MOBILITY MANAGEMENT IN WIRELESS IP NETWORKS

Mathalenta Jane Rosa Ngwane

2005

MOBILITY MANAGEMENT IN WIRELESS IP NETWORKS

Mathalenta Jane Rosa Ngwane

(000287)

A dissertation submitted to the Faculty of Science & Agriculture in fulfilment of the requirements for the degree

MASTER OF SCIENCE

In

COMPUTER SCIENCE

Department of Computer Science

University of Zululand

2005

DECLARATION

I,…………………………………………………………….. declare that this dissertation represents research work carried out by the author and has not been submitted in any form to another University for a degree. All the sources I have used have been duly acknowledged in the text.

____
Signature of student

DEDICATION

I should thank my families for supporting me when committing myself to the dissertation. To God Almighty for giving me strength and good health, and his encouraging word.

ACKNOWLEDGEMENTS

I would like to express my thanks to all who gave me advice, encouragement and assistance in this study. I especially thank my supervisor Prof. M. O. Adigun, for his advice and patience in supervising this dissertation. I must thank my co-supervisor Mr. G.E Ojong, who gave me considerable guidance and for being very patient with me and for teaching me how to face challenges and conquer them. I know that if it was not for him, I would give up on this study. Finally, I gratefully acknowledge the useful comments from the senior research group comprised of, Mbali, Sihle and Sara for giving me encouraging support through the toughest of times. To my research colleagues, Phiwa, Mzomuhle, Bheki, Divya, Mcebo, Thulani, Paul, Reuben, Edgar and Mohammed, they indeed gave me many suggestions and support in many aspects.

I would also like to thank many contributors, who are known or anonymous, particularly Telkom for giving me sponsorship through the Centre of Excellence in the University of Zululand. Without their enthusiasm and generosity, I could not accomplish this dissertation.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Mobility management plays a significant role in the current and the future Internet in delivering effective services to the mobile users on the move. This research proposes a mobility management architectural framework that will incorporate wireless mobility architecture and call admission control. To achieve this we carried out three tasks, (i) the architectural basis for the mobility management strategy for real-time data communication was established. (ii) a suitable call admission criteria and a corresponding call admission control algorithm for the strategy was defined. (iii) a simulated performance evaluation of the strategy was developed. The results obtained are as follows: (i) this study resulted in a strategy that prioritizes calls of the guaranteed services class over calls of other classes, (ii) the higher priority of handoff calls over new calls has been achieved; and (iii) the proposed Call Admission Control (CAC) algorithm shows to have an acceptable performance with an average acceptance rate of calls of up to 83%. In conclusion this study resulted in a CAC algorithm that prioritizes real-time traffic over non real-time traffic as well as handoff calls over new calls.

CHAPTER ONE

1.0 INTRODUCTION

1.1 Overview

The issue of mobility today is of great concern since the demand for wireless communications has grown tremendously. The need for reliable seamless quality of service (QoS) has been a driving force for the increasing growth and advances in wireless communication devices. Mobility management is the critical issue to consider in providing reliable QoS anywhere any time. Requested information and services have to be delivered to mobile users with adequate QoS. This demands a closer look at how resources (bandwidth and buffer) are utilised in the Internet, and how we can maintain continuous connectivity while mobile, especially during the transmission of real-time data. Modifying the current Internet to support mobility is one of the fastest growing research areas in the global Internet research community. In this study we propose a mobility management architectural framework that will allow mobile users (real-time and non-real-time) to roam seamlessly with an open connection.

The Internet Protocol (IP)[1] is a protocol that has been proposed to provide seamless Internet communication. This protocol has been reliable to most Internet users over the last few years in making sure that communication services over the wire line Internet satisfy Internet users. The unfortunate drawback of this protocol is that it lacks mobility features since it was developed without mobility in mind. The future Internet will require mobility features that would sustain the use of mobile devices that are growing in leaps and bounds.

Considering this, the Internet Engineering Task Force (IETF) proposed Mobile IP [2], an Internet Protocol that caters for mobile users while they are in communication away from their home network. This protocol allows a mobile user to move from its home network to another network and still receive services as if it is at home. Unfortunately, Mobile IP does not state how a mobile node should behave when moving from one network to another with an open connection.

With the advent of mobility on the Internet, researchers are also looking at how real-time applications would be accommodated on a mobile Internet. This would mean the incorporation of QoS in mobility. The resource reservation protocol [27] which has been proposed for the wireline Internet is being proposed to be used with Mobile IP [4].

Today's Internet demands a reliable wireless architecture that would allow seamless communication while a user is mobile.

1.2 Statement of the problem

Many wireless architectures have been developed so far. The Global System for Mobile communication (GSM) is one of the most widely used architectures [5]. GSM provides mobile users with basic cellular services like short message service, call barring, call forwarding, call holding, call waiting etc, and it has been used for many years by cellular services providers. One of the limitations of this architecture is that it does not meet the demands of the mobile Internet users [5]. Mobile Internet users need a packet switched wireless architecture so that the cost of using the services will be reduced [6]. An everyday problem experienced by mobile users is break in communication while a user is moving with an open connection. When a node is moving with an open connection, the signals from the base station that it uses may get weaker and the mobile node needs to register with the next base station [21]. The new base station may reject the incoming call because of lack of enough resources and the user would experience a break in connection [29]. Therefore, advance reservation of resources may have to be done in the new cell. This would require the performing of call admission control (CAC) procedures. Current CAC procedures that have been proposed are very complex.

This study investigates an IP-based wireless mobility architecture that would seamlessly take into consideration the movement of a mobile host and accommodates both real-time and non real-time applications.

1.3 Need for the Research

A lot of research on CAC mechanisms has been done with an attempt to give mobile users adequate QoS [13], [14], [15]. These schemes consider handoff calls and new calls only, in their call admission procedures. They only reserve a portion of radio channels for handoffs, and new calls may only be admitted if there are channels available after handoffs have been catered for. Such schemes overlook the real-time demands of mobile users for the current Internet since nothing is said concerning real-time applications in their CAC schemes. The study considered hierarchical wireless network architectures that exist in the literature and to adopt one that best suits our call admission control requirements. A call admission control mechanism is to be proposed, that makes an efficient use of network bandwidth and support real-time and non-real-time applications. There is a great clamour for mobile users to be able to enjoy the same services while mobile, as at home or office. The research carried out in this study is going to be an added contribution to this quest. This includes improvement to support multimedia applications, and most importantly effective utilisation of resources. This research, if successful, will be a contribution to improving the mobility aspect of the current Internet.

1.4 Research Goal and Objectives

1.4.1 Goal

This study proposes a mobility management architectural framework that will enable the future Internet architecture to use suitable mobility architecture to provide support for seamless real-time data communication; and to employ a simple call admission control algorithm that will enable efficient resource utilization.

1.4.2 Objectives

The goal is refined into the following specific objectives:

To establish the architectural basis for the mobility management strategy for real-time data communication;

To define suitable call admission criteria and a corresponding call admission control algorithm for the strategy and

To develop a simulated performance evaluation of the strategy.

1.5 Organization of the Thesis

This thesis is organised as follows: Chapter two presents a background covering mobility and QoS concepts in networking. Cellular architecture and its elements are discussed, followed by Linear and hierarchical mobility architectures. Finally the impact of QoS on mobility is presented.

Chapter three presents similar work that has been done elsewhere. Mobile IP and its extensions are discussed in two categories under micro-mobility and macro-mobility.

Call Admission Control being the heart of this research, is discussed in four categories. Finally, Rainbow Services (RS) is presented here as the bandwidth allocation scheme on which the proposed CAC is based.

Chapter four discusses how the proposed model evolved. First, architecture is defined as the context in which the proposed CAC is going to be applied. Second, the derivation of admission criteria, from Rainbow Services' QoS constraints is presented using a class-based approach. Lastly, the putting together of the CAC algorithm is discussed. The chapter ends with the presentation of simulation model, simulation design and implementation.

Chapter five presents the results of the simulation for the proposed CAC algorithm. The chapter concludes with a comparison of the results with one of the similar schemes that exist in the literature.

Chapter six concludes the content of the dissertation. The limitations of the proposed CAC scheme is highlighted and future directions are outlined.

CHAPTER TWO

2.0 BACKGROUND

2.1 Introduction

Recently, the current Internet has changed drastically compared to the two decades since it came into existence. Previously when one thinks about using the Internet services, the first thing that came to mind was going home or to the office and access the Internet using stationary desktop computers. The current Internet has improved in such a way that one can access the Internet while in the bus, train or anywhere using mobile devices such as cellphones and PDAs. That change has led to much interest in Internet services by everyone; young or old. This change also has put more challenges to network designers to design networks with mobility in mind. This has brought even more constraints on QoS because mobile Internet users need to get even more reliable QoS compared to stationary users.

The challenge mostly faced by cellular designers is to design the devices that will allow their customers to be able to access Internet services. In the next sections, the focus is on cellular architectures that are employed today and their components with emphasis on the mobility architectures that exist in the literature and their structural layouts. The impact that QoS has on mobility is also considered.

2.2 Cellular Architectures

A cellular access network allows cellular subscribers to move anywhere in the country and remain connected to the **Public Switched Telephone Network (PSTN)** through their mobile devices. The first generation (1G) cellular architecture that was designed used analogue signals which enabled voice communication only. This architecture was not designed with scalability in mind and it failed to accommodate the increasing number of mobile users.

The second generation (2G) cellular architecture known as Global System for Mobile Communications (GSM) was then developed [5], [6], [34]. GSM uses digital signals for communication and it also allows an increasing number of mobile users, but it did not allow mobile users to access the Internet and some basic mobile services such as sms and e-mail. This brought an evolution to third generation cellular architectures.

A third generation GSM was developed which allowed circuit switching (voice communication through a dedicated line) and packet switching (sending data using portions of the communication media). The packet switching capability of GSM is through the General Packet Radio Service (GPRS) [35] which was incorporated in the GSM network to allow GSM to access packet switched data.

The basic cellular network architecture has a hierarchical structure and it is formed by connecting the major components that form the cellular network which are Mobile Nodes (MN), Base Stations (BS), Cells, Mobile Switching Centres (MSC) and Public Switched Telephone Network (PSTN). This architecture is illustrated in figure2.1.

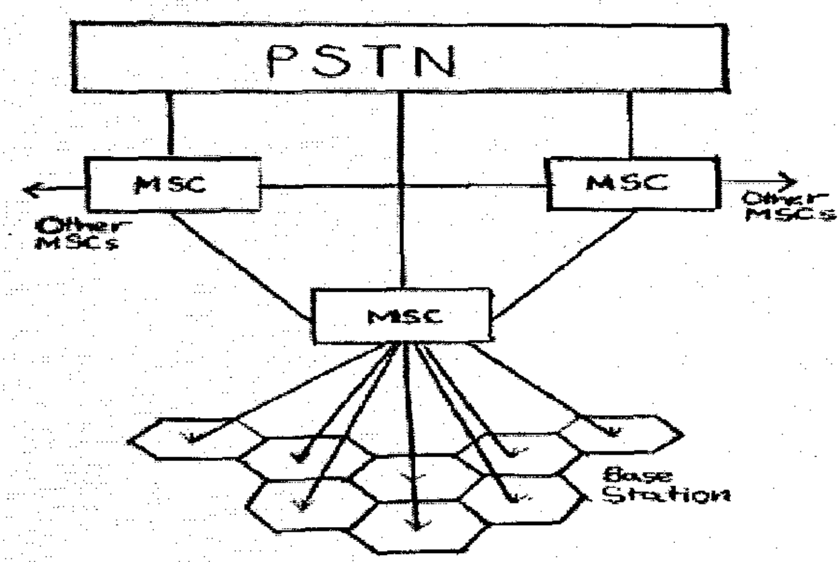

Figure 2.1: The basic cellular architecture [36]

A MN is a device that is used by a user to access services like making a call or sending an sms. A BS is a node that acts as a MN's point of connection while the MN is accessing network services. A cell is an area that may be of about few kilometres in diameter from the BS and the BS is controlling all the MNs that are in the cell region. Two or more cells may be connected to form a cluster depending on the architecture's requirements. Each BS is controlled by the main node or router known as the MSC. Each MSC may communicate with another MSC. The MSCs are connected to a wireline network known as the PSTN. This structure is known as the Cellular mobility architecture. There are many mobility architectures that exist in the literature and they have different structural layout depending on the architecture's requirements. The next section discusses the concept architecting mobility into a network.

2.3 Mobility Architectures

Mobility architectures uses mechanisms that allow a mobile user to wander about in the network with an open connection and without getting disconnected or interrupted. Early architectures that were developed with the aim of allowing mobility to mobile users did not provide satisfactory QoS. This section discusses some of the mobility architectures in existence. Two types of mobility architectures are compared, the linear architectures and the hierarchical architectures

2.3.1 Linear Architectures

Linear architectures are those architectures where components are not controlled by any component. The information that is sent or received takes a specific direction. The typical linear network architecture as discussed in [33] is depicted in figure2.2

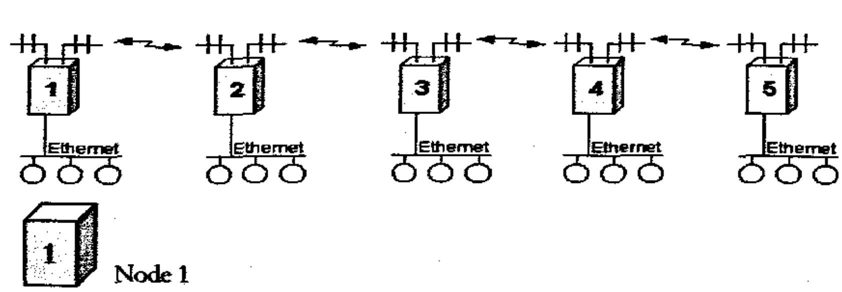

Figure 2.2: Wireless Linear architecture [33]

In this architecture each node (e.g. node1) is equipped with radio frequency ports. Each radio frequency port is equipped with a directional antenna which must face the direction of the neighbour node, say node 2. When node1 has some data to send to node5, the data has to go from node1 following the sequence of nodes from node 1 to node 5. If node 3 malfunctions the whole network may be disturbed because the signals of the directional antennas may not be powerful enough to transmit data to a far-away directional antenna.

