OFDM subcarriers. One user among the total N users in a WiMAX cell shares virtually x subchannels, x = M/N. However, one user may maximally occupy so many subchannels which can achieve up to R_{max}. The superfluous subchannels for a user who is assigned the maximum data rate should be freed for the other users. If the total number of available subchannels is not exactly divided by the number of users, the users share the remainder by TDMA. Downlink and uplink share the resource in TDD such that downlink with more traffic can share also more resources.

We assume that there is zero interference for a single WiMAX hotspot cell. Each user shares equally Pk/N TX power, where the Pk is the total transmission power of the BS. Similarly, the thermal noise for a user who is assigned one Nth channel bandwidth is also N0/N, where the N0 is the broad band thermal noise. If we assume no subcarrier diversity inside a subchannel exists, the fading of the sub carrier channel is not frequency dependent, i.e. flat fading. This is practical possible by the interleaved subchannelization mentioned in Section 2.3.3. So the SNR for a WiMAX user is given by Equation (6.2).

$$SIR_i = \frac{P_k / L_{k,i}}{N_0} \tag{6.2}$$

P_k total TX power [W] of BS k

N_0 broad band thermal noise [W]

$L_{k,i}$ path loss of user i in cell k including shadow fading and antenna gain

6.1.3 WiMAX Data Rate Determination

Similarly to GSM and UMTS, link adaptation was considered in a link level simulation and the envelope of the resulted data rate per subchannel vs. SIR mapping is taken as the WiMAX channel characteristic. Figure 6-1 draws the envelope.

Having known the number of subchannels allocated and the channel quality (SIR) for a specific user, the effective data rate is the product of the subchannel number and the subchannel rate corresponding to the SIR level.

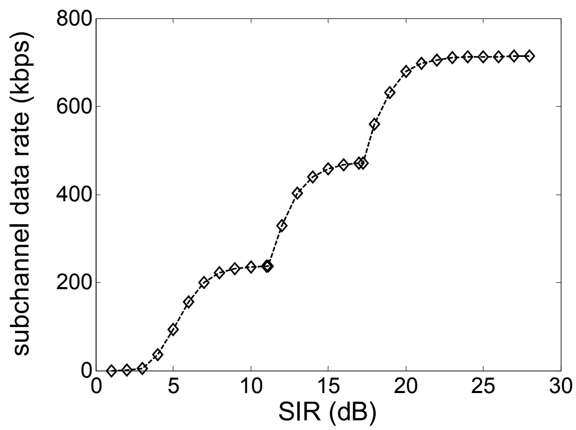

Figure 6-1: Envelope of subchannel data rate resulted from a link level simulation modelling the adaptive modulation and ARQ

6.2 Multi-Radio Access System Scenario

6.2.1 Three Radio Access Networks

The GSM and UMTS model remain the same as described in Section 4.3. WiMAX network is herein a single cell (dark circle in Figure 6-2) which overlaps partially the coverage of GSM and UMTS. The WiMAX hotspot is assumed employing an omni-antenna.

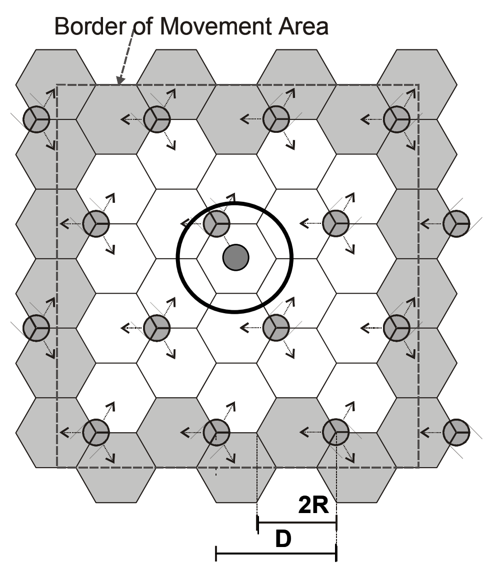

Figure 6-2: System scenario of the three access networks

A tight coupling of WLAN integrated to UMTS ([20], [31], [107]-[108]) is exploited by literature, where the integration is realised by building up an interface between the SGSN and an interworking unit (IWU, defined as an RNC emulator in [31]) of WLAN networks which is essentially a WLAN router. The network architecture is re-given in Figure 6-3. In this way, a WLAN network is regarded as an alternative access network with its own local resource management entity - IWU. The tight coupling shows several admiring features for intersystem handovers, such as short decision time, short delay time, etc. When introducing an interface between the two different networks, there exists also a challenge for a common standardisation of the interface.

We assume that the interworking of the three networks aligns with tight coupling. The standardisation work of WiMAX network supporting nomadic and mobile custom equipments is still on going. Motivated by the success of current mobile cellular networks and the vision of seamless roaming/handover in heterogeneous networks, WiMAX will most likely adopt to the tight coupling architecture..

Enabling this tight coupling architecture, the centralised or distributed MRRM introduced previously are possible to be implemented for a multi-radio access networks consisting WiMAX, GERAN and UTRAN. However, the intersystem protocol is not the concentration of this thesis and is therefore not discussed in detail.

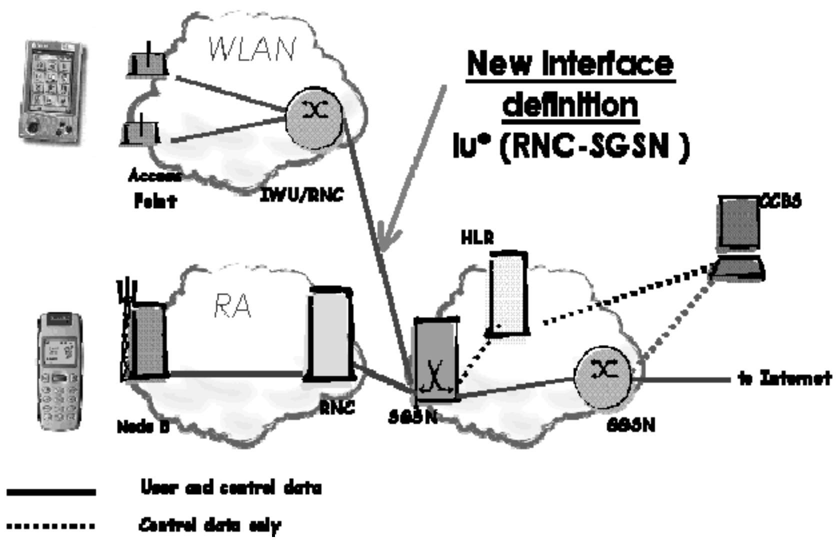

Figure 6-3: Tight coupling UMTS-WLAN [107]

6.2.2 3RAT-MRRM

6.2.2.1 User Data Rate Optimised MRRM

The MRRM algorithm is triggered per packet call. A PS session contains several packet calls and reading time in between. Upon each arrival of a packet call during a PS session, e.g. the start point of downloading a web page at web browsing, an MRRM algorithm is triggered to select the most suitable RAT to accommodate it. If the selected RAT is different from the last serving RAT (source RAT), an IS-HO will be executed. This IS-HO controlled by the network side is named inter-system cell change order (IS-CCO). So the evaluation and decision of an IS-CCO occurs per packet call. In addition, a user moving out of or into the WiMAX coverage may also trigger an IS-CCO. An IS-CCO triggered by such events is called WiMAX coverage triggered IS-CCO.

Since users typically prefer a fast download speed, we define an MRRM algorithm optimising the user data rate as follows:

1. The three candidate cells with the best C/I are identified by each local RRM. Note that if the user is outside of the WiMAX coverage, obviously only 2 RATs are available. Selecting the best candidate cell first of all in each RAT, is more practical than using pure load balancing, since the near-far effect is avoided.
2. The cells know their own current average user data rate R_{av} respectively, where $R_{av} \in [0, R_{max}]$.
3. The average data rate is mapped into four NRT load levels labelled as low, medium, high, and overload. This four-level NRT load information is aligned with the format of the 3GPP specification for the load information element [11]. Previously in Section 4.2.5, Figure 4-5 illustrates the mapping scheme. The mapping curve can be a linear, a convex or a concave curve. A concave curve, which gives more dense resolution in the high data rate range, is suitable for low load situation. A convex curve is appropriate for high load situation.
4. The cell with the lowest NRT load level will be selected to serve the newly arriving packet call. If all the cells have the same NRT load level, the source RAT remains the serving RAT. In this case, note that the candidate cell may be different from the last serving cell, affected by the user mobility. If the cells in the alternative RATs have identical NRT load levels which are smaller than the one in the source cell, the algorithm randomly selects one alternative RAT for the IS-CCO.

Simply put, a packet call is allocated into the cell within multi-radio access networks, which has the best signal quality and the best average data rate represented by the NRT load information at the time of its arrival. This cell is expected also offering the best data rate for the new packet call.

The relatively large hysteresis of NRT load mapping suits a dynamically varying system in reducing the IS-CCO rate and the ping-pong effect. The large hysteresis on the other hand exposes one shortcoming. For example, the MRRM algorithm will take 300kbps equal to 200kbps if the two rates are mapped to an identical NRT load value. In this case no IS-CCO will be decided to win the 100kps gain.

6.2.2.2 Joint Preference and User Data Rate Optimised MRRM

A system with a large capacity can accommodate a large number of users while assigning a substantial bandwidth for each user. If the load situations of two systems are unclear, the system with the larger capacity has higher probability to offer a wider bandwidth for a specific user. These predictable effects lead us to a two phased MRRM process:

Firstly, a preliminary default RAT selection before the initial RAT selection stage depends on a priority list of the RATs following the capacity descending order. An indicator of the high priority RAT being full will force a lower priority to become the default RAT.

Secondly, the user data rate optimised MRRM in Section 6.2.2.1 carries out the initial RAT selection as well as the necessary IS-CCO decision.

The preliminary default RAT selection results in two benefits for this MRRM: One is to reduce the IS-CCO rate effectively by directing the traffic to the high capacity RAT before the MRRM evaluation. The other is to allocate the maximum bandwidth with a high probability in case of identical NRT load evaluated in the diverse access networks.

6.3 Simulation Results

6.3.1 System Configuration

The System configuration for GSM and UMTS networks in Table 3-1 is still valid. Table 6-1 lists the additional system configuration for WiMAX. An urban environment is assumed. The appropriate propagation loss models given in Section 3.4.3 are then applied to the radio access networks of HSDPA and EDGE. In order to distinguish a heavy loaded hotspot, additional traffic is uniformly generated in the coverage of WiMAX in particular in conjunction with the traffic uniformly generated in the movement area for GSM and UMTS in general. Users are supposed to randomly select an air interface. Therefore the initial traffic is modelled to be equally distributed to the three networks, i.e. the input traffic density in three networks is identical.

