COMMON RADIO RESOURCE MANAGEMENT: JOINT CALL ADMISSION CONTROL AND TRAFFIC OFFLOADING METHOD

KHUZAIRI MOHD ZAINI

UNIVERSITI SAINS MALAYSIA

2018

COMMON RADIO RESOURCE MANAGEMENT: JOINT CALL ADMISSION CONTROL AND TRAFFIC OFFLOADING METHOD

by

KHUZAIRI MOHD ZAINI

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosphy

July 2018

ACKNOWLEDGEMENT

I owe my deepest gratitude to my supervisors, Dr. Azizul Rahman and Professor Dr. Azman Samsudin. This work would not have been possible unless the support and guidance of them. Dr. Azizul enthusiasm and his insightful advices on this research, and continued patience during completion of this work have greatly improved my research skills.

I also would like to thank Ministry of Higher Education Malaysia for their scholarship support through my studies and to my employer Universiti Utara Malaysia for having trust in me to complete this study. My sincere appreciation goes to the office and administration staff of School of Computer Sciences Universiti Sains Malaysia for ensuring the countless forms and paperwork were correctly completed and filed.

I am deeply indebted to my friends, Dr. Zheng Shi from University of Macau, Dr. Azizah and Dr. Norhayati of Universiti Utara Malaysia, and Dr. Zainor Yahya from Universiti Malaysia Perlis, for their help with my research, and more for their great friendship and encouragement during my studies.

Finally, my heartiest gratitude goes to my family, to my mother and mother-in-law who always have faith in me and pray for my success, to my beloved wife Saridatun Nur for her understanding, support, and love, and last but not least to all my children Izzati Aini, Izzati Fathini, Izzati Akeela and Hudhayfah for being so sweet and loving.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3GPP	3rd Generation Partnership Project
AHP	Analytic Hierarchy Process
AMPS	Advanced Mobile Phone System
ANDSF	Access Network and Discovery Function
AP	Access Point
BLTI	Bell Lab Traffic Index
CDf	Cumulative Distribution Function
CR	Cognitive Radio
CRN	Cognitive Radio Network
CRRM	Common Radio Resource Management
CTMC	Continuous Time Markov Chain
DSA	Dynamic Spectrum Access
DSL	Digital Subscriber Line
DSSS	Dynamic Sequence Spread Spectrum
eNB	E-UTRAN Node Bs
ETSI	European Telecommunications Standards Institute
E-UTRAN	Evolved Universal Terrestrial Radio Access Network
FCFS	First Come First Serve
FDMA	Frequency Division Multiple Access

FM Frequency Modulation

FSK Frequency Shift Keying

GGSN Gateway GPRS Support Node

GMSK Gaussian Minimum Shift Keying

GPRS General Packet Radio Service

GRA Grey Relational Analysis

GSM Global System for Mobile Communication

GTP GPRS Tunneling Protocol

GW Gateway

HetNet Heterogeneous Network

HLR Home Location Registrar

HSDPA High Speed Downlink Packet Access

HSPA High Speed Packet Access

HSUPA High Speed Uplink Packet Access

HWN Heterogeneous Wireless Networks

IP Internet Protocol

ITU International Telecommunication Union

JCAC Joint Call Admission Control

JRRM Joint Radio Resource Management

LTE Long Term Evolution

MAC Media Access Control

MADM	Multi-Attribute Decision Making
MEW	Multiplicative Exponential Weighting
MIMO	Multiple-Input Multiple-Output
MME	Mobility Management Entity
MN	Mobile Node
NRM	Network Reconfiguration Manager
NRT	Non Real-Time Application
OECD	Organisation for Economic Co-operation and Development
PDf	Probability Distributive Function
QoS	Quality of Services
RAN	Radio Access Network
RATs	Radio Access Technologies
RNC	Radio Network Controller
RT	Real-Time Application
SAE	System Architecture Evolution
SAW	Simple Additive Weighting
SCTP	Stream Control Transmission Protocol
SDR	Software Defined Radio
SGSN	Serving GPRS Support Node
TDMA	Time Division Multiple Access
TOPSIS	Techniques for Order Preferences by Similarity to an Ideal Solution

TRM Terminal Reconfiguration Manager

UMTS Universal Mobile Telecommunications System

UTRAN Universal Terrestrial Radio Access Network

VHO Vertical Handover

VLR Visiting Location Registrar

WCDMA Wideband Code Division Multiple Access

WiFi Wireless Fidelity

WLAN Wireless Local Area Network

WWW World-Wide Web

PENGURUSAN SUMBER RADIO BERSAMA: KAEDAH KAWALAN KEMASUKAN PANGGILAN BERSAMA DAN PEMINDAHAN TRAFIK

ABSTRAK

Konsep Pengurusan Sumber Radio Bersama (CRRM) telah dicadangkan oleh 3GPP untuk mengurus sumber radio yang terkumpul secara efisien untuk setiap teknologi capaian radio dalam pelbagai rangkaian tanpa wayar (HWN). Cabaran utama CRRM untuk HWN adalah untuk mencari sambungan yang terbaik bagi perkhidmatan yang diminta, membolehkan pemindahan panggilan daripada satu teknologi radio ke teknologi radio yg lain dengan lancar dan menggunakan kesemua sumber radio yang sedia ada semaksima mungkin. Oleh itu, tesis ini mencadangkan kawalan panggilan panggilan bersama (JCAC) dan algoritma pemindahan trafik dalam CRRM untuk rangkaian tanpa wayar dan mudah alih bersepadu. Terdapat tiga sumbangan utama di dalam penyelidikan ini. Pertama, model mobiliti pengguna dalam rangkaian tanpa wayar berlapis adalah dicadangkan. Rumus masa kediaman di WLAN diperoleh dan kebarangkalian keluar dari kawasan liputan dan masa tinggalnya dikira. Seterusnya, persamaan *vertical handoff* (VHO) dihasilkan. Prestasi dinilai secara matematik menggunakan teori masa berterusan Markov Chain (CTMC). Kedua, JCAC terdiri daripada pemilihan rangkaian permulaan dan VHO juga dicadangkan. Pergerakan pengguna, tempoh penggunaan, ciri-ciri perkhidmatan dan beban sistem telah dimasukkan ke dalam parameter membuat keputusan JCAC dan dinilai dengan menggunakan tiga dimensi Markov Chains. Ketiga, algoritma pemunggahan trafik diperkenalkan dalam CRRM yang dicadangkan. Proses hierarki analitis (AHP) digunakan untuk penugasan pemberat didalam kriteria membuat keputusan dan fungsi kos digunakan untuk pemilihan pengguna dalam algoritma pemindahan. Keseimbangan pertukaran di antara metrik prestasi