2.3.2 Hierarchical Architectures

Hierarchical network architectures are those architectures that have components that are organised in different levels. A number of such architectures have been developed so far [20], [21], [24], [28]. Telecommunications Enhanced Mobile IP (TeleMIP) [20] is one of the hierarchical network structures that were developed in order to accommodate increasing QoS demands for mobile users. This architecture is a two-level hierarchical structure. The first level is composed of cells that are connected to form a domain. Each cell is controlled by one Subnet Agent (SA) (which basically is called a Base Station). All the SAs in a cluster are controlled by a node that is called a Domain Foreign Agent (DFA). A DFA controls two or more domains depending on the architecture's requirement. TeleMIP architecture has proven that hierarchical network architectures improve the efficiency of the network. Hierarchical network architectures accommodate as many mobile users as possible depending on availability of resources in the network, compared to linear network architectures. Through research, mobility has been shown to be the pre-requisite for any network infrastructure. As mobility continues to be part of our daily life, QoS also continues to demand extra attention everyday. The next section discusses QoS and its impact on mobility.

2.4 Impact of QoS on Mobility

With the rapid growth and excitement of using mobile devices today, the network designers and network services providers are faced with great challenge of meeting the mobile users QoS requirements. These challenges are on how network designers will make sure that the network architectures they design will accommodate the growing number of mobile users. It is also on how network services providers will make sure that the service they provide to mobile users is reliable.

We define QoS as a general term that incorporates bandwidth, latency, and jitter to describe a network's ability to customise the treatment of specific classes of data. For example, QoS can be used to prioritise real-time transmissions (e.g. video) over non real-time traffic, such as electronic mail transmission. Advanced networks can offer greater control over how data traffic is classified into classes and greater flexibility as to how the treatment of that traffic is differentiated from other traffic. The technique that best accommodates these QoS parameters is known as call admission control (CAC). CAC is the networking technique that accepts or rejects incoming calls on the network depending on the QoS constraints provided by the incoming call in advance. CAC mechanisms are supposed to be prioritising real-time traffic over non real-time traffic and prioritise handoff calls over new calls. A handoff call is a call that started in another cell, and because of the mobility of the user the call may have to be handed off to the next cell according to the direction of the user. This is a great challenge that researchers are using the development of CAC mechanisms to address this phenomenon.

CHAPTER THREE

3.0 LITERATURE REVIEW

3.1 Introduction

The current Internet demands, and user mobility today, present a great challenge to developers of mobile wireless architectures. The big question they need to answer is "how existing mobile wireless architectures could be modified to satisfy mobile users' increasing QoS requirements?" One of the appropriate answers will seriously consider call admission control mechanisms that should be applied on mobile wireless architectures. This chapter gives an overview of the literature on Mobile IP and its extensions. First, Mobile IP is discussed being our base mobility protocol. Second, the existing mobile wireless architectures are also discussed especially hierarchical architectures. This is because today's cellular network is represented as a hierarchical structure. Lastly we discuss existing call admission control mechanisms.

3.2 Mobility

Mobility allows mobile users to have access across a wide range of networks for a wide range of needs. Mobility solutions provide several benefits, but also bring challenges that mobile network developers must address within network architecture.

Mobility can be divided into two types: micro-mobility and macro-mobility which are discussed in the next two sections respectively.

3.2.1 Macro-mobility

Macro-mobility is the type of mobility that considers the movement of a MN across different domains in various geographical regions. Most of these architectures use Mobile IP for macro-mobility requirements. This section discusses some of the existing macro-mobility architectures.

Internet Protocol (IP) is the protocol that was designed to serve stationary Internet users for many years. This protocol has been reliable for this type of Internet usage, but with the evolution of mobility, real-time multimedia requirements of the current Internet has brought challenges that could not be met using IP. The IETF devised Mobile IP (MIP) as a protocol which was meant to accommodate mobile users. The first version of MIP's routing mechanism is called triangle routing.

Mobile IP's first version, Mobile IPv4 [3], has been adopted but it had some underlying drawbacks because there was considerable delay in the packets destined for the Mobile Node (MN). The delays are caused by the way packets are routed to the mobile node. The Correspondent Node (CN) which is the node that has packets to send to the mobile node does not recognise even if they are in the same network.

Packets sent by the CN to the MN go through the Home Network (HN) then to the Foreign Network (FN) where they can be delivered to the mobile node.

The mobile node cannot receive messages directed to it, before it is connected to a foreign network. The home agent (the router in the mobile node's home network) tunnels and redirects data meant for a mobile node, to the foreign agent (the router in the mobile node's foreign network). The foreign agent then directs the data to the mobile node. The mobile node then replies by directing the data straight to the CN. This routing mechanism is called Triangle Routing [3]. It causes a lot of delay in the mobile node's data (figure 3.1).

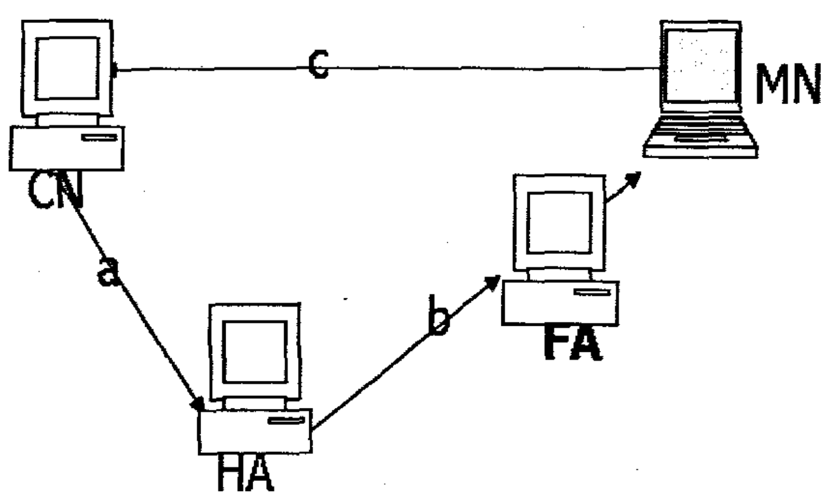

Figure 3.1: Mobile IP's Triangle Routing

To solve the problem of Triangle routing, the IETF proposed the Route Optimisation [7], [8] mechanism.

Route optimisation [7], [8] (figure 3.2) is an improvement of the triangle routing. The first process of routing packets is the same as that of triangle routing, where packets from the CN destined for the MN go through the HA. They are then tunnelled to the FA and delivered to the MN. After the packets from the CN have reached the MN, the MN communicates its care-of-address to the CN.

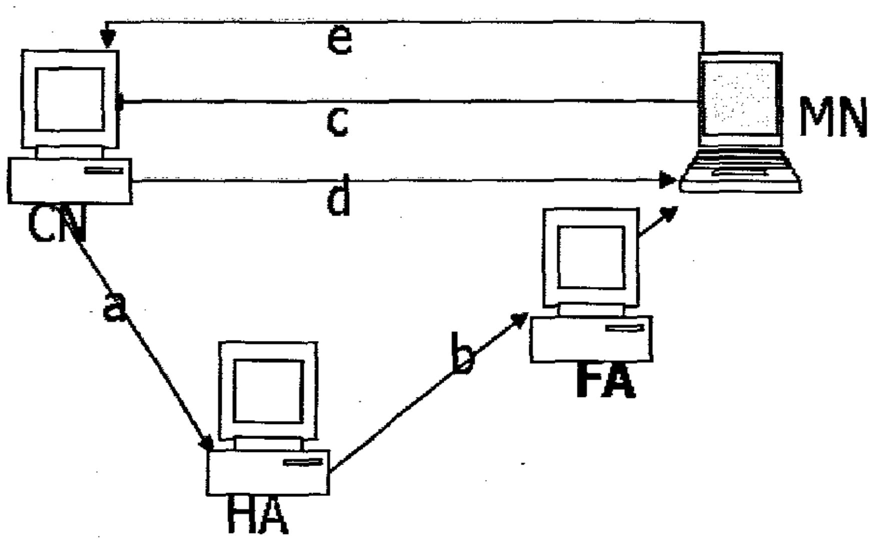

Figure 3.2 MIP's Route Optimisation

Then the CN starts communicating directly without passing through the HA. This protocol does not cater for a mobile that moves from one network to another with an open connection. It also does not state what happens when the MN is moving from one FN to another. In an attempt to improve on the capabilities of MIP's route optimisation, Multicast-based Mobile IP [9] was proposed.

IP Multicast-based Mobile IP [9] is a mobility protocol that works like Mobile IP but it has more mobility features compared to Mobile IP. When away from home, a MN gets a new care-of address by stateless (given by a Foreign Agent) or stateful (Dynamic Host Configuration Protocol) address auto configuration, just as in MIP. It then registers its new care-of address with the Multicast group (*MG),* defined by its multicast address (*IPMCAST*), using Internet Group Membership Protocol (IGMP). Each time the MN moves from one subnet to another, it joins the Multicast group, with its new care-of-address and leaves MG with its previous care-of address. If the MN can simultaneously access its previous and new point of attachment, it should not leave *MG* with its previous care-of address until it starts receiving data on its new care-of-address. As a result, packets will be multicast to both care-of-addresses. This architecture consumes a lot of network resources by allowing duplicate packets to be multicast in two IP addresses.

Another MIP extension is the RSVP-Mobile IP scheme which was proposed in [4]. This scheme makes use of MIP and RSVP together. This model suggests that the MN uses two IP addresses, Domain Care-of Address that is going to be recognised globally and Local care-of-Address that is going to be used in a local access network. There is also a RSVP-MP which is capable of tunnelling packets to the MN. The only limitation of this model is that it places a heavy burden in one access router.

These protocols were developed in an attempt to accommodate the evolving mobility of users and their real-time multimedia requirements, but there are no built-in mechanisms to provide suitable QoS to mobile users. Presented in the next section are mobility and schemes that have been proposed for the current Internet to accommodate mobility. The concept of QoS provisioning on the Internet is also covered.

Cellular IP [24], [28] is a lightweight, robust, host mobility protocol that supports micro-mobility and frequently migrating hosts but efficiently interworks with Mobile IP to provide macro-mobility. Cellular IP is a new mobile host protocol that is optimised to provide access to a Mobile IP enabled Internet in support of fast moving wireless hosts. Cellular IP incorporates a number of important cellular principles but remains firmly based on IP design principles. The Cellular IP's network architecture consists of Cellular IP mobile node (MN), which is a mobile device that implements the Cellular IP protocol. Cellular IP base station (BS) serves as a wireless access point for the mobile host and routes IP packets while performing all mobility-related functions. They are built on regular IP forwarding engine with the exception that IP routing is replaced by cellular IP routing and location management.

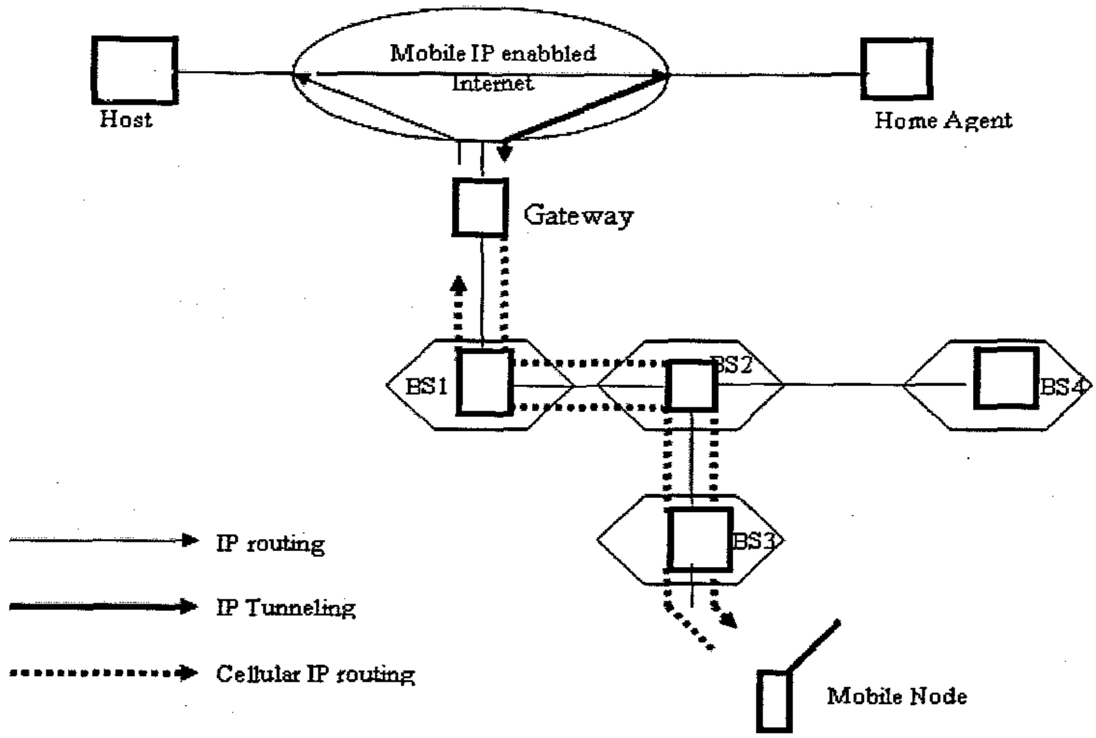

Figure 3.3: Cellular IP architecture

Cellular IP BS is responsible for authentication of the mobile hosts and routing of IP packets inside the Cellular IP network. It also communicates with mobile hosts via a wireless interface. Cellular IP Mobile Agent (gateway) is a Cellular IP node that is connected to a regular IP network by at least one of its interfaces.