WiMAX BS TX power	3.16W
WiMAX omni antenna gain	11dBi
WiMAX thermal noise	-104.6dBm
WiMAX max. subchannel number	30
WiMAX downlink resource sharing	80%
WiMAX cut off SIR for coverage	5dB
Hotspot radius	1200m
Maximum data rate terminal supported	384kbps

Table 6-1: System configuration

User mobility is taken into account. The mobility models remain the ones introduced in Section 4.3.2.

System performance in the white areas is investigated for avoiding the boundary effect. According to the topology parameters, the hotspot is about 2.72 times GSM/UMTS cell area.

6.3.2 Capability of Separate Systems

Simulation results of separate systems, see Figure 6-4, show the mean data rate as a function of carried load inside the hotspot for our configuration. Inside the hotspot area, the maximum carried load of WiMAX is triple of the GSM and more than double of the UMTS carried load, respectively. For the same load, e.g. 2Mbps, the mean data rate of WiMAX users (limited by the maximum MS supported data rate) is around double of UMTS users while fourfold of GSM users.

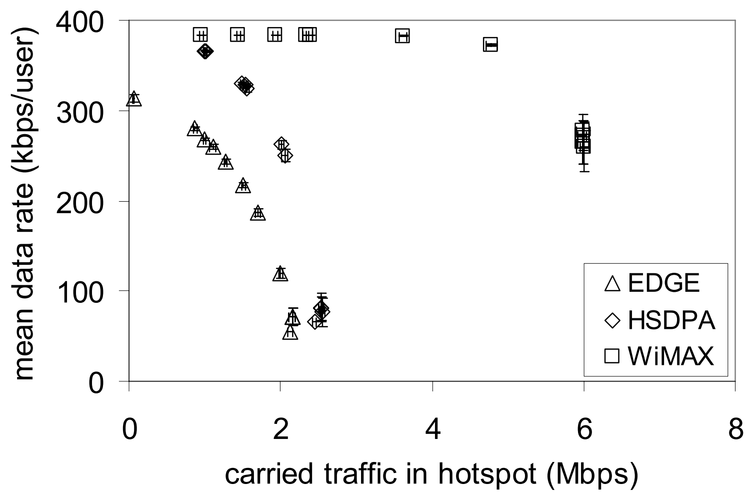

Figure 6-4: System capability of the three separate radio access networks, where 99% confidence intervals are shown for both variables

The mean data rate descends clearly with increasing traffic load. The larger the gradient of a curve (which is a negative value in Figure 6-4), the larger the data rate a network can offer at a particularly given load.

If the joint preference and user data rate optimised MRRM is implemented, WiMAX should have the highest priority in the first phase of the default RAT selection. The second should be UMTS. GSM should have the lowest priority.

6.3.3 WiMAX Integrated to GSM and UMTS

We study the following scenarios:

1. Separate systems: a baseline without any MRRM strategy with random RAT initialisation by users denoted rand_ss
2. Separate WiMAX: denoted rand_sw, where GSM and UMTS are integrated by user data rate optimised MRRM and share load initiated in them via user data rate optimised MRRM despite the separation of WiMAX. Users also randomly select the initial RAT to access.

3. WiMAX integrated to GSM and UMTS: The user data rate optimised MRRM is denoted rand_3rat.
4. WiMAX integrated to GSM and UMTS: The joint preference and user data rate optimised MRRM is denoted pref_3rat.

Four fixed amounts of the overall input traffic are used to investigate the MRRM effects. MRRM effects on load redistribution, mean data rates and IS-CCO activity are investigated. The mean data rate is calculated as the average user data rates of all completed packet calls.

6.3.3.1 Load Redistribution

The MRRM algorithms introduced in Section II have implicitly redistributed the load among the three networks, see Figure 6-5. Pursuing the maximum data rate possible, every packet-call and consequently every packet session is directed to the most appropriate network decided by the algorithms.

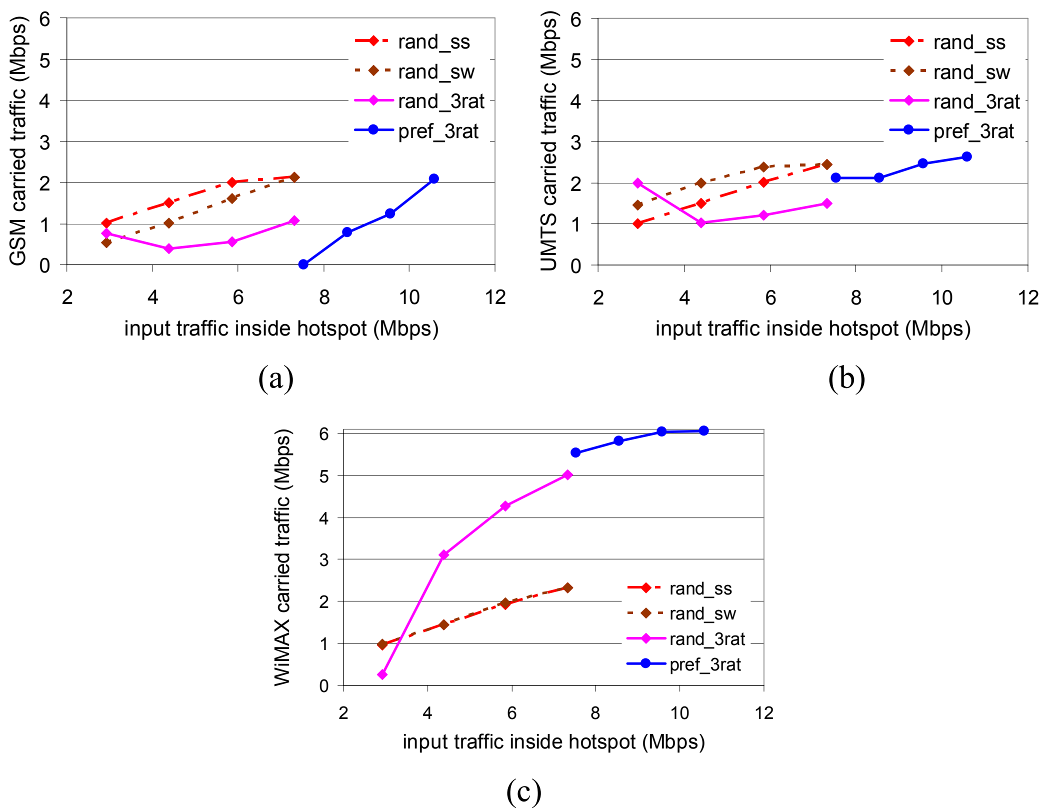

Figure 6-5: MRRM effect on the traffic redistribution inside the hotspot, (a) GSM, (b) UMTS, (c) WiMAX

MRRM redistributes the traffic from GSM to UMTS when WiMAX is separated, while it redistributes the traffic from both GSM and UMTS to WiMAX when WiMAX is integrated. The system capacities in terms of system throughput are around 2Mbps, 2.5Mbps, and 6Mbps inside the hotspot area for the GSM, UMTS, and WiMAX, respectively.

At 7Mbps input traffic, the GSM and the UMTS networks are saturated with rand_ss and rand_sw. With WiMAX integrated, the joint preference and user data rate optimised MRRM clearly distributes more traffic to WiMAX and UMTS than GSM.

6.3.3.2 Mean Data Rate

The mean data rate as a function of the input traffic in the hotspot area is shown in Figure 6-6 is shown for the four scenarios. If users select a RAT to initiate a session, the same traffic density will be resulted in three networks. This leads to the situation that before WiMAX is half full, GSM and UMTS are saturated. Due to the GSM/UMTS overload situation in the large area outside the hotspot, the curves corresponding to users' random RAT selection stop at less than 8Mbps. One might expect that the mean data rate decreases with the increasing input traffic. However, the left most point of the curve rand_3rat indicates a lower mean data rate than the next point with a higher input traffic. This can be understood from the joint effects of large hysteresis in NRT load mapping and the random traffic filling. At this low load situation (around 1Mbps each network inside the hotspot area), the average user data rate in an EDGE/HSDPA cell results in a relative high level, which corresponds to a low NRT load information due to the large hysteresis. Although the user data rate in WiMAX is larger, see Figure 6-4, we believe the corresponding NRT load information is the same as in EDGE/HSDPA. This can especially be understood from the fact that we have three NRT load mapping thresholds and therefore a large hysteresis. Thus, users are staying in EDGE/HSDPA but are not re-allocated to WiMAX for the possible data rate gain.

If users select a RAT to initiate a session, the same traffic density will be resulted in three networks. This is shown by rand_ss and rand_sw in Figure 6-5. However, with separate WiMAX, the free capacity in WiMAX is not available for the other two networks. So WiMAX cannot be used to solve an overload situation of GSM/UMTS. Since the traffic is symmetrically initiated in GSM and UMTS, the relatively similar capacity limits of GSM and UMTS result in limited MRRM gain for the mean data rate, referring to the narrow gap between rand_ss and rand_sw in Figure 6-6. It can be seen from Figure 6-5 that MRRM has in fact redistributed the load between GSM and UMTS at the left side, up to 6Mbps input traffic. The load shifts by around 0.5Mbps from GSM to UMTS. Nevertheless, the reduction and the increase of data rates in the two networks almost average out since the gradients of GSM and UMTS datarate vs. load curves in this range are more or less same, see Figure 6-4.

Integrating the high capability WiMAX network into GSM and UMTS, the MRRM gain on user data rate is quite significant. Especially when the two networks get into overload, which resulting in low data rate in rand_sw and rand_ss cases, the WiMAX resources joined draw the data rate back to a high level. The underlying reasons are twofold. On the one hand, the WiMAX capacity is definitely larger than the two networks. On the other hand, the derivative of data rate with respect to load in the WiMAX network is quite small, which means a large amount of increased load sacrifices the data rate slightly.

The benefit of the joint algorithm on mean data rate is even more significant. We take insight to the quantitative gain by looking at some selective data points in Figure 6-6. At 7Mbps input traffic, the GSM and UMTS networks are overloaded with separate WiMAX and separate systems. However, the mean data rate using the joint algorithm (Curve pref_3rat) already reaches the maximum data rate terminals support (384kbps given in Table 6-1). The gain is 30% over the mean data rate yield by the user data rate optimised MRRM (Curve rand_3rat), while 130% over the mean data rate of separate WiMAX (Curve rand_sw) or separate systems (Curve rand_ss). The default RAT selection after the list of RAT priority compensates the shortcoming of the relatively large hysteresis of the four levels NRT load information mapping and effectively promotes the MRRM performance.