sistem yang berbeza diperiksa dengan skim CRRM yang dicadangkan dengan dan tanpa algoritma pemindahan. Hasil yang diperolehi menunjukkan bahawa skim JCAC yang dicadangkan mengurangkan kebarangkalian kadar sekatan panggilan yang baru sebanyak 24% dan pengurangan sebanyak 95% bagi VHO, serta meningkatkan 11% keluaran sistem keseluruhan. Selain itu, skim yang dicadangkan juga mengurangkan 2.5% kos perkhidmatan dan mengurangkan perbezaan kadar penggunaan rangkaian keseluruhan sistem sebanyak 7%. CRRM yang dicadangkan dengan skema pemindahan dapat menerima lebih ramai pengguna masuk ke sistem dan meningkatkan 6% keupayaan keseluruhan. Akhirnya, penemuan CRRM yang dicadangkan yang terdiri daripada algoritma JCAC dan pemindahan telah menunjukkan kesan yang bermunafaat terhadap prestasi sistem, dan memberikan garis panduan dalam rekabentuk untuk penyelidikan masa depan dalam domain ini.

COMMON RADIO RESOURCE MANAGEMENT: JOINT CALL ADMISSION CONTROL AND TRAFFIC OFFLOADING METHOD

ABSTRACT

The concept of Common Radio Resource Management (CRRM) has been proposed by 3GPP in order to efficiently manage the common pool of radio resources that are available for each of the existing radio access technologies in the heterogeneous wireless networks (HWNs). The main challenge of CRRM for HWNs is to search the best connection for the demanded services, enabling calls transference from one interface to another seamlessly and utilizing the availability of all radio resources. Thus, this thesis proposes a joint call admission control (JCAC) and traffic offloading algorithms in the CRRM for an integrated cellular and mobile network. The contributions of this work are threefold. First, a user mobility model in the wireless overlay network is proposed. The equation of dwelling time in the WLAN is derived and the probability of moving out from the coverage area and its sojourn time is calculated. Next, the vertical handoff (VHO) rate equation is produced. The performance is assessed analytically using the theory of continuous time Markov Chain (CTMC). Second, JCAC comprises of initial network selection and VHO is proposed. User mobility, usage duration, service characteristics and system load have been incorporated in the JCAC decision making parameters and evaluated by using three dimension Markov Chains. Third, traffic offloading algorithm is introduced in the proposed CRRM. An analytic hierarchy process (AHP) is applied for weight assignment to the decision making criteria and the cost function is used for the user's selection in the offloading algorithm. The tradeoff between different system performance metrics are examined between proposed CRRM scheme with and without offloading. Results obtained show that the proposed JCAC

scheme reduces the new and VHO calls blocking probabilities by 24% and 95% respectively, and increases 11% of the overall system throughput. Furthermore, the proposed scheme also minimizes 2.5% of the service cost and improves 7% of the variance between network utilizations of the entire system. The proposed CRRM with offloading scheme can admit more users into the system and improve the overall throughput by 6%. Finally, the findings of the proposed CRRM which consist of JCAC and offloading algorithms have shown the impact on different attributes of the system performance, providing insights and design guidelines for future research in this domain.

CHAPTER 1

INTRODUCTION

This chapter begins by explaining the main drivers for this research in terms of wireless evolution and the importance of next generation heterogeneous wireless network research. After that the problem statement is formulated, followed by the research objectives, contributions, and the dissertation organization.

1.1 Introduction

The emerging wireless technologies has been accompanied not only by basic voice services but also by growth and development of a wide variety of applications, ranging from the simple text-based utilities to the most bandwidth intensive such as video conferencing and streaming. The projected global mobile data traffic will increase nearly 10 times between 2014 and 2019, and nearly three quarters of the mobile data traffic will be video traffic by 2019 [3, 4]. According to the Bell Labs Traffic Index (BLTI), this excessive growth is driven by several technological and usage factors [5].

Since the introduction of the first generation or 1G cellular technology, it has now gone through tremendous enhancements till the birth of 4G. The evolution of wireless or cellular technology could be divided into three phases [6]. In the first phase, the cellular technology was mainly used for basic communication such as voice call and short messaging. The prime objective was to have mobility where people can communicate anytime and anywhere. The rise in Internet services influenced the second phase of wireless technology. People started to have Internet connection while on the move where they can read and reply email instantly, and browse their favorite websites. Meanwhile, in the third wave of wireless evolution involved changes of architectural aspects by integrating different networks under common operations.

According to the data by OECD, wireless Internet subscriptions had exceeded half a billion by the end of 2010, an increment of 10% in just six months among the OECD countries [7]. This growth was partly fuelled by inexpensive, flat-rate mobile data plans [8]. Due to the overload of wireless data traffic, Verizon, on the other hand, planned to remove the unlimited mobile data tariff by replacing with per customer usage plan [9].