BSs within the cellular IP access network periodically emit beacon signals. Every beacon signal contains parameters related to BS: Cellular IP network identifier, gateway IP address and ID of the paging area. When the mobile host first enters the access network, it listens to beacon signals sent from the BS. The beacon signals allow the mobile host to know the nearest Base Station and registers with it. The Base Stations record the interface through which they received the beacon signal. After receiving the beacon signal, the mobile host has to send the data packets to the gateway through the recorded interface. The data packet will contain the mobile host's Home address and the subnet ID. Cellular IP access networks are connected to the Internet via Cellular IP Mobile Agent (gateway). Mobility between gateways (macro-mobility) is managed by Mobile IP while mobility within the access networks (micro-mobility) is managed by Cellular IP. Assuming Mobile IP and no route optimisation, packets will be first routed to the host's HA and then tunnelled to the gateway. The gateway detunnels packets and forwards them toward a BS. Inside a Cellular IP network, mobile hosts are identified by their home address, and data packets are routed without tunnelling or address conversion. The Cellular IP routing protocol ensures that packets are delivered to the host's actual location. Packets transmitted by mobile hosts are first routed toward the gateway and from there on to the Internet.

The Cellular IP architecture saves network resources by reducing registration and control messages by the mobile host. This protocol does not clearly tell us how the Home Agent will know to which gateway to send the packets destined for the mobile host. It also does not state how the Home Agent will know the current location of the mobile host before it sends the first packet to it.

Handoff Aware Wireless Access Internet Infrastructure (HAWAII) [22] is another proposal dealing with the optimisation of Mobile IP. This architecture was developed with Mobile IPv4 in mind and has many similar features with Cellular IP.

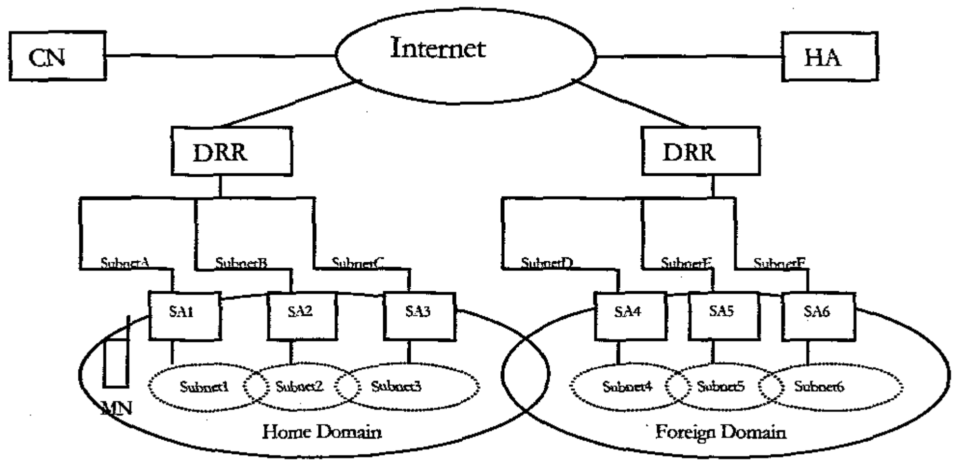

Figure 3.4: HAWAII architecture

The HAWAII has the following elements: there is a Domain Root Router (DRR) which acts as a gateway to the global level. Domains are organised in a hierarchical structure. Each domain consists of subnets and each subnet is controlled by routers. MN uses its IP address if in its Home Network (HN). MN uses a co-located care-of-address (which is the IP address of the foreign agent) when in the Foreign Domain (FD).

When a MN enters a new domain it registers with a DRR by giving its Home IP address to the DRR of that domain. The DRR replies by giving the MN a co-located CoA (a care-of address which is the IP address of the foreign agent). The MN registers the co-located CoA with its HA. All the incoming packets destined for the MN are sent to the specific DRR according to the subnet address. When the MN moves within the domain it uses its Home IP address. When it moves to a new domain, MIP caters for that movement. The DRR creates MN specific routes in each router from the MN to the DRR. Packets are sent to and from the MN using these host specific routes. To maintain the MN's states when idle or it first powers up, the MN sends path hyphenate update messages to the DRR of that domain.

This architecture reduces the number of control messages by the mobile node. It caters for macro-mobility by using Mobile IP.

3.2.2 Micro-Mobility

Micro-mobility is the type of mobility that describes the movement of a MN within a subnet. This type of mobility is sometimes referred to as intra-domain mobility in some mobility architectures. In this section, some of the micro-mobility architectures that exist in the literature are presented.

The Intra-Domain Mobility Management Protocol (IDMP) [21], [30] uses a hierarchical structure to manage node mobility in future IP-based cellular networks. By aggregating multiple subnets into a mobility domain, IDMP localises the scope of most location update messages and drastically reduces both the global signalling load and the update latency. The IDMP architecture (figure 3.3) consists of the Mobility Agent (MA) which acts as a gateway or domain-wide point for packet redirection; a SA which is responsible for subnet specific mobility services; a subnet, which is an area that consists of a number of cells. The area is controlled by an SA. A cell is an area that is controlled by one Base Station (BS). A cell may have many mobile nodes that are being served by the BS.

When a mobile host first enters a subnet, it registers with the serving SA during which it is given a Local Care-of-Address (LCoA) and a designated Mobile Agent (MA). The mobile host then communicates its LCoA with the MA. The MA in its reply gives the MH the Global Care-of-Address for global communication. The MH is responsible for communicating its GCoA to its Home Agent (HA) for routing of packets. When a MH moves within the subnet, it only registers with the serving SA and gets a new LCoA. The MH updates the MA of its new local binding on every subnet change.

The MA then forwards packets to the MH using the Internet routing tables. GCoA can be used for global routing and it is also not mandatory to tunnel from CH or HA to the MA. Packets from the sending host are forwarded to the GCoA and are intercepted by the MA. The MA then tunnels the packets to the MH's current LCoA. When a MH is moving to a new subnet, the old SA or the MH sends a Movement Imminent message to the MA informing it to broadcast all the packets destined for the MH. The MA broadcasts all the packets that are destined for mobile host to all neighbouring subnets.

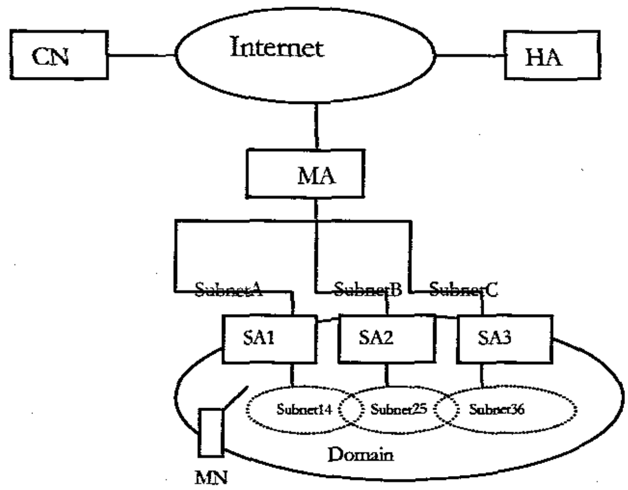

Figure 3. 5: IDMP architecture

The SAs buffer all the incoming packets to hosts specific buffers. When the mobile host registers with a new SA, the latter SA delivers the buffered packets to the MH even before the MH communicates its location update with the serving MA.

Paging support for IDMP assumes that subnets are grouped into Paging Areas (PA) identified by some unique identifier. An MH that is in the passive or idle state can be able to detect changes in its PA by listening to identifier broadcasts. When a MH moves from one subnet to another in the same PA, no subnet specific registration has to be done. When packets destined to a MH arrive in the MA, the MA sends a Page Solicitation message to all the subnets in the MH's PA and buffers those and other incoming packets. When a MH switches to an active state and registers, the packets are forwarded to it.

Since global binding updates are generated only when the MH changes from one domain to another and obtains a new GCoA, this approach reduces signalling load. The IDMP protocol does not cater for inter-domain mobility. The MH is the one that does domain-specific registration and HA notification. This protocol does not tell us what happens when an MH moves from one PA to another. We are not told how other SAs know when to discard the buffered packets during handoff or during paging. Another scheme was proposed in [20] that uses IDMP for micro-mobility.

Telecommunication Enhanced Mobile IP (TeleMIP) is a two-level hierarchical mobility management architecture [20]. This architecture separates micro (intra-domain) mobility from macro (inter-domain) mobility for next-generation cellular networks. The network of this architecture is partitioned into domains and subnets.

When a MN enters a new domain it registers with the serving BS by replying to agent advertisements sent by an SA. The BS sends the agent advertisement reply to the serving SA/FA. The serving SA/FA gives the MN the LCoA and assigns it a designated MA. The MN sends a registration request to the MA. The MA replies by giving the GCoA to the MN. The MN does a global binding update with its HA. All the packets from the CN to the MN go via the HA, the HA send them to the GCoA and are intercepted by the MA. The MA then sends them to the MN. When the MN moves from one subnet to another, it only changes the LCoA and does the binding update with the serving MA. The GCoA remains the same. When the MN moves from one domain to another, its inter-domain mobility is managed by traditional Mobile IP and the GCoA changes. TeleMIP uses IDMP for intra-domain mobility.

Since global binding updates are generated only when the MH changes from one domain to another and obtains a new GCoA, this approach reduces signalling load. The MH is the one that does domain-specific registration and HA notification.

Many mobile wireless architectures use a hierarchical structure to represent their network functionality and improve on QoS for their mobile users. In a real-time application or service such as the one targeted in this work, registration and control activities, if not minimised, could degrade the system's QoS. Therefore, the contribution of this work is to provide strict QoS to mobile users by allowing mobile users to request reservation of resources as the network performs admission control on the request. In the architecture we shall adopt, the CAC is the major aspect that is going to be emphasised because CAC is one of the mechanisms that can improve the efficiency of the network especially when the network accommodates both real-time and non real-time applications.

3.3 Call Admission Control

Call admission control (CAC) is a technique that is used to provide QoS in a network by restricting the access to network resources. An admission control mechanism accepts a new call request provided there are enough free resources to meet the QoS requirements of the new call request without violating the committed QoS of already accepted calls. There are many call admission control mechanisms that have been proposed so far. Existing call admission control mechanisms can be grouped into the following categories: Policy-based CAC, Threshold-based CAC, History-based CAC and Statistical CAC. Presented in the sections that follow are the four categories described in detail.

3.3.1 Policy-based Call Admission Control

Policy-based CAC schemes use a certain policy in admitting different calls on the network. For example the scheme in [12] proposes An Adaptive Algorithm for Call Admission Control in Wireless Networks which is built upon the concept of guard channel policy. The guard channel policy is also discussed in [13]. It solves the problem of handoff call-blocking by monitoring the handoff blocking rate at each base station. If a base station experiences a high handoff blocking rate, the number of guard channels is increased until the handoff blocking rate drops below a threshold. If the base station is seen to be using a fraction of guard channels over a period of time, the number of guard channels is decreased until most of them are used frequently. By so doing, the handoff blocking rate will be brought close to its threshold. New calls can only be admitted if the total number of on-going calls is less than the threshold of admitted calls even if some guard channels are available. With this scheme, it can be seen that where there are many handoff calls waiting to use the network and admitted calls take too long to release the channels, new calls may never be admitted into the network.

A similar scheme was also developed in [18] which is an admission policy for two classes of service where each class occupies a different number of channels. They used fractional guard channel policy for a single class. This policy states that new calls are accepted with a certain probability that may depend on the channel occupancy. The call is lost when buffer size reaches a threshold.

3.3.2 Threshold-based Call Admission Control

The schemes under this category use a threshold to measure the performance of their CAC schemes. Javenski et al [14] proposed an Admission Control for QoS Provisioning in Wireless IP Networks. Here, a new call is accepted if the sum of both reserved resources in the cell and bandwidth demanded by that call is lower than a threshold. Handover is accepted in the target cell if the sum of occupied bandwidth by the other active flows in the cell and the requested bandwidth for that call is lower or equal to a threshold. This scheme makes an efficient use of network bandwidth by not violating the QoS constraints of already admitted calls, but this scheme does not consider the type of traffic whether it is real-time or not.