At 300kbps mean data rate, the comparison of the input traffic of various MRRM strategies reveals the effect of traffic distribution. Based on users' random selection of the initial RAT, i.e. same traffic density in every network, the load afforded is increased from 4.6Mbps (Curve rand_ss and rand_sw) and so on to 7.3Mbps (Curve rand_3rat) by implementing the user data rate optimised MRRM among 3 RAT. Based on the default RAT selection after the list of RAT priority, i.e. large traffic density in a network with high capacity, the load afforded is further increased to 9.5Mbps (Curve pref_3rat).

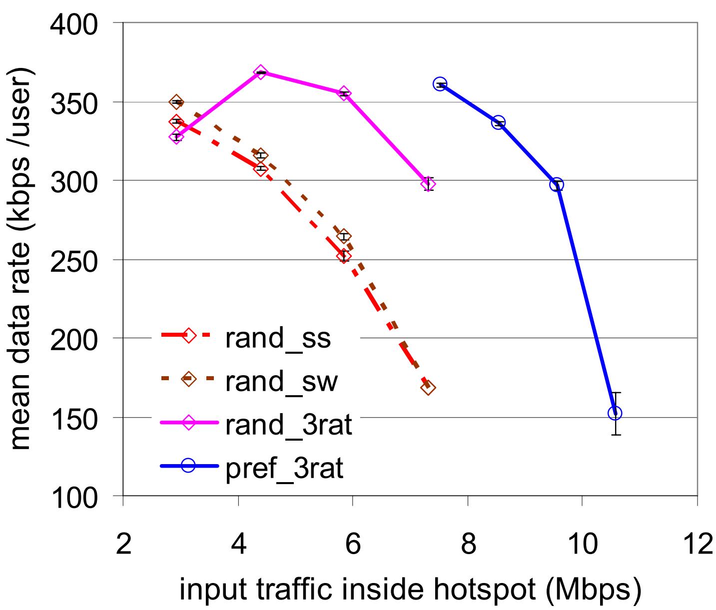

Figure 6-6: MRRM effect on the mean data rate where 99% confidence intervals are shown

6.3.3.3 IS-CCO Activity

To study the IS-CCO activity, we look at the number of IS-CCO per PS session. One PS session has statistically 5 packet calls which may trigger IS-CCOs. The overall input traffic inside the movement area drawn in Figure 6-2 is counted for the IS-CCO rate in Figure 6-7, if the traffic is the subject of MRRM.

The default RAT selection after the list of RAT priority of the joint algorithm is not counted as IS-CCO.

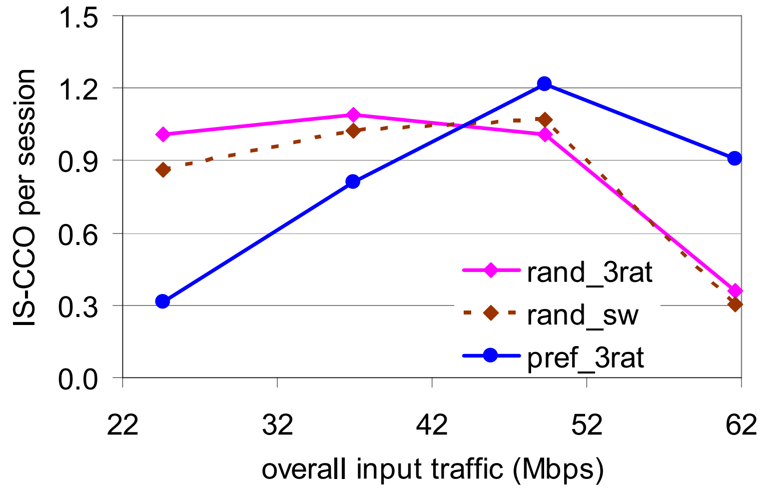

Figure 6-7: IS-CCO activities

On average, the IS-CCO rate is around once per session. WiMAX integrated into GSM/UMTS heterogeneous networks does not increase the IS-CCO rate too much in the case of random default RAT selection. The very low IS-CCO rate for low input traffic of pref_3rat indicates that the default RAT selection after the list of RAT priority, though simple and direct, has achieved a nearly correct constellation of the traffic and has meanwhile avoided frequent IS-CCOs. In contrast to the low load side, the joint algorithm is associated with high IS-CCO rate at the high load side. This can be explained by the large traffic density at the hotspot area (see Figure 6-5) where three networks are available for IS-CCO.

6.4 Summary

Multi-radio resource management (MRRM) for HSDPA, EDGE and WiMAX is studied in our network simulator. In this simulator, the dynamic radio channel conditions and the packet scheduling mechanisms are modelled for the three networks. Due to the fact of disparate capacity of these three radio access networks, our assertion is that a gain of data rate can be obtained if MRRM re-allocates the traffic to the network which is supporting the highest data rate. An MRRM algorithm based on user data rate related NRT load information has shown the ability of improving the user data rate, redistributing load in heterogeneous networks. The simulation results here show that the more disparate the capacity of the underlying systems, the higher the MRRM gain.

An additional simple default RAT selection after a prioritised list of RAT in descending order of capacity increases the MRRM gain of the mean data rate by 30%. Compared to the mean data rate without MRRM, the gain is 130%. If we take as a reference point that the heterogeneous networks maintain a high level mean data rate at 300kbps, the load supported inside the hotspot area is increased by 59% with the user data rate optimised MRRM, and by 107% with the joint preference and user data rate optimised MRRM.

The reason for the additional gain from the joint preference and user data rate optimised MRRM algorithm is the disparate gradients of data rate vs. load in various networks. For example, if a certain amount of traffic is transferred from GSM to WiMAX, the mean data rate gain in GSM is far greater than the mean data rate loss in WiMAX. The larger the gradient of a data rate vs. load curve, the larger the data rate a network can offer at a particularly given load.

Chapter 7

Conclusion

Sharing of scarce radio resources, initial RAT selection and seamless intersystem handover are envisaged as salient features of the next generation networks comprising desperate multi-radio access networks. The corresponding radio resource management functionality is called multi-radio resource management (MRRM). This thesis contributes the MRRM decision algorithms for integrated services and the implementation of MRRM in the realistic multi-radio access networks consisting GSM/EDGE, UMTS/HSDPA and WiMAX. The issues discussed are as follows:

- Local radio resource management mechanisms
- Theoretical analysis for verifying the simulation model via a case study: pure voice service in GSM and UMTS systems
- Performance of integrated service of heterogeneous networks where various MRRM strategies are applied to the heterogeneous networks
- Distributed MRRM implementation utilising the legacy signalling capability of the 3GPP networks
- MRRM strategies extended to three radio access networks for NRT services

This thesis studies the MRRM decision algorithms and MRRM implementations mainly based on event-driven simulations since the complex air interface specific radio characteristics, which are considered essence to represent a local radio access network, are easily adopted into a simulation model.

7.1 What Is the Local Radio Resource Management Good for?

The local radio resource management is usually a group of management mechanisms, including admission control, power control, packet scheduling, intrasystem handover, and load control. The local radio resource management of GSM/EDGE and UMTS/HSDPA systems already are built in features of the air interfaces. A single radio access network typically differs from others by the layer 1 and layer 2 technologies and the RRM adaptive to the two layers. Various mechanisms of radio resource management therefore define the characteristics of the radio access network in conjunction with other access technology specific air interface characteristics. We have firstly considered the local radio resource management as the intrinsic features of each single underlying network in the heterogeneous environment.

Other groups, who have studied the common radio resource management, to our knowledge do not base the CRRM studies on a complete functional set of local RRM. For example, the complete air interface characteristics are ignored by Tölli et al. [25], where even the radio propagation and the MAC/RLC are not modelled. IST EVEREST project gives another example. Researchers of this project have explicitly discussed that the interaction levels of local RRM and CRRM will go deep with the developing integration of diverse radio access networks [23]. However, the interaction level of their simulation tool is same as the level addressed in this dissertation. That is the MRRM entity manages initial RAT selection and intersystem handover whereas the local RRM entities manage admission control, handover, scheduling, power control and congestion control (also known as load control). In their simulation tool, local RRM mechanisms such as resource reservation, service priority, and soft handover are not considered.

In the case study of voice service in GERAN and UTRAN in Chapter 3, we have also learned the quantitative impact of resource reservation, mobility, and handover criteria on the system performance, through both analytical and simulation models. This helps to verify the significance of the local RRM also. It is worth considering the realistic local RRM in the study of MRRM.

7.2 Evaluation Methods

To evaluate MRRM is a quite complex issue. The admission control in a single network alone is already a sufficient topic for research as studied in [37]-[52]. We have use analytical models to learn some basic RRM issues, such as hard handover, resource reservation for handover traffic, and mobility influence. The analytical models are used to validate our simulator tool with corresponding features as well. The evaluation of the MRRM approaches is done through a dynamic event-driven simulation.

The complexity of a real network makes it difficult for an analytical model to reveal every feature of the network. An analytical model is usually based on many simplifications and assumptions. Some additional features involved into the simulation model have shown that the system features do vary the system performance significantly. Examples of these features are sectorised antenna, mixed mobility pattern (e.g. co-existing pedestrian and vehicular users), handover hysteresis and handover retry etc. Moreover, the MRRM at call level is difficult to be studied with an analytical model.

The study about MRRM in the literature is mainly based on various simulation tools, see all the references listed later concerning MRRM or CRRM. With respect to the underlying RAT, GSM/UMTS, UMTS/WLAN and GPRS/WLAN heterogeneous scenarios are typically investigated. WiMAX has not been considered by others. GERAN is used to label the GSM/EDGE radio access network by the IST EVEREST project. However, GPRS is simulated for serving data services in the GSM system. Although EDGE is mentioned by [28], the multiservice capacity of GSM/EDGE is derived from the pure voice capacity and pure WWW capacity. The derived capacity limit of GSM/EDGE, namely data users vs. voice users, is therefore a linear curve, which is different from the non-linear curve resulted in our event-driven simulation tool. With regard to the mobility pattern, others have considered a merely pedestrian pattern; a mixed mobility pattern is enabled in our simulation tool.