Internet applications also have gone through tremendous change. The World Wide Web (WWW), one of the most widely used applications in the Internet has moved to Web 2.0 which is more user rich content. Social networking sites, blogs, wikis, and video sharing sites such as YouTube are examples of Web 2.0. Although Web 2.0 has no impact on its specification, it has substantially influenced network infrastructure [10]. It changes the application architecture from client server to a hybrid client/server and peer-to-peer system. Furthermore, Web 2.0 is not only famous for social types of application, it has now becoming a tool for conducting business in the Internet [11]

Increasing traffic on wireless networks and higher bandwidth consumption may reduce network performance, motivating operators to invest in increased network capacity and advanced technologies to allow for faster speeds and a higher level of simultaneous access. As such, static usage of services from a single operator may no longer be capable of satisfying user demands in the near future, although in reality, there exist variety of wireless technologies that are accessible by them. Several options to eliminate the above problems have been identified [12], which include

a. Installing new base stations. Generally, devices within the same network technologies share the same spectrum. Therefore deploying new base stations that comprise macrocells/microcells, picocells, femtocells and relays, may not require new spectrum licences [13].

b. Upgrading networks. An example is the migration from current network such as LTE to Wimax. This is considered as multi-RAT network components [13].

c. Offloading. Since there is a partially overlapped area between cellular and WLAN, some of the traffic in the cellular network could be switched to WLAN.

Setting-up new network infrastructures involve planning, configuration, optimization, dimensioning, testing and maintenance, which are labor and cost intensive [14]. Therefore, it is more preferable to utilize WiFi networks for offloading the burden of cellular network [15], which will then lead to maximizing profits.

1.2 Traffic Offloading

As mobile traffic volume continues to expand, network operators are seeking viable mechanisms to augment their network capacity to meet the ever increasing demand. These mechanisms include upgrading the access networks to LTE/4G with better spectral efficiency, improving backhaul capacity, and offloading traffic from mobile access to alternative channels such as WiFi and femtocells [16]. Being one of the techniques in network management, traffic offloading has been adopted by mobile network operators to alleviate pressure on their networks overloaded by the surge of mobile traffic. For instance, owing to the advantage of low-cost and technical maturity, WiFi offloading has become a popular choice. In was predicted by 2016, 63 percent of all traffic from mobile-connected devices will be offloaded through WiFi or femtocells [3]. The key enabler for mobile traffic offloading is the fast development of wireless communication technologies.

Nowadays, smartphones can support a rich set of communication technologies such as LTE, HSPA, WiFi, and Bluetooth. On the network side, mobile operators are upgrading their infrastructure to LTE and LTE-advanced for enhanced performance. WiFi and femtocells are gaining popularity in metropolitan areas by offering diverse and convenient wireless access. Following the trend, 3GPP has put considerable effort in standardizing IP-based offloading solutions for a tight integration to 3GPP network architecture [17]. Research communities also showed convincing results on the effectiveness of mobile traffic offloading through experimental studies [18, 19, 20].

1.3 Heterogeneous Wireless Network Challenges and Development

In order to allow more freedom of choices to satisfy user demands, next generation heterogenous wireless network (HWN), as illustrated in Figure1.1, has been proposed for future development of wireless architecture. HWN allows possible integration of multiple wireless networks by utilizing dynamic selection of networks to satisfy user quality of services and ensure always best connected services. The HWN concept proposes a flexible and open architecture for a large variety of different wireless access technologies, applications, and services with different QoS demands. In addition, the evolution of wireless technologies have not only focused on increasing data rates as shown in Figure 1.2, but incorporated into the overall systems to operate intelligently. This concept is known as 5G where it allows seamless wireless world interconnection, convergence and cooperation while serving at very high data rates [1].

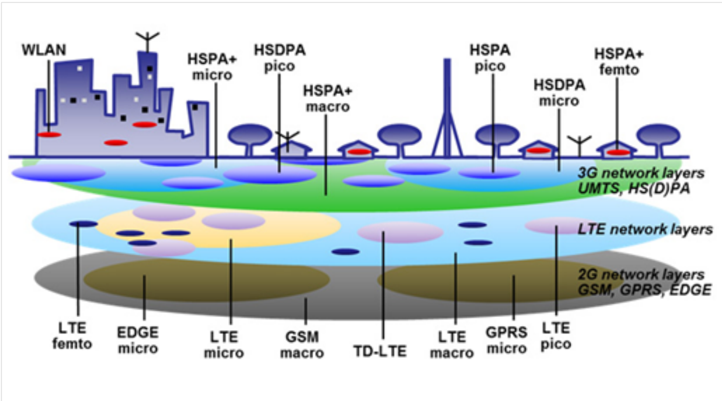

Figure 1.1: Heterogenous Wireless Network Architecture

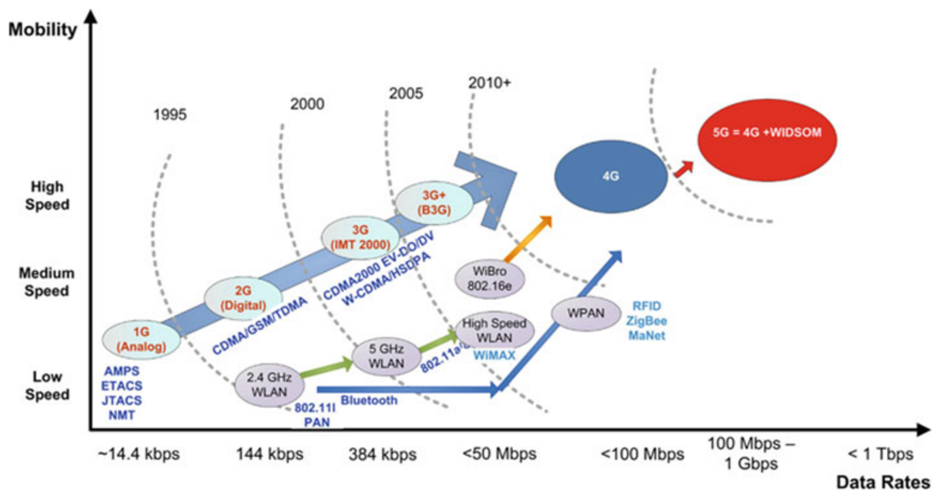

Figure 1.2: Evolution of Wireless Technologies [1]