In Distributed Adaptive Admission Control in Mobile Multimedia Network [15], each cell monitors the number of multi-class handoff call attempts and handoff failure occurrences as a function of time. When the number of a certain class of calls in the observed cell is seen to be greater or equal to a threshold, and the call request of that same class is placed and not immediately rejected in the corresponding cell, there is a possibility of the grant of request for that call. The cell also monitors if a certain class of calls is greater than or equal to the maximum number of channels reserved for that class of calls. If the latter is true and the new call request is posed to the corresponding Base Station, it could be rejected or put on a queue for sometime until a channel is free. This scheme does not have a way of making sure that real-time traffic is accommodated before accepting other classes of traffic.

3.3.3 History-based Call Admission Control

In history-based CAC schemes the network uses the movement history of the MN to anticipate the MN's future behaviour. Lim et al. [16] discussed an approach in which a base station is assumed to have some knowledge of the movement pattern of the Mobile Node placing the new call based on the Mobile Node's previous call history. This information is used by the base station to check if bandwidth reservation in the appropriate cells is possible. If it is possible the base station contacts the cell next to the current cell of the Mobile Node and cells that are further but in the direction of the Mobile Node. If the Mobile Node always follows similar moving patterns, a lookup table is used before a new call can be admitted; otherwise the new call is blocked. This scheme does not classify traffic. It only considers movement patterns.

3.3.4 Statistical Call Admission Control

Statistical CAC schemes use certain statistics or calculations to monitor the behaviour of the total load that the network accommodates. Lohi et al.[17] discussed Handover Issues and Call Admission Control in Cellular Systems. Here they considered the overload probability of the cell concerned when a new call arrives, or handover call is about to take place. They also considered the overload probability of the adjacent cells surrounding the concerned cell. They defined the overload probability as the probability that a cell cannot support any more calls after all the available channels have been used up. This scheme does not reserve any channels for handoff calls and it does not consider the type of traffic that occupied the channels. This leads to inefficient use of resources because one may find that channels are occupied by low priority non-real-time traffic at the expense of real-time traffic.

Tseng et al. [19] take the reserved flows in adjacent cells into account and calculate the cell overloading probability. They defined the cell overloading probability as the probability that the number of connections in the target cell is greater than the cell's capacity. If the overloading probability is greater than the highest tolerable overloading probability, the connection request is rejected. This scheme does not take into account the type of traffic.

Call admission control mechanisms that are discussed above take different views of call admission control. Most of the call admission control mechanisms prioritise handoff calls over new calls but the way they prioritise is such that new calls may never be admitted at some point. These schemes, again, are sensitive to calls already admitted on the network, but do not give preferences to real-time traffic over non real-time traffic. A call admission control mechanism is required that will prioritise handoff calls over new calls and real-time traffic over non real-time. The Rainbow services (RS) bandwidth allocation architecture in [11], [31] separates real-time traffic from non real time traffic.

3.5 Rainbow Services

Rainbow Services [31] is a novel QoS architecture recently proposed for the future Internet. It allows different classes of traffic to make efficient use of network bandwidth while at the same time guaranteeing the QoS of each traffic class. In addition to the Best Effort class, the developers of Rainbow Services have proposed four additional classes [32]. There are Guaranteed Constant Rate (GCR) for constant flow rate, Guaranteed Minimum Rate (GMR) for real-time variable rate flows, Assured Delivery Service (ADS) for flows that may tolerate delay, and Assured Low loss Service (ALS) for flows that may tolerate loss according to the loss profile. This study is about an admission control that takes into consideration the classes of the Rainbow Services [11], and uses them in prioritising the type of traffic to be considered first. Presented in the following section is the proposed mobility architecture in which the envisaged CAC algorithm will operate.

3.6 Overview of the Proposed Mobility architecture

Mobility management requires an architecture that will be able to cater for micro-mobility and macro-mobility and be able to apply call admission control without overloading the network with registration messages from the Mobile host. Looking at mobile wireless architectures that exist in the literature, it can be seen that wireless architectures divide the network into hierarchy of domains and subnets to provide better services to mobile users. The aim of the mobile wireless architectures is to minimise the registration process that the mobile node has to undergo before it can access services in its foreign network. TeleMIP architecture has been adopted as our base architecture. Another important aspect of mobility management is call admission control. From the discussion in section 3.5, it is clear that an efficient call admission control (CAC) mechanism is required, that will prioritise handoff calls over new calls and prioritise guaranteed service class over other classes. The proposed CAC mechanism is going to use RS service classes to accomplish this task and the proposed CAC mechanism is called Rainbow Services Call Admission Control (RSCAC).

CHAPTER FOUR

4.0 MODEL DEVELOPMENT

4.1 Introduction

This chapter presents a mobility management architecture that uses hierarchical structure of domains and subnets, but tries to do away with many registration messages that the mobile host has to send before accessing the services in its foreign network. This architecture also admits calls based on the Rainbow Services [11] service model which accommodates both real-time and non-real-time applications.

4.2 Proposed Architecture

Our architecture is a hierarchical structure that is composed of Subnets, Domains, subnet agents (SA), mobility agents (MA) and mobile nodes (MN), (figure 4.1). This architecture is based on the TeleMIP architecture [20]. We have chosen a hierarchical structure instead of a linear structure because as discussed previously, hierarchical structures perform better than linear structures. This is validated by the trend in the recently proposed mobility architectures. The majority use a hierarchical structure. The subnet is controlled by one SA which will act as a mobile node's point of attachment.

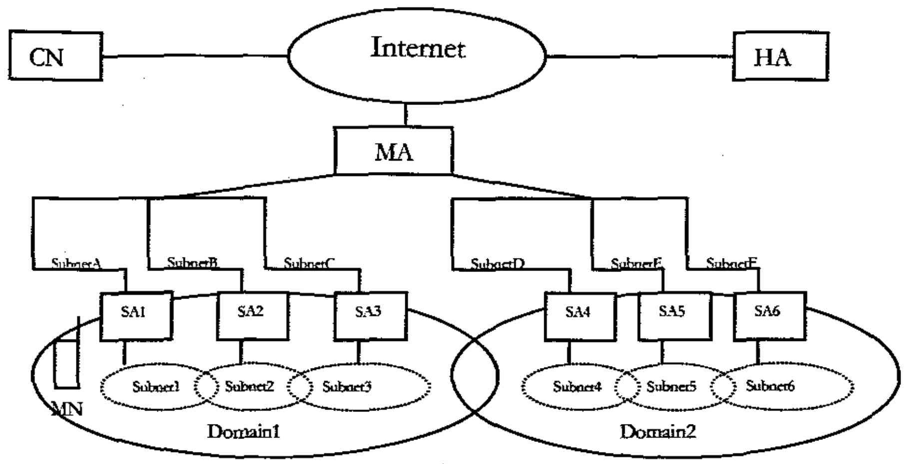

Figure 4. 1: The Proposed architecture

The Subnet Agent is also responsible for broadcasting beacon signals which are agent advertisements to all the mobile nodes that are coming from external domains. When the mobile nodes have responded to agent advertisements, the SA again will be responsible for local registration of every mobile node that enters its subnet and serving the MNs that are in its subnet.

The Domain is a wide area that is composed of seven subnets and is controlled by one Mobility Agent (MA). We chose to use seven cells as our domain because the standard cell structure accommodates six cells. We wanted to monitor the behaviour of RSCAC when calls are coming from all possible directions of the centre cell. The structure of a domain is depicted in figure 4.2. A MA is the node that controls two domains (to better monitor all the MNs) and is responsible for global registration for every mobile node that enters any of its domains, binding updates with the Home

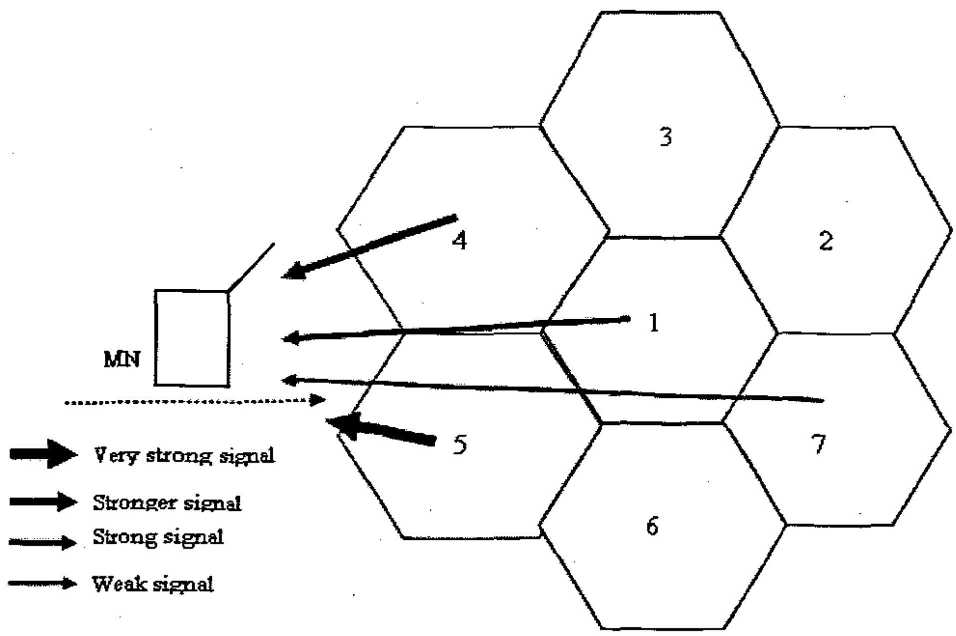

Figure 4. 2: The structure of a domain

Agent for the Mobile Node and serving the mobile nodes that are in its domain. The Mobile Node (MN) is the node that asks for services from the SA and can also move to another subnet or domain with or without an open connection.

When a MN first enters a new domain, it first listens to beacon signals broadcast by the SAs from surrounding subnets. Beacon signals are agent advertisements that contain the IP address of the serving SA for that subnet and the IP address of the serving MA for that domain. The MN decides to register with the SA according to the signal strength. The MN chooses to connect with the SA that has a stronger signal compared to other signals because it is assumed that when the signal is stronger it symbolises that the SA is nearer.

4.3 Call Admission Control Criteria

Our architecture takes into consideration the service classes of each individual flow. The service classes used are those proposed in the Rainbow Services architecture [16]. We first discuss the criteria for admitting flows that belong to each service class and then present our admission control algorithm. Figure 4.3 contains all the acronyms that are used in the proposed admission control algorithm.

B_{GCR} : Bandwidth allocated to GCR classes
B_{GMR} : Bandwidth allocated to GMR classes
B_{ADS} : Bandwidth allocated to ADS classes
B_{ALS} : Bandwidth allocated to ALS classes
$B_{class\ res}$: Bandwidth reserved for a class including handoffs of that class.
$B_{class\ new}$: Bandwidth required by a class making a new call
$B_{class\ total}$: Total Bandwidth assigned to that class
$B_{class\ used}$: Bandwidth used by on-going calls of that class
B_{total} : Total link bandwidth
Flow bandwidth : Data rate of the flow
$B_{GCRhand}$: Bandwidth reserved for GCR handoff
$B_{GMRhand}$: Bandwidth reserved for GMR handoff
$B_{ADShand}$: Bandwidth reserved for ADS handoff
$B_{ALShand}$: Bandwidth reserved for ALS handoff
L_{GCR} : Loss rate of a GCR class
R_{GCR} : Rate of a GCR class
D_{class} : The total delay experienced by a flow of a class

Figure 4.3: Acronyms for the proposed CAC algorithm

4.3.1 Admission Criteria for flows of GCR class

GCR class is allocated a fixed bandwidth on the network. This bandwidth may be used by flows of other low priority classes if flows of GCR classes are not using the whole bandwidth. When a flow of GCR class arrives, flows belonging to the lower classes that are using the channel will have to relinquish the channel for the GCR flow [31]. The flows of GCR will be admitted on the network if the following conditions are met:

I. The network maintains a constant rate throughout the session;

II. The flow's loss rate should be below a threshold e.g. 10^{-6}, since this is small enough for practical situations. The flow's loss rate is the rate at which a GCR class can tolerate packet loss;

III. The network makes sure that there are no handoff calls requesting some services;

IV. Delay bound D_{GCR} of the requesting GCR flow and of already active flow of GCR, would not be violated. D_{GCR} is the total delay experienced by a GCR flow and

V. The sum of GCR flows' reserved bandwidth, the bandwidth demand of the requesting GCR flow and the bandwidth used by the on-going GCR flows is less than the total bandwidth assigned to the GCR class.

$$B_{GCR\,res} + B_{GCR\,used} + B_{GCR\,new} < B_{GCR\,total} \quad \text{.................................4.1}$$

4.3.2 Admission Criteria for flows of GMR class

GMR class is allocated minimum bandwidth on the network. This bandwidth may be used by flows of other low priority classes if flows of GMR classes are not using the whole bandwidth. When a flow in the GMR class arrives, the lower classes that are using the GMR resource will have to relinquish the resource for the GMR flow [31]. The flows of GMR will be admitted on the network if the following conditions are met:

I. The flow states its loss profile. GMR's minimum loss rate L_{GMRmin} is greater than the network's average loss rate L_{Nave}.

$$L_{GMRmin} > L_{Nave} \quad \text{.......} \quad 4.2$$

II. The delay bound D_{GMR} of the requesting GMR flow and of the active flows of GMR and GCR should not be violated. D_{GMR} is the total delay experienced by a GMR flow;