7.3 Performance Gain Brought by LIAM

MRRM manages initial RAT selection and intersystem handover (IS-HO) if the local RRM manages the admission control, scheduling, handover, power control and load control in each single system. It is a common agreement that various services have different quality requirements. Thus, the MRRM should concern the service type. The service-based MRRM policy has been adopted by many researchers, e.g. [29]-[31]. Load balancing is another popular strategy for MRRM adopted by [24]-[26], [30] and [32]. Nevertheless, the load definition itself is varying in literature based on the assumptions made. Moreover, authors of [32] addresses the load balancing for the voice service though handover cost and QoS are jointly compared, i.e. voice IS-HO decision by a cost function. CRRM algorithms for IS-HO are especially proposed in [24]-[25] and [30]. However, all are based on the overlay network assumption without particular consideration for radio access technologies. The others merely study initial RAT selection. In this dissertation, MRRM algorithms for both initial RAT selection and intersystem handover are contributed, providing quality of service for integrated voice and data services in heterogeneous networks. Offering more opportunities of resource sharing than merely initial RAT selection, joint initial RAT selection and IS-HO better exploits the diversity of the heterogeneous networks.

By minimising the resource consumption which ensures a minimum QoS, the capacity of multi-radio access networks is maximised [27]-[28]. However the capacity optimisation should not be the solitary goal of MRRM. Especially in cases when the traffic load is less than the capacity limit, MRRM should optimise the performance of every individual service. The contributions of the Load Information Aiding MRRM (LIAM) approach in this thesis lie on several points:

- LIAM is a comprehensive MRRM for multiservices considering the interplay of multiservices. In our case study of service mix, i.e. integrated voice and web-browsing services, the voice service has absolute priority over the web-browsing service. This priority is guaranteed by particularly setting the cell load equal to the RT load at the MRRM evaluation for every voice call.
- LIAM is suitable for both initial RAT selection and IS-HO. The capacity gain and the performance gain are further evaluated via event-driven simulations, which model realistic systems at call level described in the beginning of Chapter 4.

- LIAM uses various methods to optimise the QoS. Firstly, imperative IS-HO/initial RAT selection is adopted for RT services, minimising the delay caused by MRRM. The load-aware MRRM evaluation maximises the success of the MRRM decision compared to another imperative IS-HO/initial RAT selection, namely "Blind". The capacity gain of RT services developed by the LIAM approach is shown not less than the gain of the load balancing algorithm [101]. Secondly, packet call arrivals are selected to trigger IS-HO of NRT services. Packet calls containing several IP datagram have reading time (idle time) in between. We believe that this triggering behaves rightly from the aspect of timing and packet loss minimisation due to MRRM. Thirdly, NRT load in our approach is based on the user data rates instead of the one based on the resource consumption in literature [26]-[28]. This type of NRT load information leads to a direct result in user data rate optimisation and consequently transmission delay minimisation.
- LIAM makes the decision mainly based on the load information. The format of the load information is aligned with the format of information element carried by the 3GPP signalling defined in the 3GPP standard. This contributes one advantage that the implementation of LIAM in GSM/EDGE and UMTS/HSPA may utilise the existing 3GPP signalling.

We have carried out the comparison between LIAM, currently available MRRM, and service-based MRRM through simulation investigation. Results show that LIAM has achieved the maximal performance gain but a sub-optimal capacity gain of mixed service and an optimal capacity gain of pure voice service. Except the case when every single network is fully loaded, LIAM becomes the best solution for MRRM.

7.4 Distributed MRRM: Benefit and Degradation

Distributed MRRM locates at each radio access network, most likely co-located with the local radio resource management entity. The concept of distributed MRRM was already elaborated both by 3GPP [24] and IST EVERST [23]. But, what we proposed is the method gathering the necessary load information for MRRM decision based on a distributed MRRM architecture. Merely from the distributed architecture, the local load information is automatically available at the MRRM since the information is anyway available at the local RRM. We propose that the distributed MRRM works with a database which stores the load information concerning the alternative systems. The pair of load information stored is updated by the latest IS-HO involving the respective cells, realistic or dummy.

The main benefit from the proposal is the mitigation of - up to complete elimination of - additional signalling for load information gather, at the expense of a few percent degradation of the system performance.

This type of distributed MRRM is especially suitable for short term implementation of MRRM in the 3GPP radio access networks. Limited by the signalling procedure, the available information leads to sub-optimal MRRM performance. Nevertheless, this distributed MRRM enables the feasibility of multiple MRRM strategies if multiple operators control the underlying radio access networks.

7.5 3RAT MRRM: Implementation and Performance

Multi-radio resource management (MRRM) for heterogeneous networks comprising HSDPA, EDGE and WiMAX is studied in our network simulator as well. Due to the fact of disparate capacity of these radio access networks, our assertion is that a gain of data rate can be obtained if MRRM re-allocates the traffic to the network which is supporting the highest data rate. The user data rate optimised MRRM algorithm is in fact the strategy dealing with data services of LIAM used for the integrated services. So the usage of NRT load information also contributes to the user data rate gain, redistributing load in heterogeneous networks.

From our simulation study, we have learned that the more disparate the capacity of the underlying systems, the higher the MRRM gain will be. We also propose a joint preference and user data rate optimised MRRM which selects a default RAT following a prioritised list of RAT in descending order of capacity before starting the user data rate optimised algorithm. This joint algorithm further increases the MRRM gain of the mean data rate by 30%. Compared to the mean data rate without MRRM, the gain is 130%.

The joint algorithm combines the utilisation of system-level resource efficiency and the optimisation of user data rate. Moreover, the network-based MRRM dynamically detects the system-level resource efficiency based on the capacity information obtained. This prevents the difficulty of the detection of resource efficiency at call level required by Furuskär's proposal [27]. On the other hand, our proposal is also flexible than the predefined assumption of resource efficiency in service-based policy.

7.6 Future Work

Future work might be explored on two aspects: developing advanced MRRM algorithm and developing an analytical model of a heterogeneous environment for the algorithm investigation.

The analytical model is anticipated to be compromised by several functional components including one component for each access network considering various admission control policies in individual systems and air interface attributes of individual systems, and one component for the common multi-radio access control (initial RAT selection and intersystem handover) policies/algorithms. The suitability of the optimisation algorithms aimed relies on the issues to be optimised, e.g. the system throughput, the user perceived performance, the economical utility, etc.

There are many advanced admission control schemes for multiservices in a single system can be adopted, like for example the resource sharing policies among multiservices, the prioritisation of multiservices, the guard bandwidth reservation for handover traffic, minimising long term outage probability and graceful degradation of QoS instead of hard rejection, and so on. Dynamic resource boundary (moving resource boundary) for multiservices can be adopted. For instance, voice service shares 50% - 70%; streaming service shares 10% - 20%; non-real-time service shares 10% - 30%. The handover decision making needs also attentions. Selecting the most expensive call as a handover candidate should be more resource efficient scheme than requesting handover directly by the call which is facing reduced signal quality.

Based on the network characteristics obtained by the individual system model for multiservices, some advanced common multi-radio access controls can be evaluated. Such advanced common multi-radio access control should result in adaptive resource sharing based on the stochastic current system status including signal quality, load status, predicted resource consumption, resource efficiency. Some economic strategies should be also included into the algorithm, such as the utility function of different services and the cost of intersystem signalling. Furthermore operator and user preference may also involve into the algorithm.

Considering so many aspects of a common access control, weighting factors are required to differentiate the importance or priorities of the different aspects/parameters. Tuning the weighting factors, the algorithm can achieve optimisation for different objectives, e.g. capacity optimisation or QoS optimisation. Because of the potential challenge met by the comparability of variant sub networks, the fuzzy logic is desired rather than direct comparison of the aforementioned stochastic system status in different sub networks.

Appendix A. Log-normal Sum Approximation

According to [83], the moment generating function (MGF) of a random variable X is defined as

$$\Psi_X(s) = \int_0^{\infty} exp(-sx) p_X(x)dx, \ (Re(s) \geq 0) \tag{7.1}$$

The MGF of the sum of independent random variables also possesses this desirable property that it can be written directly in terms of the MGFs of the individual random variables as follows:

$$\Psi_{(\sum_{i=1}^{K} X_i}(s) = \prod_{i=1}^{K} \Psi_{X_i}(s), \ (Re(s) \geq 0) \tag{7.2}$$

Gauss-Hermite integration is used to express the log-normal MGF:

$$\Psi_{X_i}(s) = \int_0^{\infty} \frac{\xi \, exp(-sx)}{x\sigma_Y \sqrt{2\pi}} exp\left[-\frac{(\xi \ln x - \mu_Y}{2{\sigma_Y}^2}\right] dx \tag{7.3}$$

$$\approx \sum_{n=1}^{N} \frac{\varpi_n}{\sqrt{\pi}} exp\left[-s \, exp(\frac{\sqrt{2}\sigma_Y a_n + \mu_Y}{\xi}\right]$$

Where μ_Y and σ_Y are the mean and the standard deviation of the Gaussian random variable $Y = 10 \log_{10}(X)$, N is the Hermite integration order, ξ is a scaling constant, ϖ_n and a_n are weights and abscissas of Hermite integration tabulated [109] as follows:

No. N	***Abscissas*** a_n	***Weight*** ϖ_n	***Total Weight*** $\varpi_n \, exp({a_n}^2)$
1	-5.38748089001	2.22939364554E-013	0.898591961453
2	-4.60368244955	4.39934099226E-010	0.704332961176
3	-3.94476404012	1.08606937077E-007	0.62227869619
4	-3.34785456738	7.8025564785E-006	0.575262442852
5	-2.78880605843	0.000228338636017	0.544851742366
6	-2.25497400209	0.00324377334224	0.524080350949
7	-1.73853771212	0.0248105208875	0.509679027117
8	-1.2340762154	0.10901720602	0.499920871336
9	-0.737473728545	0.286675505363	0.493843385272
10	-0.245340708301	0.462243669601	0.490921500667

11	0.245340708301	0.462243669601	0.490921500667
12	0.737473728545	0.286675505363	0.493843385272
13	1.2340762154	0.10901720602	0.499920871336
14	1.73853771212	0.0248105208875	0.509679027117
15	2.25497400209	0.00324377334224	0.524080350949
16	2.78880605843	0.000228338636017	0.544851742366
17	3.34785456738	7.8025564785E-006	0.575262442852
18	3.94476404012	1.08606937077E-007	0.62227869619
19	4.60368244955	4.39934099226E-010	0.704332961176
20	5.38748089001	2.22939364554E-013	0.898591961453

Table A-1: Weights and abscissas of Hermite integration

The log-normal sum $\sum_{i=1}^{K} X_i$ can now be approximated by a log-normal variable $X = 10^{0.1Y}$, where Y is a normal distribution with mean μ_Y and standard deviation σ_Y, by matching the MGF of X with the MGF of $\sum_{i=1}^{K} X_i$ at two different values of s: s_1 and s_2. The μ_Y and σ_Y are then the solutions of two independent equations:

$$\sum_{n=1}^{N} \frac{\varpi_n}{\sqrt{\pi}} exp\left[-s\, exp(\frac{\sqrt{2}\sigma_Y a_n + \mu_Y}{\xi} \right] \tag{7.4}$$

$$= \prod_{i=1}^{K} \Psi_{X_i}(s_m; \mu_{X_i}, \sigma_{X_i}), for\ m = 1, 2$$

Appendix B. Link Level Channel Capacity

B.1 HSDPA Channel Capacity

For HSDPA transmissions, 15 spreading codes with SF16 are reserved for a pure PS traffic case while 10 spreading code (SF16) are reserved in the scenario with integrated services. Link adaptation, i.e. adaptive modulation and coding scheme, and retransmission were studied in a link level simulation and the resulted effective data rate per 15 (10) codes vs. SIR mapping is adopted to determine user data rate based on the channel conditions. The resulted mapping is shown in Figure B-1.