There are two different ways of integrating heterogeneous wireless networks known as tightly coupled and loosely coupled interworking [2], as depicted in Figure 1.3. The integration can be between cellular and cellular networks, or cellular and WiFi networks. A tightly coupled system allows the integration with other networks through interfacing directly to their system. Therefore, these two networks could use the same authentication, mobility, and the billing system. In a loosely couple system, the integration of the two networks are done at the Internet level. Thus, different mechanisms and protocols are used for authentication, mobility and billing. However, Piamrat et.al. [21] addressed the concern in terms of alleviating load-balancing, billing and authentication, as well as on how the real implementation could be done across different network operators and users.

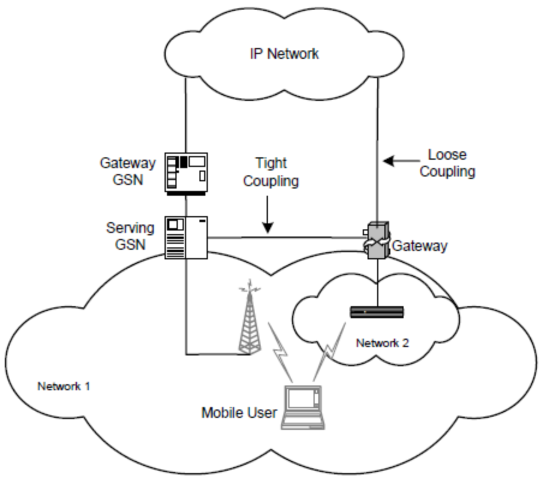

Figure 1.3: Interworking of Heterogeneous Wireless Networks [2]

The successful implementation of HWN has also been assisted by the introduction of dynamic spectrum access (DSA). Solving the problem of scarcity and inefficient usage of spectrum is the main motivation of DSA. The key enabling technology of DSA is cognitive radio CR, a software defined radio (SDR) that is reconfigurable through software embedded in microprocessors and able to sense and adapt to the available frequency. The idea of cognitive radio was first promoted by Mitola who defined cognitive radio as a radio that has the ability to understand its surrounding and adapts intelligently [22]. The main idea of cognitive radio is not only related with opportunistically using the unused spectrum efficiently while meeting the heterogeneous and dynamic requirements of the users, but it also assists the user with its activity. The use of CR as the implementation tool has also been considered in Long Term Evolution (LTE) [23, 24] and in 5G [1].

Another important aspect of future wireless networks is the autonomic operations that move from legacy centralized systems to more dynamic and distributed systems [25]. Through a central coordinator, a network could decide the most suitable resources for users. The proposed resources may or may not adhere to the user's requirements. In case the network could not meet the user's objectives, it will suggest the nearest accepted level of resources the network can provide to the users. On the other hand, future wireless networks should embrace the concept of context-awareness, as discussed in [26]. Context-aware mobile and wireless networking (CAMoWiN) should be seen not as a set of numerical values only as the traditional understanding in pervasive computing, but as a flow of information to infer knowledge. Furthermore, CAMNoWiN will be the base for autonomic decision making, allowing the entire system to be self-optimized where the network, user and application shall be able to make decisions by themselves.

1.4 Radio Resource Management Research Challenges

Since there will be different technologies with various characteristics in HWN, a common platform to manage the resources is required, known as Common Radio Resource Management (CRRM). The next generation heterogeneous wireless networks' requirements have force the Common Radio Resource Management (CRRM) or Joint Radio Resource Management (JRRM) to play a more important role than before. The challenges do not only on refer to the technical efficiency but also whether different constraints given by users' preferences are met, and this will have adverse impact on the entire system. The ability to change parameters adaptively and automatically without operator intervention and sharing resources across other operators are amongst the new challenges to be addressed by CRRM. CRRM has been proposed to cater for multiple choices of wireless access network and the challenge would be how to achieve the overall optimum trade-off between resource utilization and quality of service in heterogenous wireless networks [27].

The introduction of HWN has allowed users to have more choices of which network to be connected to that suits their preferences. Similarly, network operators also, have the option to allocate which access network to the users in order to achieve network operators' objectives. Thus, striking the balance of satisfying users and network operators is a paramount task.

Accepting users in the network by considering a single criterion from the network only, is no longer applicable in HWN. In order to support the efficient management of spectrum resources, ETSI through its working group 3 (WG3) has proposed to consider user behavior related requirements such as personalization, context-awareness and pervasive computing [28]. IEEE P1900.4 [29] suggested a set of parameters for radio resource optimization including the user's aspects. However, there are too many input which may result in complexity of the system. Thus, critical parameters need to be identified to lessen too much information exchange.

According to [30], adapting system behavior in accordance to network operator policies, service level agreements, user preferences and location-awareness [31] yields significant improvements in end-to-end performance and resource utilization. As such more parameters need to be taken into consideration instead of only traffic pattern in designing a better framework of resource management for cognitive wireless networks. They proposed the concept of context-enhanced heterogenous access management and from the preliminary results, it showed an improvement of network performance.

1.4.1 Seamless Terminal Mobility

In HWN, mobile users should be able to connect to services at anytime and anywhere. Therefore, seamless terminal mobility needs to be supported, where users will be able to roam from one network to another without service disruption. Changing from one network to another with the same technology is called horizontal handover, while changing to different technology is known as vertical handover (VHO). In a legacy system, VHO is an ongoing session attached to a specific network during the intial network selection, but needs to change the attachment due to its mobility. In HWN, however, VHO may occur due to degrading link quality, system load or any other aspect deem appropriate in the view of the administrator. Triggering VHO not because of mobility is known as traffic offloading [32]. Therefore, identifying suitable time of action and mechanism to invoke traffic offloading without affecting system performance need to be addressed.