III. The network makes sure there are no handoff calls requesting some services and

IV. The sum of GMR flow's reserved bandwidth, the bandwidth demand of the requesting GMR flow and the bandwidth used by the on-going flows is less than the total bandwidth that can be used by flows of GMR class

$$B_{GMR\ res} + B_{GMR\ used} + B_{GMR\ new} < B_{GMR\ min} + B_{GCR\ total} \quad \text{.......} \quad 4.3$$

4.3.3 Admission Criteria for flows of ADS class

An ADS class is allocated a fixed bandwidth on the network. This bandwidth may be used by other low priority classes if flows of ADS classes are not using the whole bandwidth. When a flow of the ADS class arrives, flows of the lower classes that are using the channel will have to relinquish the channel. If the whole bandwidth is occupied by ADS flows, incoming flows may be queued for some time or rejected [31]. The flows of ADS will be admitted on the network if the following conditions are met:

I. The sum of ADS flows' reserved bandwidth, and the bandwidth demand of the requesting ADS flow and the bandwidth used by the on-going flows is less than the total bandwidth assigned to the ADS class and

II. The network must make sure that there are no handoff calls requesting some services.

$$B_{ADS\ res} + B_{ADS\ used} + B_{ADS\ new} < B_{ADS\ total} \qquad \text{.......... 4.4}$$

4.3.4 Admission Criteria for flows of ALS class

ALS class is allocated a minimum bandwidth on the network. Flows of this class may also use bandwidth allocated to flows of GMR, GCR and ADS classes if there is some free bandwidth that belongs to these classes. When flows of any of these classes arrive, the flows of ALS class that are using the channel will have to relinquish the channel [31]. The flows of ALS class will be admitted on the network if the following conditions are met:

I. The flow of the ALS class must state its loss profile. ALS's minimum loss rate L_{ALSmin} must be greater than the network's average loss rate L_{Nave}.

$$L_{ALSmin} > L_{Nave} \quad \text{.......... 4.5}$$

II. The network must make sure that there are no handoff calls requesting some Services and

III. The sum of reserved bandwidth for ALS flows, the bandwidth demand of the requesting ALS flow and the bandwidth used by the on-going ALS flows is less than the total link bandwidth. BE bandwidth is assumed to be negligible.

$$B_{ALS\ res} + B_{ALS\ used} + B_{ALS\ new} + B_{GMR\ total} + B_{GCR\ total} + B_{ADS\ used} < B_{total} \quad \text{.......... 4.6}$$

4.4 Proposed Call Admission Control

Consider a wireless network which can support multiple types of services. To provide the desired QoS for each service, the network assigns channels according to the priority level. At a base station (BS), if there are different traffic types of call arrivals (including new calls and handoff calls), the network will give priority to handoff calls over new calls irrespective of the service class the handoff calls belong to. When a new call of any of the service classes wants to use a channel, the network will first check whether there are handoff calls also requesting to use the channels. If there are handoffs waiting, the network will give priority to handoff calls.

Our proposed CAC algorithm makes efficient use of the network by accommodating as many calls as possible, and maintaining a reasonably high level of network utilization. Many researchers have worked on the subject of AC and most of them view handoff calls as of higher priority compared to new calls [2], [8], [10], [11]. Many of their ideas centre around the efficient use or partitioning of the network bandwidth.

4.4.1 Admission Algorithm for new calls only

Assuming there are no handoff calls waiting to use the channel, the admission algorithm for new calls only will take place:

When a new request for any class is posed to the network, and the network will check whether or not it can provide reliable service to that class according to the class' admission criteria. If so, the network is going to accept that call. Otherwise the incoming call will be rejected figure 4.4.

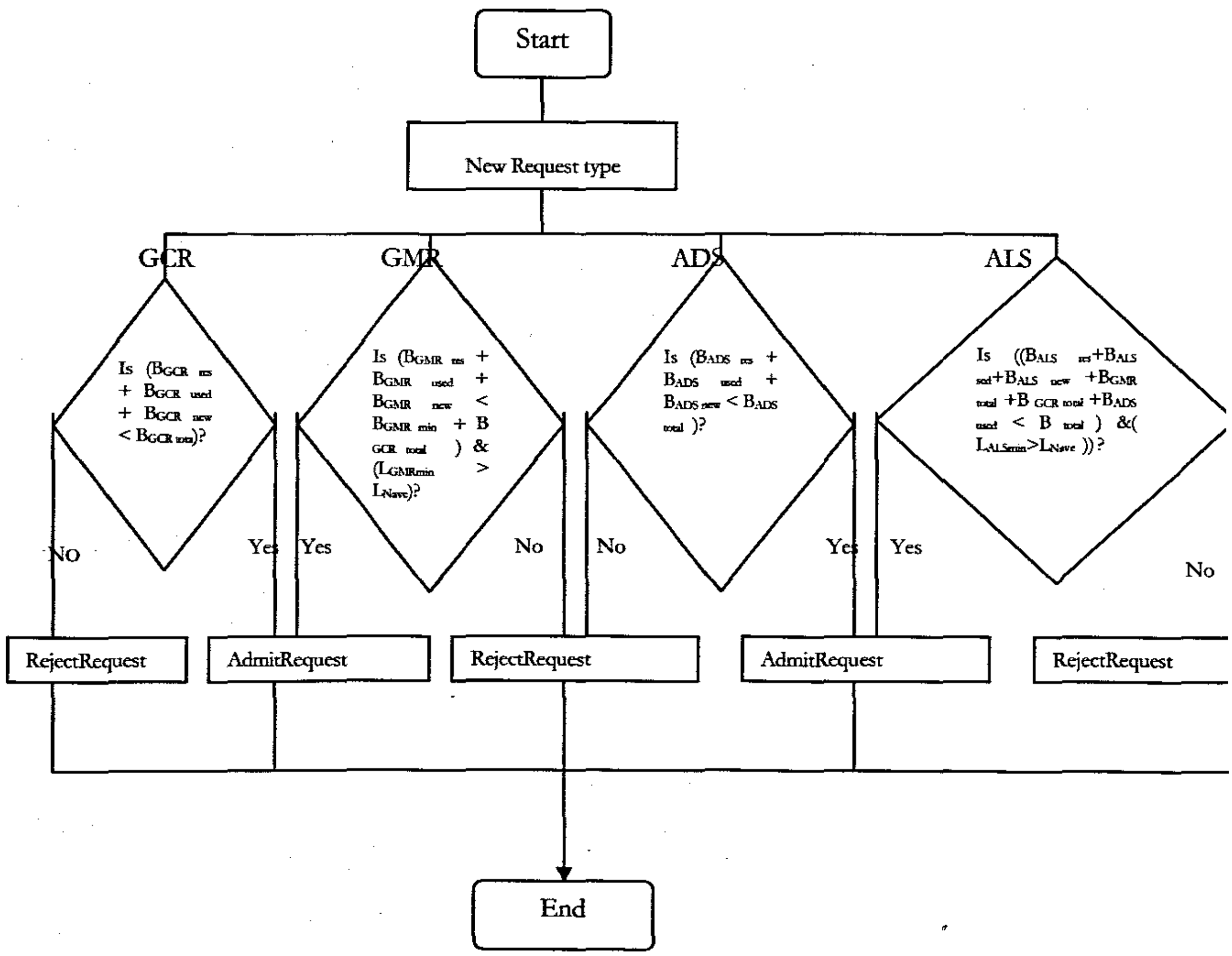

Figure 4.4: Admission algorithm for new calls only

4.4.2 Admission Algorithm for Handoff Calls only

Assuming that there are no new calls requesting some services, the admission process for handoff calls only is as follows:

Select Case HadoffRequest
Case GCRHandoff
If($(B_{GCRres} + B_{GCRused} + B_{GCRhand}$
$+ B_{GCRnew} < B_{GCRtotal})$ & $(R_{GCR} = \text{constant})$ &
$(L_{GCR} < 10^{-6})$ **) then**
Admit Handoff
Else
Reject Handoff
End if
Case GMRHandoff
If($(B_{GMRres} + B_{GMRhand}$
$+ B_{GMRused} + B_{GMRnew} < B_{GMRtotal} + B_{GCRtotal})$ **&** $(L_{GMRmin} > L_{Nave})$**) then**
Admit Handoff
Else
Reject Handoff
End if

Case ADSHandoff
If$(B_{ADSres} + B_{ADSused} + B_{ADSnew} + B_{ADShand} < B_{ADStotal})$ **then**
Admit Handoff
Else
Reject Handoff
End if
Case ALSHandoff
If($(B_{ALSres} + B_{ALSused} + B_{ALSnew} + B_{ALShand} + B_{GMRtotal} + B_{GCRtotal} + B_{ADSused} < B_{LStotal})$ &
$(L_{ALSmin} > L_{Nave})$ **) then**
Admit Handoff
Else
Reject Handoff
End if
End Select

Figure 4.5: Admission algorithm for handoff calls only

If there is a handoff request for any class, the network will check whether the admission criteria for that class can be met e.g. if the handoff request is for a GCR class the system will use the admission criteria for a GCR class as described in section 4.3.1. If the admission criteria is met, the network is going to accept that handoff call. Otherwise the incoming handoff call will be rejected, if the handoff request is for a GMR class the system will use the admission criteria for a GMR class as described in section 4.3.2. If the admission criteria is met, the network is going to accept that handoff call. Otherwise the incoming handoff call will be rejected, if the handoff request is for a ADS class the system will use the admission criteria for a ADS class as described in section 4.3.3. If the admission criteria is met, the network is going to accept that handoff call. Otherwise the incoming handoff call will be rejected, if the handoff request is for a ALS class the system will use the admission criteria for a ALS class as described in section 4.3.4. If the admission criteria is met, the network is going to accept that handoff call. Otherwise the incoming handoff call will be rejected (Figure 4.5).

4.4.3 Admission Algorithm for New Calls and Handoff Calls

When a new call request arrives at the network, the network will check whether there are any handoff calls requesting some services. If there are handoff calls requesting services, the network will give priority to handoff calls and the request is processed (Figure 4.5), otherwise, the new request will be granted (figure 4.4). The pseudo code is presented in figure 4.6

```
Algorithm:        RS CAC

Purpose:     This algorithm prioritises guaranteed service classes over other
service
                  classes. It also prioritises handoff calls of any service class
over
                  new calls of any service class.

Pre-condition:     The network meets the admission criteria for each service
class

Process Body:  If request = New then
                  Check if there are any Handoffs requesting
                    If Handoff = True Then
                       Process Handoff
                     Else
                       Process New
                     End if
                 Else
                    Process Handoff
                 End if

Post-condition:  Class admitted/rejected
```

Figure 4.6: Pseudo code for new calls and handoff calls admission

4.5 Simulation Model

In the evaluation of the proposed CAC scheme, a simulation was carried out. The simulation model used in this dissertation assumed that enough resources have been reserved for handoff calls and for guaranteed services' new calls according to the guaranteed services admission criteria.

The simulation network consisted of 7 wireless cells as shown in Figure 4.2. In the simulation cell 1 is considered the current cell where the MN resided and where new calls were generated. All the surrounding 6 cells (cell 2 to 7,) were generating handoff calls and new calls to this centre cell. We adopted a hexagonal cell layout in our model (Figure 4.2), because this is the cell layout that is used in current networking environments [33] to represent a cell. In the simulated CAC algorithm RS service classes were used. Both the GCR and GMR service classes are used as guaranteed services, while ADS and ALS are used as assured services, and BE as best effort class stand on its own.

The percentage of total calls accepted according to their service classes over the total load that was put on the network was monitored. The total load was defined as the total number of calls that were accepted and rejected for each service class. The percentage of handoff calls rejected due to insufficient amount of available resources in the target cell and the percentage of overall handoff calls and overall new calls rejected due to insufficient amount of available resources in the target cell were also monitored.

4.6 Simulation Design and Implementation

All classes that make up our simulation were written in the Visual Basic.net 2003 programming language using Microsoft Visual Studio.net 2003 environment. Visual Basic.net is chosen over other programming languages simply because Visual Basic.net is an Object Oriented Programming (OOP) Language. A Personal Computer that runs Windows XP operating system was used to run our simulator. The simulator had a visual component that allowed the experimenter to continuously view the network for the duration of the simulation. The first step of identifying objects and classes began by generating a set of candidate classes and objects using the classical approach and behaviour analysis. Abstractions that relate to system function points revealed during behaviour analysis are shown in Table 4.1.

Table 4.1: Abstractions that relate to system function points

Behaviours
1. The system checks if a call is a new call or a handoff call. Prioritise handoff call over new call. 2. If new call the system checks the type of class that is making a call. Prioritise guaranteed service classes over low priority classes. 3. The system checks if there are enough resources for the call to be accepted. If enough then accept else reject.

Looking at the objects that we implemented we find that the base station (BS) was determined to be the main object. The BS is responsible for monitoring all the MNs that arrive in its cell, the QoS requirements of the MN's services and initiates the MN's handoff if required. A class diagram showing the components that interact in the simulator and the user interface of the simulator are shown in Appendix A.

4.7 Limitations of the simulator

The limitations of the simulation are as follows:

1. We only used random numbers to test our algorithm; we did not use the sequential order.
2. We could not show the transmission of calls, we only showed the type.