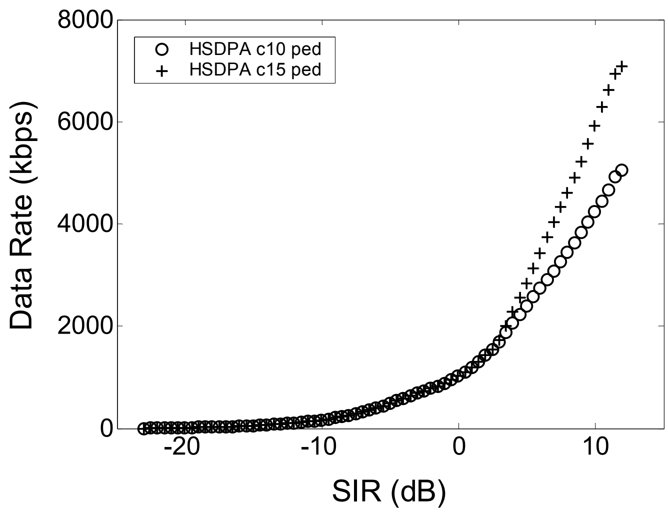

Figure B-1: Data rate and SIR mapping obtained by HSDPA link level simulations

B.2 EDGE Channel Capacity

Link adaptation was considered in a link level simulation and the resulted data rate per time slot vs. SIR mapping is taken to determine the user data rate based on the channel conditions. The resulted mapping is shown in Figure B-2.

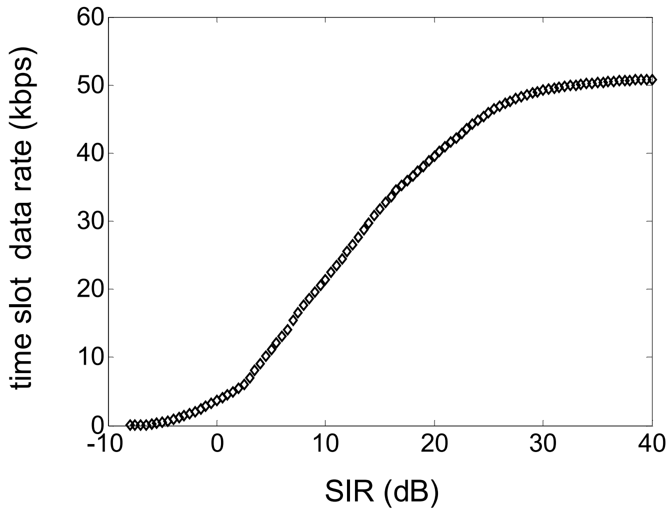

Figure B-2: Data rate and SIR mapping obtained by EDGE link level simulations

B.3 Channel Capacity Dependence on Location

In contrast to the circular WiMAX omni-antenna, HDSPA and EDGE use 120° sectorised antenna. Obviously, the direction-dependent antenna gain differ the SIR's of users located on the circumference of a circle centred at the BS. The datarate per channel was calculated along the direction of two angles, for bore sight in the centre axis of the antenna, and for an angle of 6/π to the bore sight, which corresponds to a 30° angle to the centre of the antenna.

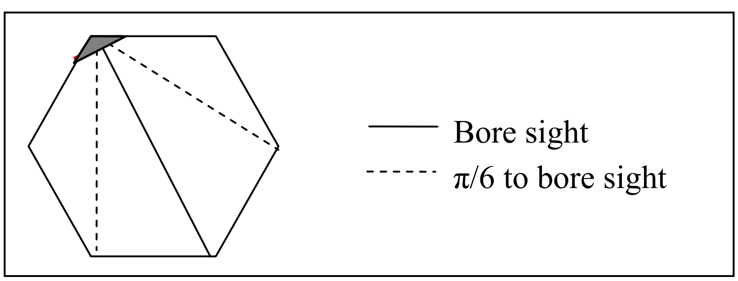

Figure B-3: Illustration of the bore sight and 30° to bore sight of a 3-sectorised antenna

Figure B-4 shows the HSDPA cell capacity assuming a single stationary user at different distances along the direction of bore sight and along the direction of 6/π (30°) away from bore sight. In the case that there should be more than one user in the cell, then this cell capacity would have been divided by the number of active users in order to obtain the datarate per user. The channel data rate is location dependent and HSDPA users at different locations encounter also different interference. At larger angles (30° to 60°) away from the bore sight direction, the datarate decreases even further than shown in this Figure B-4 for the 30° direction.

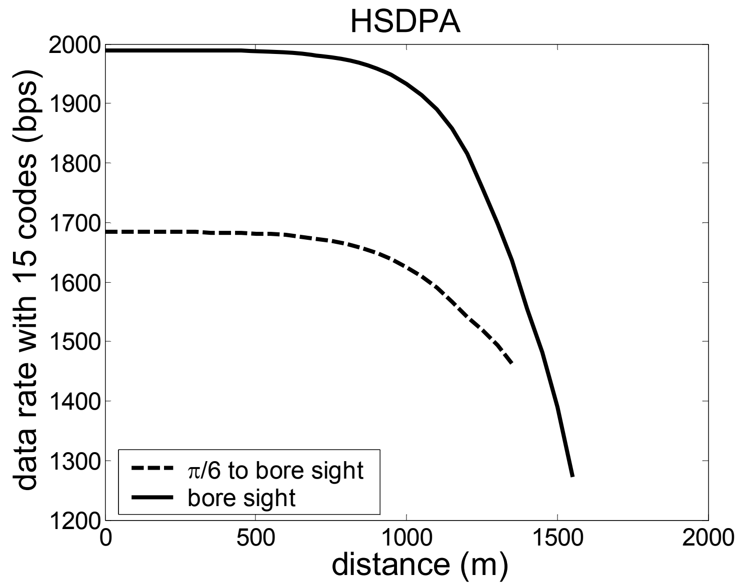

Figure B-4: HSDPA channel datarate versus distance

Figure B-5 shows the EDGE timeslot capacity assuming one single stationary user at different distances from the base station along the bore sight and along the direction of 6/π (30°) away from the bore sight. One user may occupy up to 8 timeslots. The timeslot data rate is mainly distance dependent similar to the situation for as WiMAX with its circular omni antenna. EDGE users at different locations encounter a fixed interference.

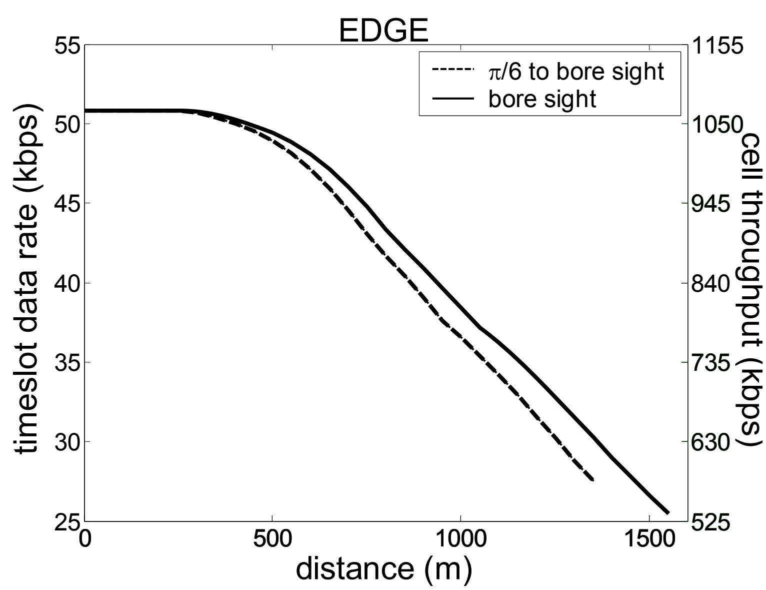

Figure B-5: EDGE timeslot datarate versus distance

As a consequence, the user datarate in HSDPA and EDGE can vary by a certain amount, depending on where the user is located in the cell and in which angle from the base-station antenna. In contrast, the implemented WiMAX air-interface operates with a circular omni antenna, which provides a uniform datarate only depending on the distance and not on the angle to the antenna.

B.4 WiMAX Channel Capacity

The datarate for a single WiMAX subchannel is obtained by a calculation of the WiMAX pathloss and by a mapping of the SNR value to the received user datarate, also considering a cut-off SNR value of here 5 dB. So far, physical link layer simulations were available for three modulation schemes, which then lead for a 10 MHz bandwidth to the following subchannel/cell throughput as shown in Figure B-6.