1.4.2 Challenges

It is important that the CRRM proposed is able to jointly managed resources. For instance, if most of users prefer one radio access technology over another, this would result in overloading the selected network and underutilizing the other. Thus, the preferred network will face higher blocking probabilities and consequently degrade the performance of the system. Furthermore, when selecting more users to a smaller coverage area such as WLAN, there might be too much handover when the users are moving out from overlapped areas.

The newly designed CRRM has to minimize false rejection and false admissions as envisioned for future wireless networks [33]. False rejections occur when the system blocks a session although there is available resource in another network to meet the session's requirement. False admissions happen when the system accepts a session although it cannot meet the session's objective.

Depending on the degree of CRRM interaction, functions that appear in the CRRM may consist of admission control, VHO, congestion control, scheduling or power control. In the context of this research, admission control, VHO, and congestion control functionalities were considered, while power control and scheduling were located in the specific technology radio resource controller. Therefore, a holistic evaluation of the proposed CRRM need to show on how the interplay between these functions affect the performance of the system.

1.5 Problem Statement

In the existing technical literature, many related studies on CRRM were reported in a survey paper by [21, 34, 35]. Despite this effort, the majority of existing works did not consider mobility in the evaluation, although it is an important parameter in the wireless network selection procedure [36, 37]. Failing to infuse mobility in the CRRM will result in too many handoffs in a heterogeneous network, and also increase the imbalance of system utilization. Several works took the mobility calculation by adapting a fluid mobility model in their CRRM proposals [38, 39, 40, 41, 42, 43]. Literatures [40], [41] and [42] classified user mobility as either vehicular or non-vehicular based on a preset speed threshold (V_{th}). If the mobile node (MN) speed is less than the speed threshold, there is higher probability that the user will be allocated to WLAN resources. However, speed alone may not accurately consider vertical handoff (VHO) to occur. For example, there are cases where the user is traveling more than the threshold speed but able to complete the task prior to moving out from the overlapped coverage. On the other hand, there are users who travel less than the speed threshold but may use resources a longer time even after leaving the overlapped area. This may cause resource wastage and unnecessary VHO for both scenarios.

What has been neglected so far in the current CRRM solutions, is to fully exploit the advantage of an overlay network. Current works of CRRM will block a new call request from a MN located in the non-overlapping area if there is insufficient the radio access resources (e.g.cellular), and this also happens to handoff calls. In an overlay network, some of the users located in the overlapping area, who are accessing the cellular radio access technology (RAT), can be offloaded to the WLAN system, thus this will grant a new or handoff request to the cellular resource. The existing offloading algorithms focused mainly on the end user or application requirements instead of the system performance. Furthermore, the current offloading approaches have not examined the interaction between the network selection and offloading functions in the CRRM.

1.6 Research Objectives

The ultimate aim of this research was to produce a CRRM that efficiently utilize network resources and maximize network operators' profit, while meeting the diverse user satisfaction and services QoS requirements. Specifically, the proposed CRRM should reduce the new call blocking probability, avoid false call rejections, and minimize the service cost of network operation. Noting that the user's behavior can be predicted with high accuracy using historical information [44, 45], the proposed CRRM combines the user mobility pattern and application usage duration together with service characteristics and system load in the decision making criteria. Below are the objectives:

a. To develop a user mobility analytical model that improve throughput and reduces vertical handoff calls blocking probability in HWN.

b. To develop a joint call admission control (JCAC) that takes into consideration user mobility, application usage duration, system load, and service characteristics which increase the throughput, reduce the blocking probabilities and lowering the service cost.

c. To develop a CRRM framework which includes the JCAC and traffic offloading features that can improve the overall system performance and increase satisfaction level.

1.7 Contributions of the Thesis

The CRRM scheme presented in this dissertation consist of three functions: Initial Network selection, VHO network selection and traffic offloading. Therefore, the key contributions of this research include:

1. A detailed mathematical model of mobility was derived while considering uniformly distributed location of users and directions. The proposed method uses a geometrical model for predicting the dwell time within WLAN and the call holding time. Velocity, cell size, and mobility direction are the main parameters that determine the dwell time. Then, the probability distribution function of dwelling time is obtained and compared with the call holding time, in order to calculate the probability of moving out from the coverage area. In order to ensure the validity of the model, the analytical results were validated with Monte-Carlo simulation and showed a good match between the analytical and simulation models. The performance of the proposed mobility model in WLAN was quantitatively evaluated by means of continous time markov chain (CTMC). The performance metrics measured are throughput, blocking probability and average handoff rate.

2. A mathematical model for the proposed policy-based CRRM algorithm was developed by using mobility, application usage, load-balance, and type of services as the main criteria. The system was modeled based on 3D Markov chain and the performance measured included new call blocking, VHO call blocking, through-

put and service cost.

3. A traffic offloading mechanism was embedded in the proposed CRRM to increase the number of accepted users to HWN and thus improve the profit of network operator. In order to avoid frequent handover and high signaling overhead, the proposed traffic offloading mechanism will only be invoked when there is insufficient resource during new calls or handover calls. An analyical hierarchy process (AHP) was used to assign weightage to each parameter of decision making and a cost function method was adopted to select the most suitable candidate of users to be migrated. Furthermore, the evaluation of satisfaction level was introduced by comparing the user's selection method between cost function and random approach. The metrics of satisfaction level is derived from the differences between the expected and available value of bandwidth, velocity and type of services. Finally, an interaction between offloading and network selection module and its impact to the CRRM performance was investigated. A discrete-event simulator was developed in JAVA programming language tool to test the newly proposed CRRM with offloading and without offloading function.