CHAPTER FIVE

5.0 SIMULATION RESULTS AND ANALYSIS

5.1 Introduction

This chapter presents four simulation experiments and their results. The first three experiments evaluate the performance of the CAC algorithm, while the fourth one is used to compare the call admission scheme with another.

5.2 Simulation Experiments

In our simulation we reserved some bandwidth by setting a threshold for ADS and ALS type of calls and we set a higher threshold for QoS sensitive type of calls. We set a higher threshold for new calls of guaranteed service because we want to minimise the rejection rate of this type of calls. We also set a certain threshold for handoff calls of any type of service class because we believe that handoff dropping rate is to be minimised regardless of the type of service class.

5.2.1 GS acceptance probability

In this experiment, the acceptance probability of a class is defined as the acceptance percentage of that class. If the acceptance percentage is high, say above 60%, that class is set to have a high acceptance probability.

5.2.1.1Test

In this test, the rate at which calls arrived, was varied starting from two calls per second (2calls/sec) to ten calls per second (10calls/sec) allowing the simulator to run for five seconds for each number of calls per second. We chose to start with 2calls/sec because we wanted to monitor the acceptance of GS service classes of the proposed CAC algorithm with at least two calls in each second for five seconds. We then collected some statistics showing the functioning of the proposed CAC algorithm. The data that we collected from the statistics is represented in Table 5.1.

5.2.1.2 Results

Figure 5.1 shows the functioning of the proposed CAC algorithm for guaranteed service calls accepted (GCR + GMR), the assured service calls accepted (ADS + ALS) and the best effort service calls accepted (BE). (Note that this test includes both new calls and handoff calls for each service).

Table 5.1: Calls accepted according to service classes vs. Total workload for each class.

Service Class	Call Arrival rate				
	2	4	6	8	10
Guaranteed	100	100	87	85	90
Assured	100	100	50	45	42
Best Effort	100	75	32	17	13

We evaluated the performance of the proposed CAC algorithm by varying the arrival rate of total calls.

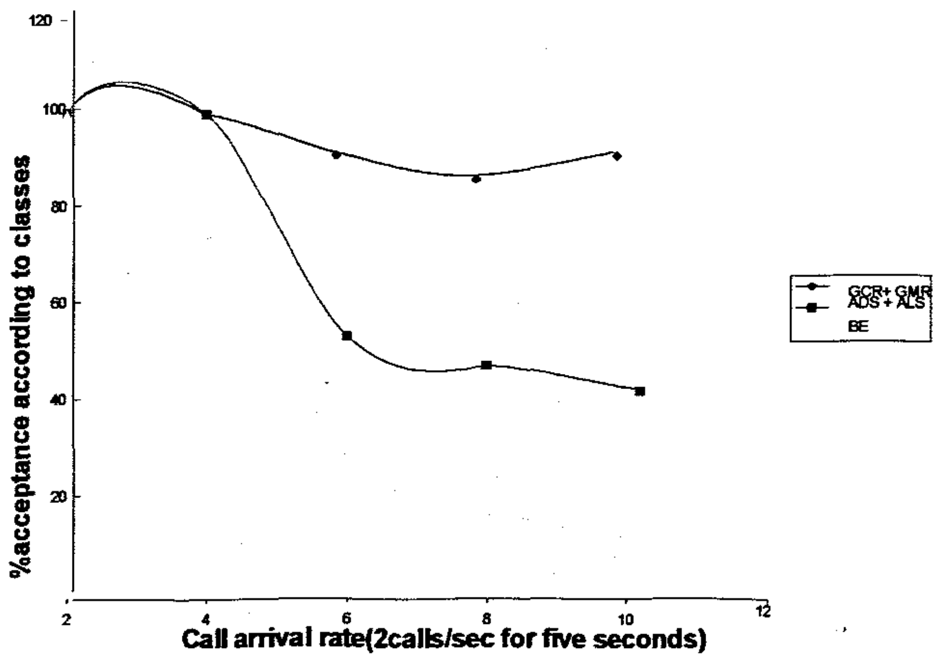

Figure 5.1: Accepted class of calls vs. Total workload

GCR and GMR classes were given higher bandwidth of 25kbps and 20kbps respectively for each class. We set 15kbps, 10kbps, and 5kbps for ADS, ALS and BE respectively. When the user started the simulator, the mobile user started sending different types of calls. The network then treated the calls according to the availability of bandwidth and according to the admission criteria for each class of call as was discussed in section 4.4. The following formula was used in calculating the percentage of acceptance for each class. This is a generic formula where (GCR+GMR) can be replaced with (ADS+ALS) or (BE) for the other groups.

$$\%CallsAccepted(GCR+GMR) = \frac{(GCR+GMR)accepted}{Total(GCR+GMR)workload} \quad \textbf{5.1}$$

5.2.1.3 Analysis

When we analyse the results that we obtained in figure 5.1 for the call arrival rate from 2 calls/sec to 4 calls/sec, we see that GS calls and AS calls have high acceptance rate compared to BE calls. The explanation for this is that few calls were generated and those calls consumed only a portion of bandwidth and not all the network bandwidth. As GS and AS calls have high acceptance rate, BE shows a drastic drop from 2 calls to 6 calls per second. This is due to the high acceptance of high priority classes. The proposed CAC algorithm allows BE calls to use the available bandwidth despite the type of class the bandwidth is reserved for. Before a call arrives, BE uses the call's bandwidth. When the call arrives, BE is dropped to accommodate the incoming call. As the call arrival rate increases from 4 calls per second to 8 calls per second we observe a drastic drop again for all the calls. This is caused by the load increase. This causes all the calls to consume their reserved bandwidth and all the incoming calls are rejected until there are released channels. Despite the load increase, GS remains at an average acceptance rate of 92%.

It can be concluded that the proposed CAC algorithm is treating calls of guaranteed service classes with higher priority, compared to other classes.

5.2.2 Handoff acceptance probability

5.2.2.1 Test

We continued evaluating the performance of RSCAC algorithm by comparing the acceptance of handoff calls with the acceptance of new calls.

5.2.2.2 Results

The data that we gathered is represented in table 5.2. Figure 5.2 is the graphical representation of the data in table 5.2.

Table 5.2: Handoff and New Calls accepted vs. Total workload for each.

Type of Call	Call Arrival rate				
	2	4	6	8	10
Handoff	100	100	95	95	91
New	75	75	76	79	50

The formula that was used in this test is shown below:

$$\%CallsAccepted(New) = \frac{NewAccepted}{Totalworkload(Newcalls)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \mathbf{5.2}$$

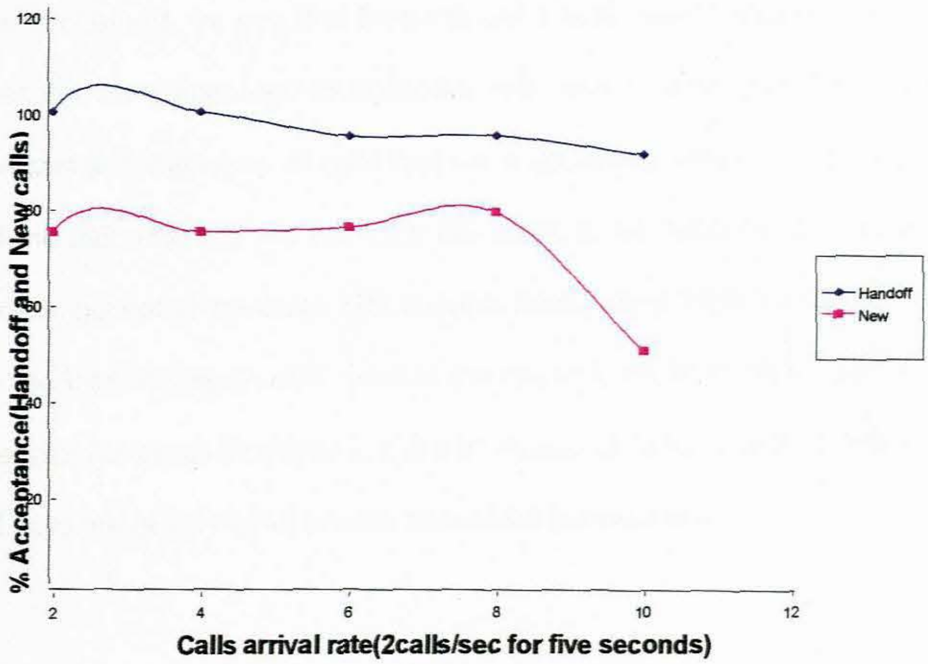

Figure 5.2: % of accepted calls (New & handoff) vs. Total workload for each category (handoff or new)

The performance of this test was evaluated by using the bandwidth provision that was explained above in section 5.2. At the start of the simulator, the mobile user started sending calls randomly (both new and handoff calls). The receiver network then accepted or rejected the incoming calls according to the admission criteria discussed in section 4.4.

5.2.2.3 Analysis

When we analyse the performance of our CAC algorithm as far as handoff calls and new calls are concerned, we see that from the call arrival rate of 4calls per second to 6calls per second the percentage of accepted calls start to decrease. This is due to the load increase and the types of calls that were accepted. When the load increase and most of the call requests are from the GS service, we expect a decrease in the number of calls accepted because GS service demands a high bandwidth on the network. As the load increases and most of the request are from other classes other than GS, we experience an increase in the percentage of calls accepted. We observe that handoff calls are given higher priority compared to new calls.

5.2.3 Overall calls accepted

5.2.3.1 Test

We then carried out the final test to evaluate the overall functioning of the proposed CAC algorithm. We evaluated the overall percentage of calls accepted. We categorised calls into accepted and rejected calls. We used the same call arrival rate of 2calls/sec in our tests.

5.2.3.2 Results

Table 5.3 shows the set of data that we obtained from our test and figure 5.3 depicts a graphical representation of the results found.

Table 5.3: Overall Calls accepted vs. Total workload

All Calls	Call Arrival rate				
	2	4	6	8	10
Accepted	90	85	84	87	68

5.2.3.3 Analysis

When we analyse the performance of our CAC algorithm for overall calls accepted (figure 5.3), we see that the overall percentage of calls accepted decreases as the load increases. This is due to the fact that more and more guaranteed service calls get rejected when the maximum bandwidth assigned to this service is used up.

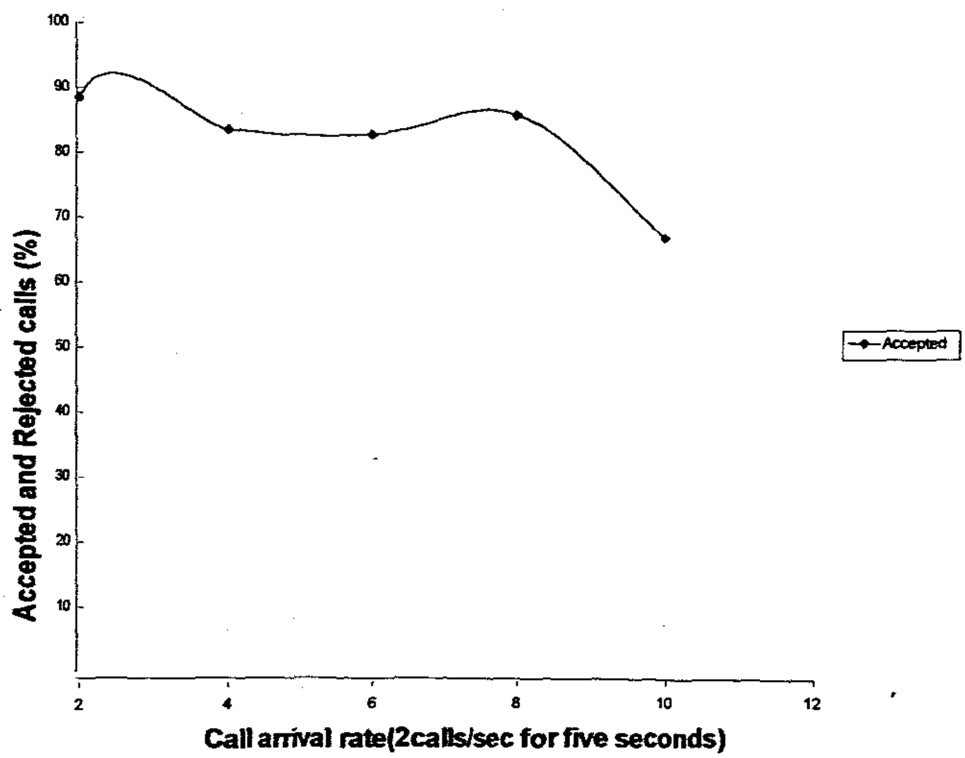

Figure 5.3: Overall Calls accepted vs. Total workload

5.3 Comparison Experiments

5.3.1 Test

We finally RSCAC algorithm with the algorithm proposed by Javenski et al. in [14]. In the scheme proposed in [14], only two types of calls were used and these are CBR and VBR real time classes. These classes are equivalent to RS's GCR and GMR real-time classes respectively. We referred to CBR and VBR real time classes in [14] as ToniGuaranteed.

5.3.2 Results

Table 5.4 gives the set of data that was obtained through test that was conducted. Figure 5.4 is the graphical representation of the test.

In the scheme in [14], on-going calls are prioritised compared to new calls. Call dropping probability is minimised by increasing the call blocking probability. This scheme does not tell us the scenario where the incoming call is CBR or VBR and the on-going call is BE.