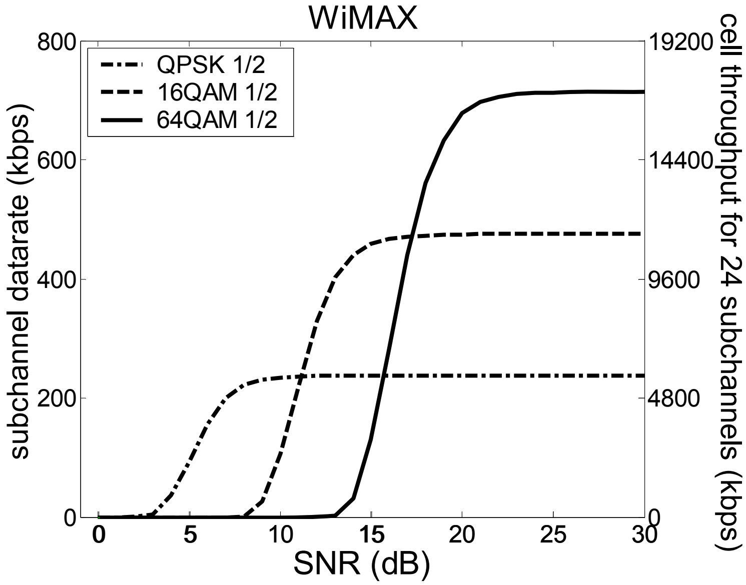

Figure B-6: Throughput in the one WiMAX hotspot with a bandwidth of 10 MHz for a single subchannel and for 24 subchannels

Since we use an omni antenna the datarate does not depend on the direction, and since we use constant interference / thermal noise values, the resulting datarate thus depends basically only on the distance.

Assuming that all stationary WiMAX users are located at the same distance from the base station, this then leads to the datarate for one single subchannel and the WiMAX cell throughput given in Figure B-7.

Thus, the WiMAX capacity per subchannel and the coverage of the WiMAX hotspot depends on the system configuration and on some simulation parameters. In particular the following ones were used in this study:

- Terrain form factor for urban and suburban scenarios.
- The bandwidth, which also influences the background thermal noise value
- Mobility: Moving users receive a penalty on the received SNR value, which reduces the datarate at a certain distance and decreased the maximum distance in which this user can still receive a sufficiently strong WiMAX signal.

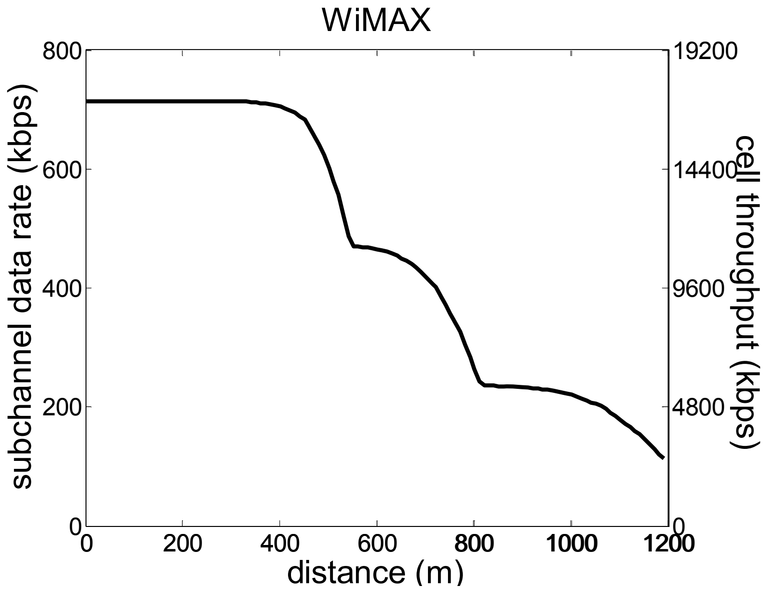

Figure B-7: Datarate versus the distance from the base station to a stationary user. The left axis shows the datarate which one stationary user receives per single subchannel. The right axis shows the datarate for 24 subchannels together, which is the cell throughput which all users would obtain, in the case that all users are located at the same distance.

Reference

[1] http://www.wimaxforum.org, Sep. 18 2006

[2] IEEE 802.16-2004

[3] A. Ghosh, D. R. Wolter, J. G. Andrews, R. Chen, "Broadband Wireless Access with WiMAX/802.16: current performance benchmarks and future potential", *IEEE Communication Magazine*, Volume 43, Issue 2, Feb 2005, pp:129 – 136

[4] C. Hoymann, "Analysis and Performance Evaluation of the OFDM-based Metropolitan Area Network IEEE 802.16", *Computer Networks* 49 (2005) pp341-363

[5] K. Aretz, M. Haardt, W. Konhäuser and W. Mohr, "The Future of Wireless Communications Beyond the Third Generation", *Computer Networks* 37 (2001), pp83-92

[6] R. Berezdivin, R. Breinig, and R. Topp, Raytheon, "Next-Generation Wireless Communications Concepts and Technologies", *IEEE Communication Magazine*, pp108.116, March 2002

[7] F. Berggren et al., "Multi-radio Resource Management for Ambient Networks", *the 16th IEEE International Symposium on Personal Indoor, and Mobile Radio Communications (PIMRC'05)*, Berlin, Germany, 11-14 Sep. 2005

[8] http://www.nokia.de/de/93/198154/e61technischedaten.pdf, 15. 05. 2006

[9] Überall dienstbereit, PC-Card für WLAN, GPRS und UMTS, http://www.heise.de/mobil/, 14. 02. 2005

[10] 3GPP Technical Specification Group Radio Access Network, "UTRAN Iu Interface RANAP Signalling", Release 5, TS 25.413, 2003, http://www.3gpp.org

[11] 3GPP Technical Specification Group GSM/EDGE Radio Access Network, "MSC-BSS interface; Layer 3", Release 5, TS48.008, 2002, http://www.3gpp.org

[12] C. Perkins, "Mobile Networking in the Internet", *Mobile Networks and Applications*, 3 (1998) 319-334

[13] P. M. L. Chan et al., "Mobility Management Incorporating Fuzzy Logic for a Heterogeneous IP Environment", *IEEE Communication Magazine*, pp42-51, December 2001

[14] F. Siddiqui and Sh. Zeadally, "Mobility Management Across Hybrid Wireless Networks: Trends and Chanllenges", *Computer Communications* 29 (2006) 1363-1385

[15] R. Yavatkar, D. Pendarakis and R. Guerin, "A Framework for Policy-based Admission Control", RFC 2753, January 2000, http://www.ietf.org

[16] F. Zhu and J. McNair, "Multiservice Vertical Handoff Decision Algorithms", *EURASIP Journal on Wireless Communications and Networking*, Volume 2006, pp1-13

[17] P. Vidales et al., "Autonomic System for Mobility Support in 4G Networks", *IEEE Journal on Selected Areas in Communications,* Vol. 23, No. 12, December 2005, pp2288-2304

[18] R. Ferrús et al., "Vertical Handover Support in Coordinated Heterogeneous Radio Access Networks", *14th IST Mobile & Wireless Communications Summit*, Dresden, 19-23 June 2005

[19] W. Zhang, J. Jaehnert and K. Dolzer, "Design and Evaluation of a Handover Decision Strategy for 4th Generation Mobile Networks", *VTC2003-Spring*, 21-24 April 2003, Jeju, Korea

[20] 3rd Generation Partnership Project, "Feasibility Study to Wireless Local Area (WLAN) Interworking", V1.1.0, 3GPP TS 22.934, 2002-03

[21] M. Lott et al., "Cooperation of 4G Radio Networks with Legacy Systems", *14th IST Mobile & Wireless Communications Summit*, Dresden, 19-23 June 2005

[22] N. Nafisi el al., "Extending QoS Policy-based Mechanisms to B3G Mobile Access Networks", *IST Summit on Mobile & Wireless Communications*, Lyon, France, 27 – 30 June 2004

[23] X. Gelabert et al., "Radio Resource Management in Heterogeneous Networks", *3rd International Working Conference on Performance Modeling and Evalutation of Heterogeneous Networks*, HET-NET'05, Likley, UK, July 2005

[24] 3GPP Technical Specification Group Radio Access Network, "Improvement of RRM across RNS and RNS/BSS", TR 25.881-500, 2001, http://www.3gpp.org

[25] A. Tölli, P. Hakalin, and H.Holma, "Performance of Common Radio Resource Management (CRRM)", *in Proc. IEEE Int. Conf. Communications 2002 (ICC 2002)*, vol.5, Apr.-May 2002, pp 3429-3433

[26] X.Gelabert, J. Pérez-Romero, O. Sallent, R.Agusti, "On the Suitability of Load Balancing Principles in Heterogeneous Wireless Access Networks", *Wireless Personal Multimedia Communications, WPMC' 05*, Alaborg, 18-22 Sep. 2005

[27] A. Furuskär and J. Zander, "Multiservice Allocation for Multiaccess Wireless Systems", *IEEE Transaction on Wireless Communications*, Vol. 4, No. 1, January 2005

[28] G. Fodor, A. Furuskär and J. Lundsjö, "On Access Selection Techniques in Always Best Connected Networks", *Specialist Seminar on Performance Evaluation of Wireless and Mobile Systems, ITC16*, Antwerp, Belgium, 31 Aug. –02 Sep. 2004

[29] J. Pérez-Romero, O. Sallent and R. Agusti, "Policy-based Initial RAT Selection Algorithms in Heterogeneous Networks", *7th International Conference on Mobile and Wireless Communication Networks (MWCN'05),* Morocco, 19-21 Sep. 2005

[30] W. Zhang, "Performance of Real-time and Data Traffic in Heterogeneous Overlay Wireless Networks", *19th International Teletraffic Congress (ITC19)*, Beijing, China, 29 Aug. – 2 Sep. 2005

[31] O. Andrisano et al., "UMTS and WLAN Integration: Architectural Solution and Performance", *the 16th IEEE International Symposium on Personal Indoor, and Mobile Radio Communications (PIMRC'05)*, Berlin, Germany, 11-14 Sep. 2005

[32] A. Pillekeit, F. Derakhshan, E. Jugl and A. M-Thiel, "Force-based Load Balancing in Co-located UMTS/GSM Networks", *VTC-2004fall*, LA, USA, 26-29 Sep. 2004

[33] P. M. L. Chan, R. E. Sheriff, Y. F. Hu, P. Conforto and C. Tocci, "Mobility Management Incorporating Fuzzy Logic for a Heterogeneous IP Environment", *IEEE Communication Magazine,* December 2001, pp42-51

[34] L. Giupponi et al., "A Novel Joint Radio Resource Management Approach with Reinforcement Learning Mechanisms", *IEEE International Workshop on*

Radio Resource Management for Wireless Cellular Networks (RRM-WCN), Phoenix, April 2005

[35] L. Giupponi, R. Agusti, J. Pérez-Romero, O. Sallent, “Joint Radio Resource Management Algorithm for Multi-RAT Networks”, *IEEE GLOBECOM 2005*, St. Lious, USA, 28 Nov.- 2 Dec. 2005