1.8 Organization of the Dissertation

This dissertation is organized in seven chapters as described in the following:

Chapter One provides an overview on the thesis as a whole. It outlines the issues related to the research and highlights the importance of the research. This chapter forms the problem statement of the research and address the challenges in doing the research. Research objectives and contributions are also stated in this chapter.

Chapter Two consists of a literature review that examined background materials on CRRM, which defines the general framework for this research. It touches the development of wireless technologies and standard organization specification on supporting the idea of performing multi-RAT management and traffic offloading. The chapter reviews the existing related works in the area of policy-based and non-policy based decision making approaches.

Chapter Three addresses the methodology used in carrying out the research. The chapter begins by introducing the Design Research Method (DRM), followed by explaining the proposed CRRM.

Chapter Four introduces the derivation of user mobility model as the first objective outlined before. The chapter first discusses on how the probability distribution function (PDf) and cumulative distribution function (CDf) of dwelling time is derived. Once the probability of moving out and the dwell time are obtained, a VHO rate can be derived. The results and analysis are discussed briefly in their dedicated section.

Chapter Five describes the proposed CRRM by applying the mobility model developed in the previous chapter. Prediction of moving out from hotspot area, together with type of service and the load-balance mechanism are incorporated in the proposed CRRM decision making element. Furthermore, the newly VHO rate is also applied in calculating the transition rate from the inner to outer areas. Based on the 3D Markov chain steady state distribution, performance metrics such as blocking probability, throughput and service cost are analyzed.

Chapter Six explains the proposed CRRM with offloading capability. It starts with a brief discussion about the proposed CRRM. Then, it discusses on how to apply AHP for assigning the appropriate weightage for each decision making items. All the weightage that have been derived are assigned to calculate the cost function in order to select the most suitable user to be migrated to another network. This chapter shows the evaluation of satisfaction level and finally evaluate the performance between the proposed CRRM with and without the offloading scheme.

Chapter Seven discusses and concludes the research. This chapter recalls the developed mobility model and CRRM, and highlights the contribution made throughout the research. Finally, the dissertation ends with suggestions for future works for any researchers wanting to embark on this line of research.

CHAPTER 2

LITERATURE REVIEW

This chapter begins by discussing the fundamentals of cellular and WLAN, followed by explaining the concept of next generation heterogeneous wireless network and the standard organizations involved. Next, this chapter delves into the details of CRRM and evaluates several network selection strategies available in the literature. Finally, the findings of the research gap, are summarized.

2.1 Cellular Technology Evolution

Since the cellular system was created, a constant evolution in the device architecture, system networks and services has been observed. This evolution was always driven by the need of better system capacities than the actual system can offer. The following subsections give a brief description of this evolution path.

2.1.1 First Generation

In the late 1970s, the first generation was designed for voice communication using analog signals. The AMPS was one of the leading analog cellular systems in North America which use FDMA to separate channels in a link. The system used two separate analog channels, one for forward (base station to mobile station) communication and one for reverse (mobile station to base station) communication. The band between 824 and 849 MHz carried the reverse communication; the band between 869 and 894 MHz carried forward communication (see Figure 2.1). Each band was divided into

832 channels. However, two providers can share an area, which means 416 channels in each cell for each provider. Out of these 416, 21 channels were used for control, which leaves 395 channels. AMPS has a frequency reuse factor of 7; this means only one-seventh of these 395 traffic channels are actually available in a cell. For transmission, AMPS used Frequency Modulation (FM) and Frequency Shift Keying (FSK) for modulation. Voice channels were modulated using FM, and control channels used FSK to create 30 kHz analog signals [46].

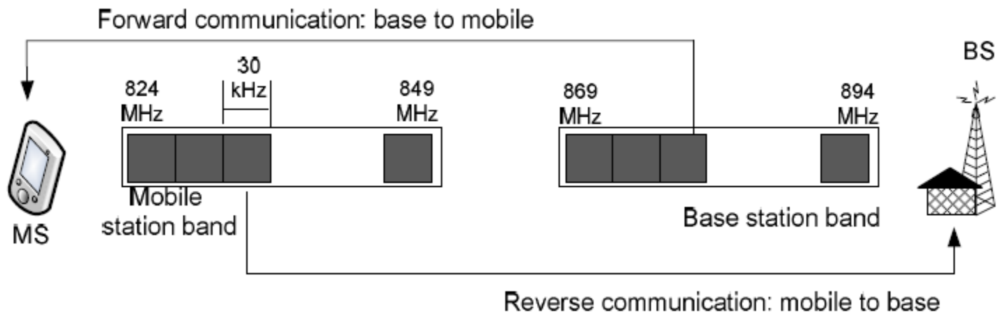

Figure 2.1: AMPS Cellular System

2.1.2 Second Generation

To provide higher-quality on mobile voice communications, the second generation of the cellular phone network was developed. While the first generation was designed for analog voice communication, the second generation was mainly designed for digitized voice. The GSM was an European standard that was developed in the early 1980s to provide a common second-generation technology for Europe. Although GSM was created as a pan-European project, it quickly attracted worldwide interest and was deployed in a great number of countries outside Europe [47]. The aim was to replace a number of incompatible first-generation technologies. GSM used two bands for duplex communication. Each band was 25 MHz in width, shifted toward 900 MHz, and divided into 124 channels of 200 kHz, separated by guard bands. Each 270.8 kbit/s digital channel modulates a carrier using Gaussian Minimum Shift Keying (GMSK) (a form of FSK used mainlyin European systems); the result was a 200-kHz analog signal. Finally, 124 analog channels of 200 kHz were combined using FDMA [46].