5.3.3 Analysis

We conducted the performance analysis by using the same call arrival rate of 2calls/sec. We only tested the performance of the proposed CAC algorithm (RSCAC) by using the RS's GS calls (GCR and GMR) and Toni's GS calls (CBR and VBR). We compared the performance of the two schemes (RSCAC and CAC in [14]). It can be seen that RSCAC guaranteed service calls show a high rate of acceptance compared to Toni's guaranteed service classes.

Table 5.4: Accepted RSguaranteed and Toniguaranteed calls vs. Total workload

Service Class	Call Arrival rate				
	2	4	6	8	10
RSGuaranteed	100	95	80	82	89
ToniGuaranteed	100	90	72	75	76

This is caused by the way Toni treats the guaranteed service classes.

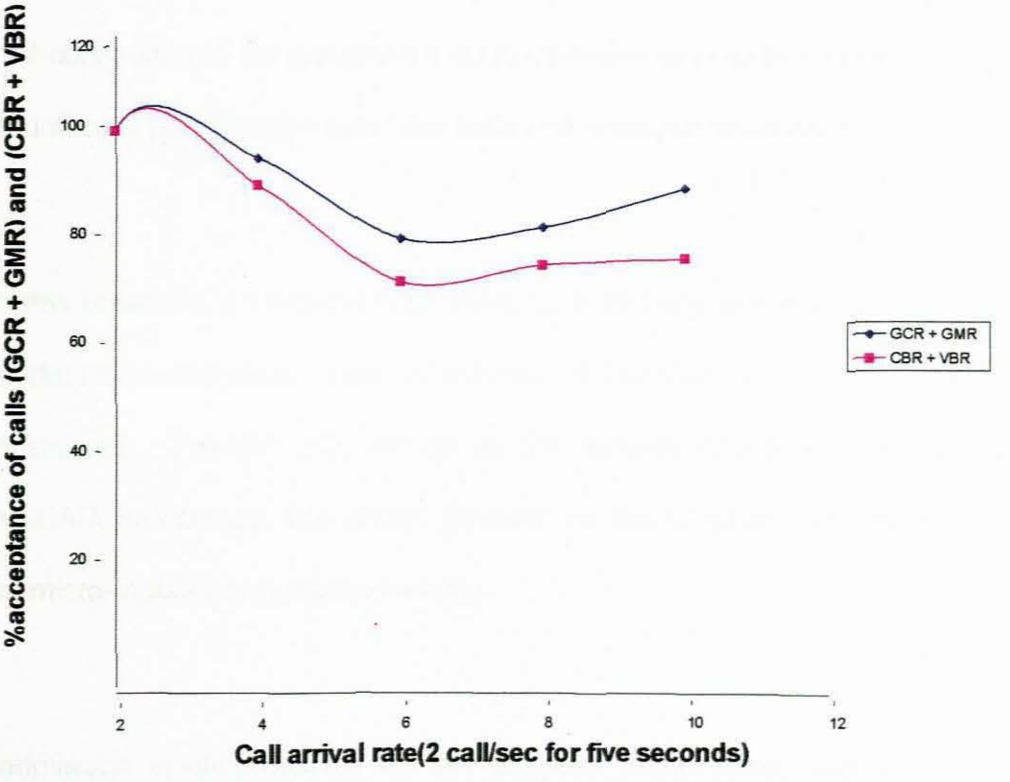

Figure 5.4: Accepted RSguaranteed and Toniguaranteed calls vs. Total workload

Looking at the graph, it can be seen that RSCAC has a marginal performance compared to that in [14].

CHAPTER SIX

6.0 CONCLUSION

6.1 Summary

In mobility management for wireless networks it is fundamental to consider the QoS needs for mobile users and how these can be provisioned. This incorporates a number of constraints to be provisioned such as real-time prioritisation over non real-time, handoff calls prioritisation over new calls and resource reservation.

Through this research, an architectural basis for a mobility management strategy for real-time data communication was established. A number of mobility architectures were considered. TeleMIP [20] served as the mobility architecture on which the proposed CAC was based. We chose TeleMIP as the base architecture because it caters for micro-mobility and macro-mobility.

A call admission control criteria for the adopted architecture, was defined. The admission control criteria that was defined was taken from the QoS constraints for each service class in RS [31]. RS defines QoS constraints for bandwidth allocation for each service class. Using those QoS constraints, the admission control criteria for RSCAC was defined. We then proposed a call admission control algorithm using the defined call admission criteria. The proposed call admission control algorithm prioritises handoff calls over new calls and GS classes over other classes. A call is only admitted on the network if its QoS constraints can be met according to its corresponding class' admission criteria.

A simulated performance evaluation of the strategy was developed where we conducted three tests. We finally conducted a test to compare the performance of RSCAC to that in [14]. The RSCAC prioritises guaranteed service calls over other calls. It also allows handoff calls to get high rate of acceptance compared to new calls.

Furthermore, in our scheme, guaranteed service calls received a high rate of acceptance compared to the equivalent scheme in [14]. We can, therefore, conclude that RSCAC can be suitable to be employed in real life scenarios because we have categorised the calls into real-time and non real-time types of classes and obtained a better performance as shown in figure 5.1.

6.2 Future Work

In our simulation we only tested a few parameters. These were: calls accepted over total workload, calls rejected over total workload, calls accepted according to class of service over total workload for each class, and handoff calls prioritisation over new calls. We understand that these are not the only parameters we can use to declare our CAC algorithm as performing. The performance of our CAC algorithm could also be tested in a case where there are no resources readily available. We also could test many parameters that exist in the literature, e.g. channel holding time for each class over speed of a mobile user. The architecture that we adopted is a micro-mobility architecture that uses Mobile IP for macro-mobility. In our simulation we simulated micro-mobility only.

This research work has the potential of being extended. It is recommended that if one develops an interest in taking this study further, the remaining parameters and QoS constraints be tested.

REFERENCES

[1] The Internet Protocol, RFC 791, www.ietf.org/rfc/rfc0791.txt

[2] Mobile IP, RFC 3012, http://www.networksorcery.com/enp/rfc/rfc3012.txt

[3] Perkins C. E, "IP Mobility Support for Ipv4", revised, Internet – Draft, draft – ietf – mobileip – rfc2002 – bis – 01.txt (2002), www.faqs.org/rfcs/rfc3220.htm

[4] Kaloxylos A., Zervas E., Merakos L., Paskalis S., "An Efficient RSVP – Mobile IP Interworking Scheme", Journal for Special Topics in Mobile Networks and Applications (MONET), vol. 8, no. 3, pp. 197-207, June 2003.

[5] Web ProForum Tutorials, "Global System for Mobile Communications (GSM)", http:// www.iec.org, (10 April 2004)

[6] Web ProForum Tutorials, "Wireless Internet Network Communications Architecture", http:// www.iec.org, (21 May 2004)

[7] Perkins C. E, Johnson D. B, "Mobility Support in Ipv6", (draft), August 2003.

[8] Johnson D. B., Perkins C., "*Route Optimization in Mobile IP*", Internet Draft, draft-ietf-mobileip-optim-09,Work in Progress, February 2000. http://citeseer.ist.psu.edu/context/15053/0

[9] Gastsfssor E., Jonsson A., Perkins C. E, "Mobile IP Regional Registration", Internet – Draft, draft – ietf – mobileip – reg– tunnel – 02.txt (2000), http://people.nokia.net/charliep/txt/regreg/regtun.txt

[10] Perkins C. E, Johnson D. B, "Route Optimization in Mobile IP", Internet – Draft, draft – ietf – mobileip – optiml – 09.txt (2002), http://k-lug.org/~griswold/Drafts-RFCs/draft-ietf-mobileip-optim-09.txt

[11] Ojong G., and Takawira F, "*Rainbow Services: A New Architecture for QoS Provisioning in the Future Internet*", SATNAC, 2003, pp. 4 on CD-ROM, September, 2003.

[12] Zhang Y., Liu D., "*An Adaptive Algorithm for Call Admission Control in Wireless Networks*", Proc. IEEE Globecom, pp.3628–3632, Dec. 2001.

[13] Chiu M.H. and Bassiouni M. A., "*Predictive schemes for handoff prioritization in cellular networks based on mobile positioning*",*IEEE Journal on Selected Areas in Communications*, vol.18, pp.510–522, Mar. 2000.

[14] Janevski T., and Spasenovski B., "*Admission Control for QoS Provisioning* in WirelessNetworks",http://www.mnlab.cs.depaul.edu/seminar/spr2003/WiQoS.pdf, (16 January 2005)

[15] Bozinovski M. and Gavrilovska L., "*Distributed Adaptive Admission Control in Mobile Multimedia Network*", http://kom.aau.dk/~marjanb/papers/mmt00.pdf, (10 February 2005)

[16] Lim S., Cao G., and Das C. R., "*An Admission Control Scheme for QoS-Sensitive Cellular Networks*", IEEE, Wireless Communications and Networking Conference, 2002. WCNC2002, pp. 296- 300 vol.1, March 2002.

[17] Lohi M., Baldo O., and Aghvami A. H., "*Handover Issues and Call Admission Control in Cellular Systems*", http://www.comp.brad.ac.uk/het-net/HET-

NETs04/CameraPapers/P44.pdf, (15 October 2004)

18] Mahmod S., Sirisena H., and Pawlikowski[2] K., "*Call admission Control for Mobile Multimedia Wireless Networks*", Proceedings of ICT 2002, Beijing, June2002, http://www.elec.canterbury.ac.nz/research/networking/documents/Mahmod_Sirisena_Call_Admission.pdf

[19] Tseng Y., Wu H., and Hsieh M., "*Connection Admission Control for QoS Guarantees in Mobile Networks*", IEEE, Computer Communications and Networks, 1999. Proceedings. Eight International Conference, pp. 542-547, November 1999.

[20] Das S., Misra A., Agrawal P.and. Das S.K, "TeleMIP: Telecommunication Enhanced Mobile IP Architecture for Fast Intra-Domain Mobility", IEEE PCS Magazine, vol. 7, no. 4, Aug. 2000, pp. 50-58.

[21] Misra A., Das S., McAuley A., Dutta A.and. Das S.K, "IDMP: An Intra-Domain Mobility Management Protocol using Mobility Agents", Internet draft, draft-mobileip-misra-idmp-00.txt, July 2000. Work in Progress, ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=989774&isnumber=21325

[22] Ramjee R. et al., "IP Micro-mobility support through HAWAII", Internet Draft, draft-ietf-mobileip-hawaii-01.txt, July 2000. Work in Progress, http://citeseer.ist.psu.edu/context/1115488/0 "Using not through"

[23] El-Malki K.and Soliman H., "Fast handoffs in Mobile Ipv4", Internet Draft, draft-elmalki-mobileip-fast-handoffs-03.txt, Sept. 2001. Work in Progress,

http://mirrors.isc.org/pub/www.watersprings.org/pub/id/draft-elmalki-mobileip-fast-handoffs-03.txt

[24] Valko A.G., "Cellular IP: A New Approach to Internet Host Mobility", Comp. Commun. Review, vol. 29, no. 1, Jan. 1999, pp. 50-65.

[25] Perkins C.E. and Johnson D. B., "Route Optimization in Mobile IP", Internet Draft, draft-ietf-mobileip-optim-11.txt, Sep. 2001. Work in Progress, www3.ietf.org/proceedings/02jul/ I-D/draft-ietf-mobileip-rfc3012bis-03.txt

[26] Gustafsson E., Jonsson A., and Perkins C. E., "Mobile Ipv4 Regional Registration" Internet Draft, draft-ietf-mobileip-reg-tunnel-06.txt, March 2002. Work in Progress, http://citeseer.ist.psu.edu/context/1914730/0

[27] Zhang L., Deering S., Estrin D., Shenker S., and Zappala D., "*RSVP: a new resource ReSerVation Protocol*," IEEE Network, 7(5):8-18, 1993. 175.

[28] Campbell A.T., Gomez J., Kim S., Valko A. G., and Chieh-Yih Wan, "Design Implementation and Evaluation of Cellular IP", IEEE Wireless Communication Magazine, vol. 7, no. 4, Aug. 2000, pp. 42-49.

[29] El-Malki K. and Soliman H., "Fast handoffs in Mobile Ipv4", Internet Engineering Task Force draft-ietf-mobileip-fast-mipv6-00.txt, February 2001, http://citeseer.ist.psu.edu/context/1914730/0

[30] Misra A., Das S., Datta A., and Das S. K., "IDMP-*based fast handoffs and paging in IP-based 4G mobile networks*," IEEE Commun. Mag., vol. 4, no. 3, pp. 138–145, Mar. 2002.

[31] Ojong G., and Takawira F, "*Bandwidth Allocation for the Rainbow Services Network*", AFRICON, 2004, 7th AFRICON Conference in Africa Volume 1, Issue , 15-17 Sept. 2004 Page(s): 285 - 288 Vol.1

[32] Helmy A., Jasseemuddin M., Bhaskara G., "Efficient micro-mobility using intra-domain multicast-based mechanisms (M&M)", ACM SIGCOMM Computer Communication Review Volume 32 , Issue 5 (November 2002) COLUMN: Technical papers Pages: 61 - 72 , 2002.