[36] L. Giupponi, R. Agusti, J. Pérez-Romero, O. Sallent, “A Fuzzy Neutral Joint Radio Resource Management in a Multi-Cell Scenario Supporting a Multiservice Architecture”, *6th International Conference on 3G and Beyond 3G*, London, November 2005

[37] R. Ramjee, D. Towsley and R. Nagarajan, “On Optimal Call Admission Control in Cellular Networks”, *Wireless Networks*, pp 29-41, Vol. 3, Nr. 1, Jan. 1997

[38] V.C.M. Leung and F. Yu, “Mobility-Based Predictive Call Admission Control and Bandwidth Reservation in Wireless Cellular Networks”, *Proc. 20th Conf. on Computer Communications (IEEE INFOCOM)*, Anchorage, Alaska, pp518-526, 22-26 Apr. 2001

[39] S. Kim, T. Kwon, and Y. Choi, “Call Admission Control for Prioritized Adaptive Multimedia Services in Wireless/Mobile Networks”, *In Proc. of IEEE VTC'00 Spring*, Tokyo, Japan, 15-18 May 2000

[40] J. R. Moorman, J. W. Lockwood, and S.-M. Kang, “Real-time Prioritized Call Admission Control in a Base Station Scheduler”, *In Proc. of ACM international workshop on Wireless mobile multimedia*, pp 28-37, Boston, Aug. 2000

[41] J. M. Peha and A. Sutivong, “Admission Control Algorithms for Cellular Systems”, *Wireless Networks*, pp 117-126, Vol. 7, No. 2, Mar./Apr. 2001

[42] P. Ramanathan, K. Sivalingam, P. Agrawal, and S. Kishore, "Dynamic Resource Allocation Schemes during Handoff for Mobile Multimedia Wireless Networks," *IEEE J. on Selected Areas in Communications*, 17(7), Jul. 1999, pp. 1270-1283.

[43] Ki-Dong Lee , Sehun Kim, “Optimization for Adaptive Bandwidth Reservation in Wireless Multimedia Networks”, *Computer Networks: The International Journal of Computer and Telecommunications Networking*, v.38 n.5, p.631-643, 5 April 2002

[44] S. Kim and P. K. Varshney, “An Integrated Adaptive Bandwidth-Management Framework for QoS-Sensitive Multimedia Cellular Networks”, *IEEE transactions on Vehicular Technology*, Vol. 53, No. 3, May 2004, pp835-846

[45] M. Hildebrand, G. Piao, K. David, G. Cristache and F. Fechter: “Efficient Estimation of Transmission Power Applied to the Simulation of the Cell Breathing Effect in CDMA-based Wireless Systems”, *IASTED International Conference "Applied Modelling and Simulation (AMS 2002)*", Cambridge, USA, 4-6 Nov. 2002

[46] M. Hildebrand, K. David, G. Piao, G. Christache, F. Fechter, R. Sigle, A. Warich, “Investigation of Cell Breathing Effect in CDMA-based Cellular Systems”, *IST Summit on Mobile & Wireless Communications*, Aveiro, Portugal, 15 – 18 June 2003

[47] H. Holma, A. Toskala (Editors): “WCDMA for UMTS – Radio Access for Third Generation Mobile Communications”, John Wiley & Sons Ltd, Third Edition, 2004

[48] N. Bambos, Sh. C. Chen, G. J. Pottie, “Channel Access Algorithms with Active Link Protection for Wireless Communication Networks with Power

Control", *IEEE/ACM Transactions on Networking (TON)*, v.8 n.5, p.583-597, Oct. 2000
[49] M. Xiao , N. B. Shroff , E. K. P. Chong, "Distributed Admission Control for Power-controlled Cellular Wireless Systems", *IEEE/ACM Transactions on Networking (TON)*, v.9 n.6, p.790-800, December 2001
[50] J. Zhang , M. Hu , N. B. Shroff, "Bursty Traffic over CDMA: Predictive MAI Temporal Structure, Rate Control and Admission Control", *Computer Networks: The International Journal of Computer and Telecommunications Networking*, v.42 n.6, p.779-795, August 2003
[51] D. Liu , Y. Zhang , S. Hu, "Call Admission Policies Based on Calculated Power Control Setpoints in SIR-based Power-controlled DS-CDMA Cellular Networks", *Wireless Networks*, v.10 n.4, p.473-483, July 2004
[52] J. Ye , X. Shen , J. W. Mark, "Call Admission Control in Wideband CDMA Cellular Networks by Using Fuzzy Logic", *IEEE Transactions on Mobile Computing*, v.4 n.2, p.129-141, March 2005
[53] J. M. Aein, "Power Balancing in Systems Employing Frequency Reuse", *COMSAT Tech. Rev.* Vol. 3 no. 2, Fall 1973
[54] S. Manji and W. Zhuang, "Power Control and Capacity Analysis for a Packetized Indoor Multimedia DS-CDMA Network", *IEEE Transaction on Vehicular Technology*, vol. 49, no. 3, May 2000
[55] D. Kim and F. Adachi, "Capacity Estimation for Overlaid Multiband CDMA Systems with SIR-based Power Control", *IEICE Trans. Commun.*, Vol. E83-B, No. 7, July 2000
[56] J. Zhou et al., "Optimization of Dynamic Allocation of Transmitter Power in a DS-CDMA Cellular System Using Genetic Algorithms", *IEICE Trans. Fundamentals*, Vol. E84-A, No. 10, October 2001
[57] J. Kim et al., "An Analysis of Forward Link Power Allocation and User Capacity for 3GPP System", *IEICE Trans. Commun*., Vol. E85-B, No. 4, April 2002
[58] WiMAX Forum™ Regulatory Working Group, "Initial Certification Profiles and the European regulatory framework – September 2004"
[59] J. Jang, KB Lee, "Transmit Power Adaptation for Multiuser OFDM Systems", *IEEE J. on Selected Areas in Communications*, Vol. 21, No. 2, Feb. 2003, pp171-178
[60] Z Shen, JG Andrews, BL Evans, "Optimal Power Allocation in Multiuser OFDM Systems", *GLOBECOM 2003*
[61] H. Kröner, "Radio Resource Allocation for Data Services in UMTS Networks", *International Journal of Electronics and Communications*, pp55-622, No. 1, 2001
[62] 3rd Generation Partnership Project, *"*Physical Layer Aspects of UTRA High Speed Downlink Packet Access*"*, Version 4.0.0, 3GPP TR 25 848, Release 4, 2001
[63] I. Forkel, H. Klenner and A. Kemper, "High Speed Downlink Packet Access (HSDPA) – Enhanced Data Rates for UMTS Evolution", *Computer Networks* 49 (2005), pp 325-340
[64] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushayana and A. Viterbi, "CDMA/HDR: a Bandwidth-efficient High-speed Wireless Data Service for Nomadic Users", *IEEE Communications magazine*, vol. 38, no. 7, pp. 70-77, 2000

[65] H. Berg, R. Litjens, and J. Laverman, ”HSDPA Flow Level Performance: The Impact of Key System and Traffic Aspects”, COST 290TD(05)007, *IST COST Project 290 Management Committee Meeting*, Colmar, Feb. 3-4, 2005

[66] G. Nair et al, “IEEE 802. 16 Medium Access Control and Service Provisioning”, *Intel® Technology Journal*, Volume 08 Issue 03 Published, 20 August 2004, pp213-228

[67] J. Huang, V. Subramanian, R. Agrawal, and R. Berry, "Downlink Scheduling and Resource Allocation for OFDM Systems" (invited paper), *Conference on Information Sciences and Systems (CISS)*, Princeton University, NJ, USA, March 2006

[68] IEEE 802.16e,

[69] Ed Agis et al., “Global, Interoperable Broadband Wireless Networks: Extending WiMAX Technology to Mobility”, *Intel® Technology Journal*, Volume 08 Issue 03 Published, August 20, 2004, pp173-188

[70] S. Ulukus and G. P. Pollini, “Handover Delay in Cellular Wireless Systems”, *ICC98,* Atlanta, 7-11 June 1998

[71] 3rd Generation Partnership Project, "Interlayer Procedures in Connected Mode", Version 5.3.0, 3GPP TS 25.303, Release 5, 2005

[72] 3rd Generation Partnership Project, “Technical Specification Group Radio Access Network; Radio Resource Management Strategies”, Version 5.3.0, 3GPP TR 25.922, Release 5, 2004

[73] 3rd Generation Partnership Project, "Radio Resource Control (RRC); Protocol Specification", 3GPP TS 25.331, Version 5.14.0, 3GPP TR 25.922, Release 5, 2005

[74] J. X. Qiu, and J. W. Mark, “A Dynamic Load Sharing Algorithm Through Power Control in Cellular CDMA”, *Proc. The 9th IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC’98)*, Boston, MA, 8–11 Sep. 1998

[75] P. Fizpatrick, C. S. Lee, B. Warfield, “Teletraffic Performance of Mobile Radio Networks with Hierarchical Cells and Overflow”, *IEEE J. Selected Areas in Communications*, vol. 15, Issue 8, pp. 1549-1557, Oct. 1997

[76] L. R. Hu, S. S. Rappaport, “Personal Communication Systems Using Multiple Hierarchical Cellular Overlayers”, *IEEE J. Selected Areas in Communications*, vol. 13, Issue 2, pp406-415, Feb. 1995

[77] L. Ortigoza-Guerrero and H. Aghvami, “A Distributed Dynamic Resource Allocation for a Hybrid TDMA/CDMA System”, *IEEE Transactions on Vehicular Technology*, Vol. 47, No. 4, November 1998

[78] C. W. Leong et al., “Call admission Control for Integrated On/Off Voice and Best-Effort Data Services in Mobile Cellular Communications”, *IEEE Transactions on Communications*, pp 778-790, vol. 52, No. 5, May 2004

[79] K. S. Gilhousen et al., “On the Capacity of a Cellular CDMA System”, *IEEE Transactions on Vehicular Technology*, Vol. 40, No. 2, May 1991

[80] D. K. Kim, D. K. Sung, “Capacity Estimation for an SIR-Based Power-Controlled CDMA system Supporting On-Off Traffic”, *IEEE Transaction on Vehicular Technology*, Vol. 49, No. 4, July 2000

[81] L. Wang et al., “Performance Analysis of an Integrated CS/PS Services CDMA System”, *IEEE Transaction on Vehicular Technology*, Vol. 54, No. 4, July 2005

[82] W. Choi and J. Y. Kim, “Forward-Link Capacity of a DS/CDMA System with Mixed Multirate Sources”, *IEEE Transactions on Vehicular Technology*, Vol. 50, No. 3, May 2001