2.1.3 From 2G to 3G Evolution

Evolution is a common term used in the context of cellular networks. The evolution from Second Generation (2G) to 3G network means the introduction of data transportation over cellular system. The development in this area started during the second half of the 1990s, with General Packet Radio Service (GPRS) introduced in GSM cellular technology and other developments, such as the Japanese standard. These technologies were often known as 2.5 Generation (2.5G) and considered an evolution of GSM [48]. The success of iMode, the Japanese data network, gave a very clear alert of the great potential for applications over packet data in mobile systems, in spite of, the fairly low data rates supported at the time. With the advent of 3G and the enhancements on the radio interface of UTRAN, a new range of services came, that actually were only hinted at with 2G and 2.5G as shown Figure 2.2.

From GSM to GPRS, some modifications were made in terms of architecture, such as the introduction of the Serving GPRS Support Node (SGSN) and the Gateway GPRS Support Node (GGSN) [49]. These modules are responsible for packet traffic compatibility and allow MSs to use the packet switching network. In order to allow the UTRAN to use the network through GPRS network, appropriate interfaces were developed. A Few modifications at the hardware level were made, only, or in the majority of cases, software adequacy was enough to make the system versatile. It was the way found to give compatibility to the system, while at the same time, reduced the financial impact. Otherwise, it would have been too expensive if the 3G system had been started from the scratch as shown Figure 2.2 [49].

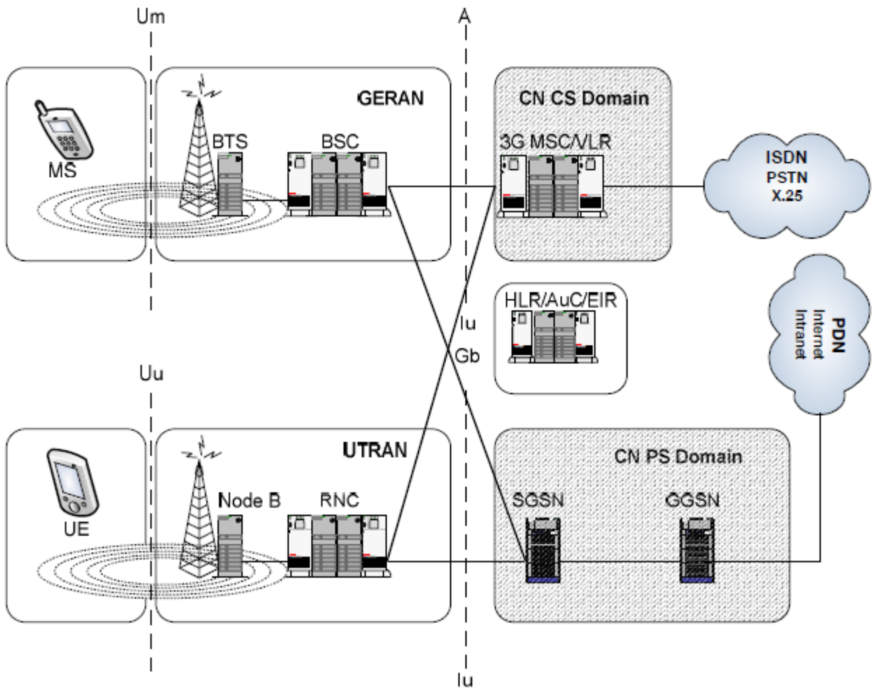

Figure 2.2: GSM/GPRS/EDGE/UMTS General Architecture

2.1.4 UMTS to HSPA Evolution

3GPP developed the architecture of UMTS towards standardization of a new system capable of increasing the speed of the traffic data with QoS. These investigations converged to a new technology, designated by High-Speed Packet Access (HSPA). The normalization of HSPA occurred in two stages, in the downlink direction (from the network to the user) and the uplink direction (from the user to the network). In the downlink, the so-called High-Speed Downlink Packet Access (HSDPA) was standardized as part of 3GPP Release 5. In the uplink, High-Speed Uplink Packet Access (HSUPA) was part of the 3GPP Release 6, where enhanced uplink packet data support was introduced. The 3GPP Release 6 also brought efficient support for broadcast services into WCDMA through the introduction of the MultiMedia Broadcast/Multicast Service (MBMS), suitable for applications like mobile TV and multimedia contents. The HSDPA can reach an initial peak rate of 1.8 Mbit/s in the terminals, and can potentially reach more than 10 Mbit/s. In HSUPA, the peak rate at an early stage can vary from 1 to 2 Mbit/s. It was expected that, in a second phase, the data rate can vary between 3 and 5 Mbit/s.

In Release 7, with the introduction of Multiple-input Multiple-output (MIMO), HSDPA and HSUPA can reach 28 and 11 Mbit/s respectively, (Table 2.1 [5]). This evolution marked a new phase of HSPA known as HSPA+. The HSPA was developed over the architecture of the WCDMA network, on the same carrier. Thus, the HSPA and WCDMA can share all elements of core and radio network, including the BS, the Radio Network Controller (RNC), SGSN and GGSN. Similarly, WCDMA and HSPA also share the locations of base stations and all the elements of antennas. With the evolution of WCDMA to HSPA, few adjustments at the hardware level were needed, both in base stations and in the RNC, to support the high data rate. However, some software packages were needed to adapt the WCDMA to HSPA. Due to the share structure between WCDMA and HSPA, the cost of changing from one technology to another was very low compared with the construction of a new standard for the network [50]. Thus, research in this area can evolve to develop technologies that can contribute for better support of QoS across the network [51].

2.1.5 LTE Convergence

In order to be prepared for the future needs, 3GPP has started the activity on the HSPA evolution. The UTRAN long-term evolution project group looks at the market introduction of the Evolved UTRAN (E-UTRAN), being the resulting specification available in Releases 8 and 9 [52]. The objective of Long Term Evolution (LTE) is to provide high data rate, low latency, and packet optimized radio access technology supporting flexible bandwidth deployments. In parallel, a new network architecture has been designed with the goal to support packet switched traffic with seamless mobility, quality of service, and minimal latency. 3G evolution consists of two parallel tracks: HSPA evolution and LTE. Figure 2.3 illustrates the relation between HSPA evolution and LTE,as it can be seen from Release 8 ahead.