[33] "Wireless Linear network", www.afar.net, (20 February 2005)

[34] "Overview of GSM", www.iec.com, (02 April 2004)

[35] "Overview of GPRS", www.cisco.com, (15 May 2004)

[36] "Existing Technology of Cellular Networks", www.doc.ic.ac.uk, (26 June 2005)

APPENDIX A

A-1 Class Description

We used eleven classes as shown below (Figure 4.5).

1. **The Cell:** monitors the MNs resident within it.
2. **The Network Class:** contains Cell objects. Seven hexagonal cells were used to construct the network.
3. **The Node Class:** This is an abstract class. It was subdivided into a base station, a mobile host, a correspondent host and a Mobility Agent class.
4. **The Mobile Node:** we gave this class an attribute of bandwidth requirements. It moved randomly within the network. A MN was able to receive and make calls while it was roaming from one cell to the next.

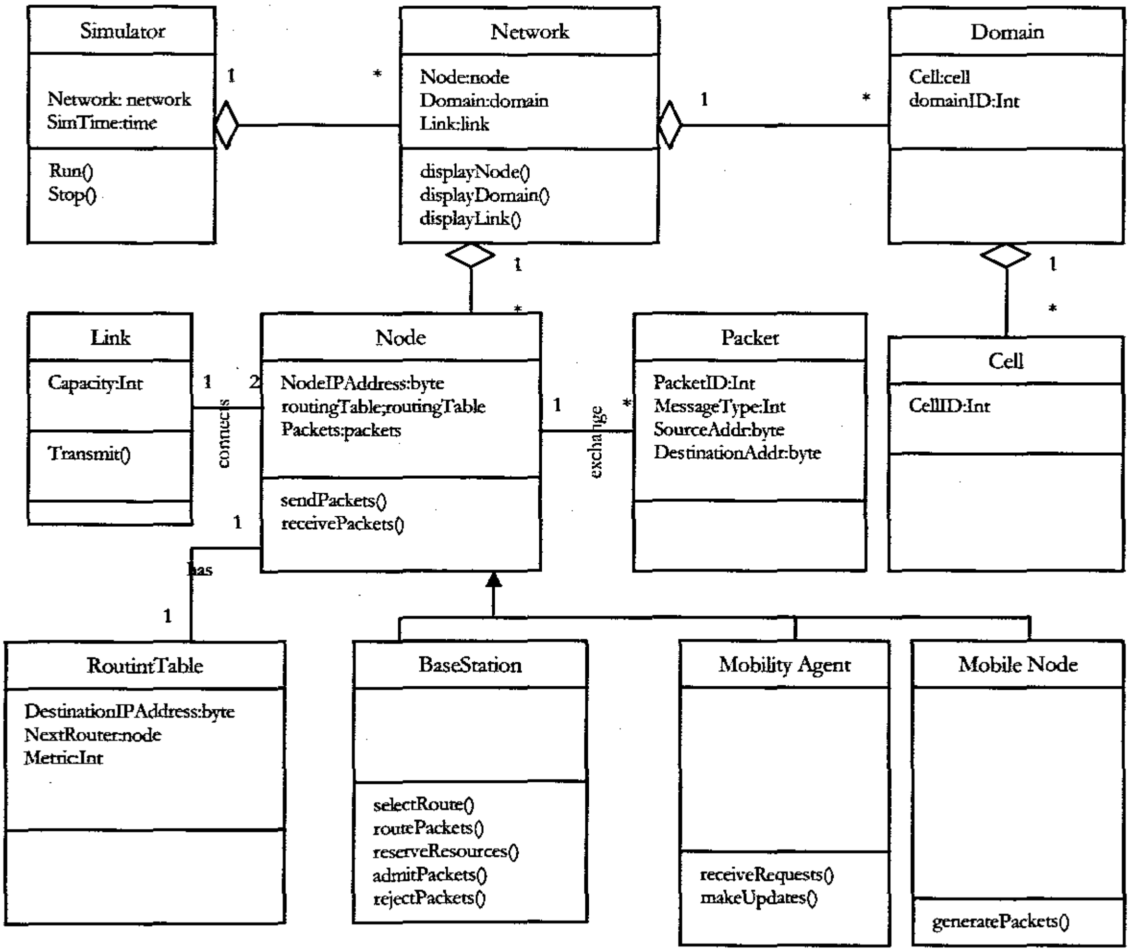

Figure A-1 The System's class diagram

5. **The CorrepondentNode:** This class is the one that send calls to the MN. It also extends the **Node** class.
6. **The BaseStation:** This class is responsible for keeping information about all MNs that enter into its cell. It is also responsible for receiving or rejecting all the calls from a correspondent host to a MN. It class also extends the **Node** class.

7. **The Simulator class:** this class created the objects that participated in the simulation, set them in motion and displayed the simulation statistics.
8. **The Call Class:** this is a simple class that represents a network packet. It only has source and destination address and type (represents QoS requirements) as parameters.
9. **The Domain Class:** this class is a container for cells objects. It is composed of seven hexagonal cells.
10. **The MobilityAgent Class:** this is responsible for monitoring all the BSs that are in its domain. It accepts the Registration requests that are sent by the BSs on the MN's behalf. It is also responsible for HA's binding updates.
11. **The Link Class:** this class is responsible for transmitting wireless signals to and from the MNs.

A-2 User Interface

When we first run our simulator, the simulation time starts at once, but no calls are shown to be made before the user can click the Start button. The initial user interface is shown in figure 4.9. The simulator starts when the user clicks the **Start** button. Once the simulator was running, the network randomly sends calls to the centre cell. We generated a random number from one to ten which will allow the network to generate new calls and handoff calls randomly. We then created another random numbers from one to ten that will

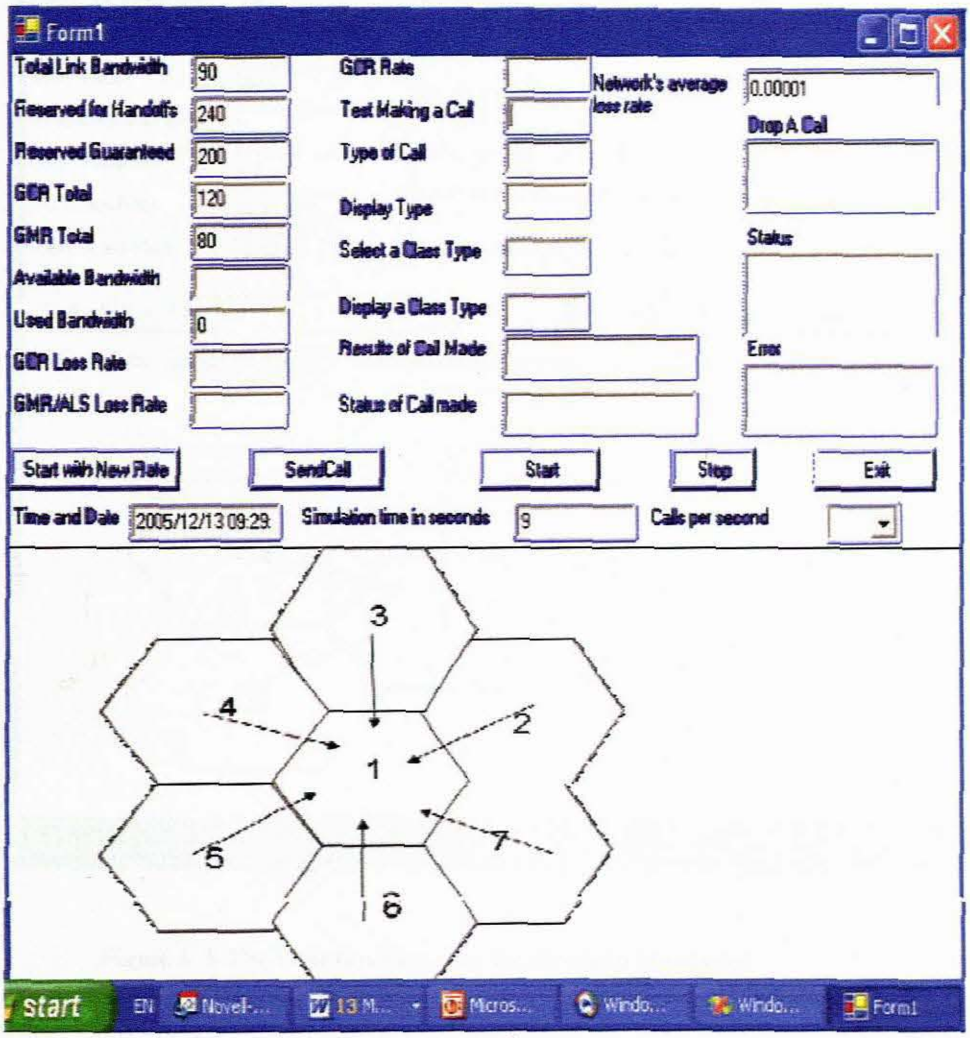

Figure A-2 The Initial User Interface

randomly generate different types of services RS service classes. Each call of a certain service class is accepted or rejected depending on the QoS constraints of each class compared to the QoS that the network can provide at that time. The QoS constraints for each service class

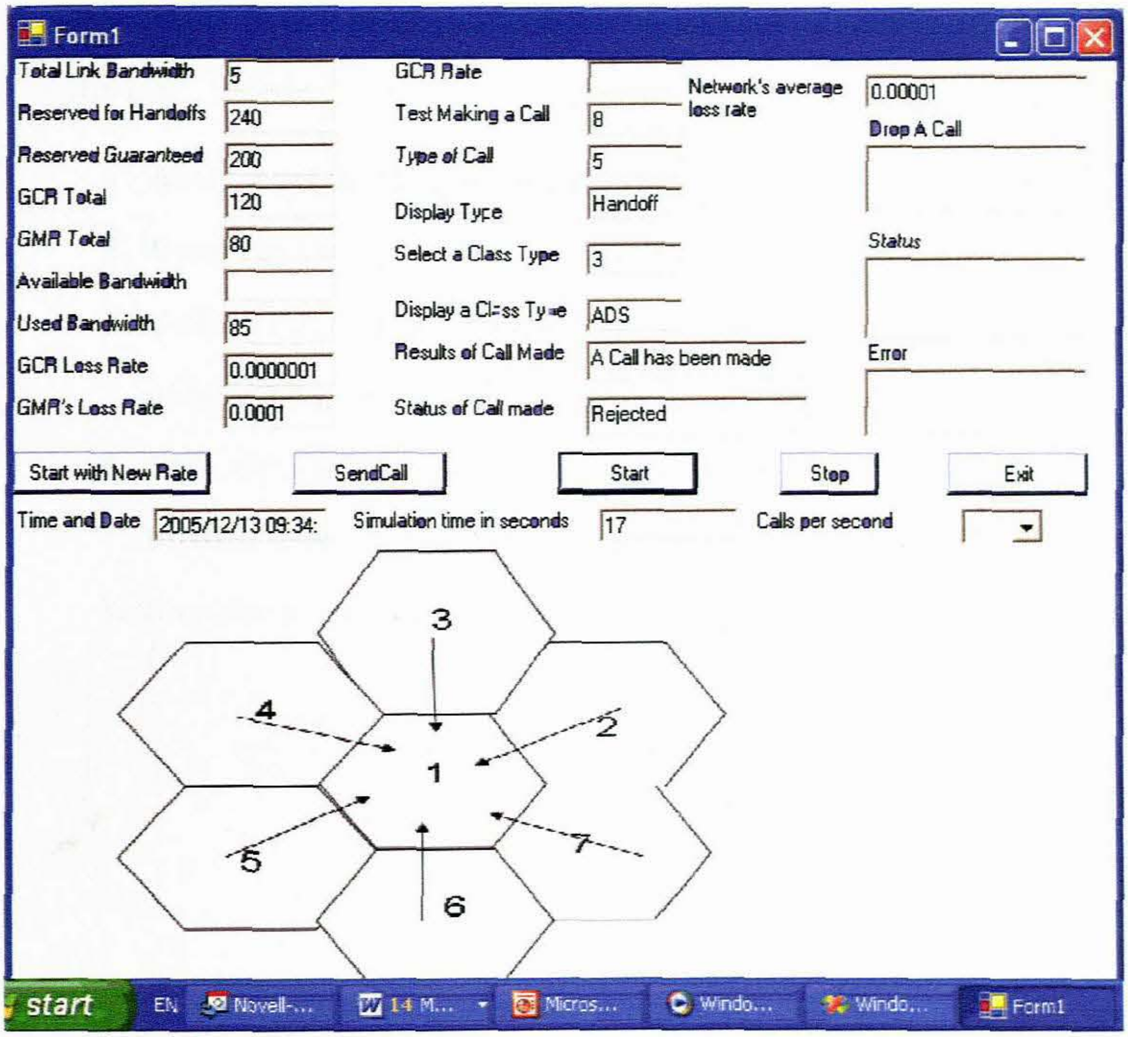

Figure A-3: The User Interface after the simulator has started

are viewed by the network as the CAC algorithm runs. Figure 4.10 shows the user interface for our CAC algorithm where the user has started the simulator by clicking the Start button. The user can also increase the calls arrival rate by selecting from the Calls per second combo box.

The user interactions with the simulation can be summarised as follows:

1. **Start button** for starting the simulation;
2. **Stop** for stopping the simulation;
3. Call arrival rate can be controlled using **Calls per second** combo box and click the **Start with new rate** button;
4. The simulation can be restarted anytime using the **Start button** after it had stopped and
5. Simulation results can be shown by viewing our database.