[83] J. Wu, N. Meha and J. Zhang, “A Flexible Log-normal Sum Approximation Method”, *IEEE GLOBECOM 2005*, St. Lious, USA, 28 Nov.- 2 Dec. 2005
[84] L. F. Fenton, “The Sum of Log-normal Probability Distributions in Scatter Transmission Systems”, *IRE Trans. Commun. Syst.*, vol. CS-8, pp. 57-67, 1960
[85] S. Schwartz and Y. Yeh, “On the Distribution Function and Moments of Power Sums with Log-normal Components”, *Bell Syst. Tech. J.*, vol. 61, pp. 1441-1462, 1982
[86] R. Barakat, “Sums of Independent Log-normally Distributed Random Variables”, *J. Opt. Soc., Am.*, vol. 66, pp 211-216, 1976
[87] N. C. Beaulieu and Q. Xie, “An Optimal Log-normal Approximation to Log-normal Sum Distributions”, *IEEE Transactions on Vehicular Technology*, vol. 53, pp. 479-489, 2004
[88] W. Yang, E. Geraniotis, “Admission Policies for Integrated Voice and Data Traffic in CDMA Packet Radio Networks”, *IEEE J. on Selected Areas in Communications*, Vol. 12, No. 4, May 1994
[89] Ch. Wu, Y. Tsai and J. Chang, “A Quality-Based Birth-and-Death Queuing Model for Evaluating the Performance of an Integrated Voice/Data CDMA Cellular System”, *IEEE Transactions on Vehicular Technology*, Vol. 48, No. 1, January 1999
[90] CW Leong, W. Zhuang, Y. Cheng, and L. Wang, “Call Admission Control for Integrated ON/OFF Voice and Best-Effort Data Services in Mobile Cellular Communications”, *IEEE Transactions on Communications,* Vol. 52, No. 5, May 2004
[91] W. S. Jeon, D. G. Jeong, “Call Admission Control for Mobile Multimedia Communications with Traffic Asymmetry between Uplink and Downlink”, *IEEE Transactions on Vehicular Technology*, Vol. 50, No. 1, January 2001
[92] Gudmundson, M., “Correlation Model for Shadow Fading in Mobile Radio Systems,” *Electronics Letters*, vol. 27, No 23, pp.2145-2146, 7 Nov. 1991
[93] T. Halonen, J. Romero, J. Melero, “GSM, GPRS and EDGE performance”, Wiley 2002.
[94] J. Brouet, “Modelling GSM Radio Propagation in MRRM simulator”, *Alcatel Memorandum UTR/L/02/0034*. 18. July 2002.
[95] European Telecommunications Standards Institute, “Radio network planning aspects”, Version 8.3.0, GSM 03.30 / TR 101 362, ETSI, France, 1999
[96] European Telecommunications Standards Institute, “Procedures for the choice of radio transmission technologies of the UMTS”, Version 3.1.0, UMTS 30.03 / TR 101 112, ETSI, France, 1997
[97] G. Piao, K. David, “Multi-standard Radio Resource Management for Integrated Voice and Data Services”, *VTC-2007 Spring*, Dublin, Ireland, 22-25 April 2007, in press
[98] S. Dixit, Y. Guo and Z. Antoniou, “Resource Management and Quality of Service in Third-Generation Wireless Networks”, *IEEE Communications Magazine*, February 2001
[99] R. Yavatkar, D. Pendarakis and R. Guerin, “A Framework for Policy-based Admission Control”, RFC 2753, January 2000, http://www.ietf.org
[100] M. Hildebrand, G. Piao, K. David, R. Sigle, D. Zeller, I. Karla, „Performance Investigation on Multi-standard Radio Resource Management for Packet Switched Services”, *VTC-2004fall*, LA, USA, 26-29 Sep. 2004
[101] M. Hildebrand, K. David, G. Piao, R. Sigle, D. Zeller, I. Karla, “Performance Investigation on Multi-standard Radio Resource Management for Circuit

Switched Services", *IST Summit on Mobile & Wireless Communications*, Lyon, France, 27 – 30 June 2004
[102] http://www.opnet.com
[103] Regulierung Telekommunikation → Frequenzordnung, Regulierung Telekommunikation → Lizenz und Wegerechte, http://www.bundesnetzagentur.de/enid/a89237d21642af1227303cf235e0bbc7, 0/Sachgebiete/Telekommunikation_2v.html, 26.01.2005
[104] G. Piao, K. David, I. Karla, R. Sigle, "Distributed MxRRM", *the 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2006)*, Helsinki, Finland, 11-14 Sep. 2006
[105] G. Piao, K. David, "MxRRM for WiMAX Integrated to GSM and UMTS Heterogeneous Networks ", *VTC-2007 Spring*, Dublin, Ireland, 22-25 April 2007, in press
[106] Hardy Halbauer, "Extension of Hata Model towards 3.5 / 5.8 GHz", *Alcatel ZFZ/A*, 22.9.2004, unpublished
[107] Gabriel Cristache et al., "Aspects for the Integration of Ad Hoc and Cellular Networks", *3rd Scandinavian Workshop on Wireless Ad-hoc Networks*, Stockholm, 6-7 May 2003
[108] M. Litzenburger, H. Bakker, S. Kaminski, K. Keil, "Very Tight Coupling of Wireless LAN and UMTS Networks: A Technical Challenge and an Opportunity for Mobile Operators", *Personal Wireless Communications Conference 2002 (PWC 2002)*, 23-25 Oct. 2002 Singapore
[109] M. Abramowitz and I. Stegun, "Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables", Dover, 9 ed., 1972

Abbreviations

2-D	2 Dimensions
2G	2nd Generation
3-D	3 Dimensions
3G	3rd Generation
3GPP	3rd Generation Partnership Project
AAS	Adaptive Antenna System
Blind_DU	Blind and allocate Data services to UTRAN
BS	Base Station
BSC	Base Station Controller
CDMA	Code Division Multiple Access
C/I	Carrier to Interference ration
COST 231	European Cooperation in the field of Scientific and Technical Research 231
CRRM	Common Radio Resource Management
CS	Circuit Switched
Eb/N0	bit energy and noise density ratio
EDGE	Enhanced Data for GSM Evolution
EIRP	Equivalent Isotropic Radiated Power
EU	European Union
FDD	Frequency Division Duplex
FTP	File Transfer Protocol
GERAN	GSM/EDGE Radio Access Network
GPRS	General Packet Radio Service
GSM	Global System for Mobile communications
HSDPA	High Speed Downlink Packet Access
HS-DSCH	High Speed Downlink Shared Channel
IETF	Internet Engineering Task Force
IMT-2000	International Mobile Telecommunications at 2000MHz

IP	Internet Protocol
IS-CCO	Intersystem Cell Change Order
IS-HO	Intersystem Handover
ISI	Inter-Symbol Interference
IWU	Interworking Unit
LIAM	Load Information Aiding MRRM
MAC	Medium Access Control
MAI	Multi-Access Interference
MCS	Modulation and Coding Scheme
MGF	Moment Generating Function
M/M/N	Markov / Markov / N (servers)
$M / M / \infty$	Markov / Markov / infinite (servers)
MRRM	Multi-Radio Resource Management
MS	Mobile Station
NRT	Non-real-time
NW	Network
OFDM	Orthogonal Frequency Division Multiplex
PCMCIA	Personal Computer Memory Card International Association
PDP	Policy Decision Point
PEP	Policy Enforcement Point
PF	Proportional Fair
PS	Packet Switched
QoS	Quality of Service
R5	Release 5
R99	Release 99
RAN	Radio Access Network
RAT	Radio Access Technology
RNC	Radio Network Controller
RR	Round Robin
RRC	Radio Resource Control
RRM	Radio Resource Management

RT	Real Time
SF	Spreading Factor
SGSN	Serving GPRS Support Node
SIR	Signal Interference Ratio
SNR	Signal Noise Ratio
SRNC-RRC	Serving RNC
SS	Subscriber Station
TCP	Transport Control Protocol
TDD	Time Division Duplex
TDM	Time Division Multiplex
TDMA	Time Division Multiplex Access
TTI	Transmission Time Interval
TX	Transmission
UE	User Equipment
UTRAN	UMTS Terrestrial Radio Access Network
UMTS	Universal Mobile Telecommunications System
VGDU	Voice allocated to GSM and Data allocated to UMTS
VoIP	Voice over IP
VUDG	Voice allocated to UMTS and Data allocated to GSM
WCDMA	Wide-band CDMA
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Access Network

Publications

[1] G. Piao, K. David, “Multi-standard Radio Resource Management for Integrated Voice and Data Services”, *VTC-2007 Spring*, Dublin, Ireland, 22-25 April 2007, in press

[2] G. Piao, K. David, “MxRRM for WiMAX Integrated to GSM and UMTS Heterogeneous Networks ”, *VTC-2007 Spring*, Dublin, Ireland, 22-25 April 2007, in press

[3] G. Piao, K. David, I. Karla, R. Sigle, “Distributed MxRRM”, *the 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2006)*, Helsinki, Finland, 11-14 Sep. 2006

[4] M. Hildebrand, G. Piao, K. David, R. Sigle, D. Zeller, I. Karla, „Performance Investigation on Multi-standard Radio Resource Management for Packet Switched Services”, *VTC-2004fall*, LA, USA, 26-29 Sep. 2004

[5] M. Hildebrand, K. David, G. Piao, R. Sigle, D. Zeller, I. Karla, “Performance Investigation on Multi-standard Radio Resource Management for Circuit Switched Services”, *IST Summit on Mobile & Wireless Communications (IST 2004)*, Lyon, France, 27 – 30 June 2004

[6] M. Hildebrand, K. David, G. Piao, G. Christache, F. Fechter, R. Sigle, A. Warich, “Investigation of Cell Breathing Effect in CDMA-based Cellular Systems”, *IST Summit on Mobile & Wireless Communications (IST 2003)*, Aveiro, Portugal, 15 – 18 June 2003

[7] M. Hildebrand, G. Piao, K. David, G. Cristache and F. Fechter: “Efficient Estimation of Transmission Power Applied to the Simulation of the Cell Breathing Effect in CDMA-based Wireless Systems”, *IASTED International Conference "Applied Modelling and Simulation (AMS 2002)*", Cambridge, USA, 4-6 Nov. 2002

[8] K. David, G. Piao, I. Karla and R. Sigle, “Method of Sharing Radio Resources, Radio Resource Management Module, and Radio Communication Network”, Alcatel patent application to the European Patent Office, *European Patent No. 06291215.9-2412*, 25.07.06