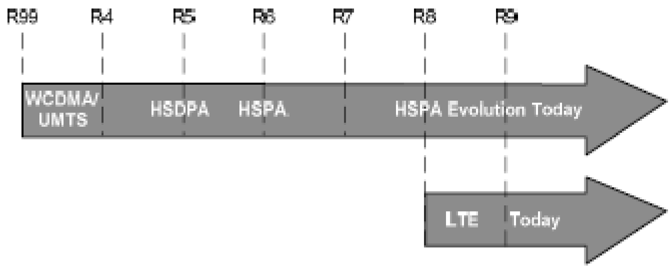

Figure 2.3: HSPA and LTE Evolution

To support the new packet data developments provided by the LTE radio interfaces, a new evolved core network was developed. The work of specifying the core network is commonly known as System Architecture Evolution (SAE). All network interfaces are based on Internet Protocol (IP). The E-UTRAN NodeBs (eNB) are interconnected by means of an X2 interface, while the Mobility Management Entity (MME)/Gateway (GW) entity and the eNBs are interconnected by means of an S1 interface, as shown in

Figure 2.4. The S1 interface supports a many-to-many relationship between MME/GW and eNBs. The functional split between eNB and MME/GW is shown in Figure 2.5. Two logical gateway entities called the Serving Gateway (S-GW) and the PDN Gateway (P-GW) are defined in this figure.

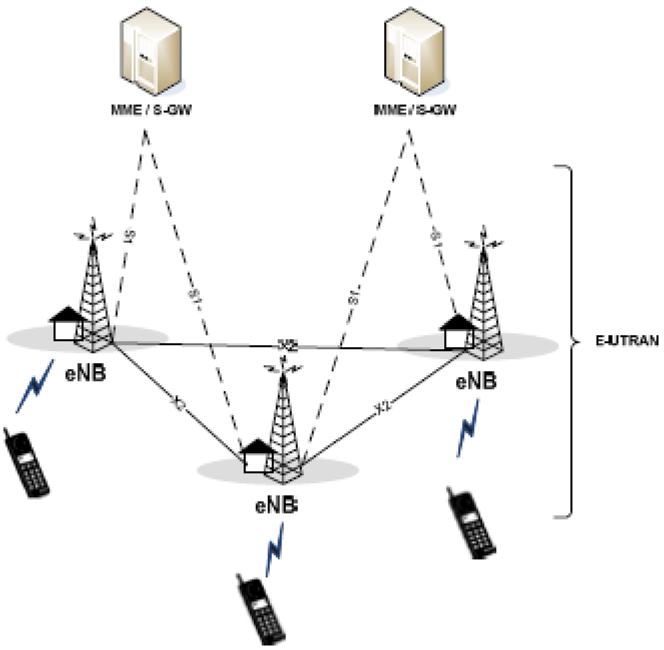

Figure 2.4: LTE Overall Architecture

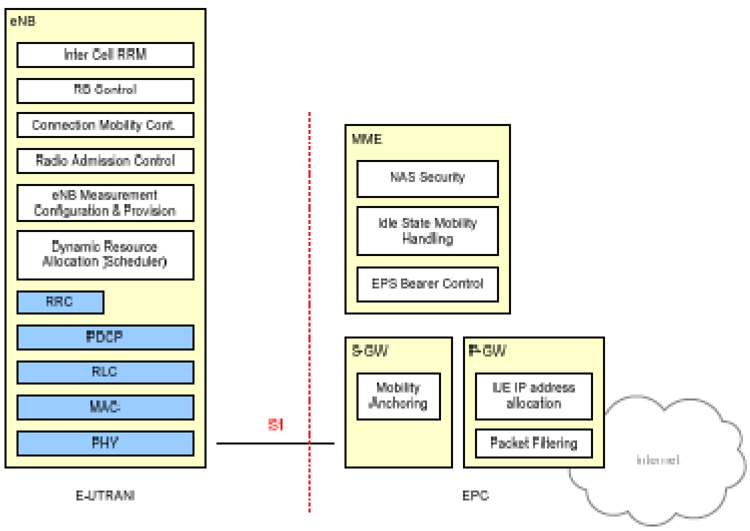

Figure 2.5: Functional Split Between E-UTRAN and MME/GW

The S-GW works as a local mobility anchor receiving and forwarding packets to and from the eNB serving the UE. The P-GW interfaces with the external Packet Data Network (PDN), such as the Internet and the IP Multimedia Subsystem (IMS). The P-GW also has several IP functions, such as address allocation, packet filtering, policy enforcement and routing. The MME is used to perform signaling and not to transport user IP packets. An advantage of a separate network entity for signaling is that the network capacity for signaling and traffic can grow independently. The main functions of MME are idle mode User Equipment (UE) reachability, including the control and paging retransmission, roaming, authentication, authorization, P-GW/S-GW selection, bearer management including dedicated bearer establishment, security negotiations and Non-Access Stratum (NAS) signaling. Evolved Node-B implements the Node-B functions as well as the protocols traditionally implemented in RNC. The main functions of eNB are header compression, ciphering and reliable delivery of packets. On the control side, eNB incorporates functions, such as admission control and radio resource management. Some of the benefits of a single node in the access network are reduced latency and the distribution of RNC processing load into multiple eNBs. The 3GPP project is working to allow LTE improvements to be seen by all kinds of access network technologies in a transparent way.

2.2 WLAN Technology

A WLAN is a communication system where a user is able to connect to a Local Area Network (LAN) using no wires. WLAN was designed as an alternative to the wired LAN to minimize the need for wired connections (see Figure 2.6). WLAN works similarly to the cellular system, offering a combination of data connectivity and