uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Deepti Dutt

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.C.S.

GRADE / DEGREE

School of Information Technology and Engineering

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Seamless Mobility Across Next Generation Heterogeneous Network

TITRE DE LA THÈSE / TITLE OF THESIS

D. Makrakis

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

Voicu Groza

Ashraf Matrawy

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Seamless Mobility across Next Generation Heterogeneous Network

By

Deepti Dutt

Thesis Submitted to the Faculty of
Graduate and Postdoctoral Studies in Partial Fulfillment
of the Requirements for the Degree of
Master of Applied Science in Computer Science

Ottawa-Carleton Institute for Electrical and Computer Engineering
School of Information Technology and Engineering

University of Ottawa

2009

Your file Votre référence
ISBN: 978-0-494-65481-1
Our file Notre référence
ISBN: 978-0-494-65481-1

ABSTRACT

A vision of future wireless networks is the coexistence technologies used by existing wireless access networks. These networks will be bound together into a single network using Internet Protocol (IP) as the glue. In order to fulfill this vision, one has to address the challenge of supporting mobility across heterogeneous wireless networks. The issues of when and how the handovers are performed while moving from one network to another, affects the performance of the mobile services. While on the move, mobile users experience connectivity disturbances, particularly when they handoff between two access points that belong to the same wireless network and also when they change from one access technology to another. This dissertation addresses the issues of seamless mobility across heterogeneous network and has successfully presented a model to support mobility across homogeneous and heterogeneous environments by integrating two major mobility protocols (MIP and SIP). The functionality of the model was simulated and performance analysis undertaken using the tool OPNET.

ACKNOWLEDGMENTS

The printed pages of my dissertation hold far more than the culmination of years of study. These pages also reflect the relationships with many generous and inspiring people I have met since beginning my graduate work. The list is long, but I cherish each contribution to this success of mine.

First of all to my advisor Dr Dimitros Makrakis, a gracious mentor who patiently guided me, who provided his supervision and expertise without which I would have been lost several times. Thanks also for providing the initial idea for the project, and introducing me to a fascinating subject, and never accepting less than my best efforts.

To my family without them this achievement, would have been impossible today, my dad Mr Ramesh .C Dutt who has been the constant source of moral and technical support, Mom Mrs. Prabha Dutt, Brother & family for their love, support and understanding during the long years of my education.

To my little cousin brothers for sharing their valuable technical knowledge, the late Sandeep Menon who helped me with this dissertation and worked on my document till two weeks before he was completely paralyzed and left all of us after having bravely fought three years against Brain tumor. Mr. Govind Dutt for his valuable suggestions and helping me edit to shape up this thesis.

To my aunt Dr. Maya Dutt for sharing with me, research experience and providing her valuable advices on research and how to write a good research document.

I would like to share my success with my co-researchers and friends, Mr. Gagandeep Singh, all at research lab, Mr. Deepesh Damodaran, Ms. Juby Jose, S, Mr. Rooshi and many more of my friends to whom I would always be indebted for their encouragement and support with various software tools which helped me expedite my work.

I should thank Dr. Yixin Dong, whose work was the platform and starting point for my work and for providing answers to all the queries I had.

To the School of Information Technology and Engineering (SITE), University of Ottawa for providing all assistance to complete my work.

Last but not the least, I thank GOD ALMIGHTY for everything that I have done and am doing.

If I have missed anyone in particular to mention, it was not intentional it is just bad printing.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

1G	First Generation
2G	Second Generation
3G	Third Generation
3GPP	3rd Generation Partnership Project
4G	Fourth Generation
AAA	Authorization Authentication and Accounting
ALG	Application Layer Gateways
AODV	Ad Hoc on demand distance vector
AP	Access Points
ATM	Asynchronous transfer mode
B2BUA	Back-to-Back User Agent
B3G	Beyond 3G
BS	Base Stations
CDMA	Code Division Multiplexing Access
CDPD	Cellular Digital Packet Data
CN	Correspondent Node
CoA	Care of Address
CR-LDP	Constraint-Based Label Distribution Protocol
DAD	Duplicate Address Detection
DHCP	Dynamic Host Configuration Protocol
DNS	Dynamic Naming System
DNS	Domain Name System
EDGE	Enhanced Data Rates for Global Evolution
FA	Foreign Agent
EAP	Extensible Authentication Protocol
FDA	Foreign Domain Agent
FEC	Forwarding Equivalence Class
GPRS	General Packet Radio Service
GSM	Global System for Mobile
HA	Home Agent
HH	Horizontal Handoff
HMIP	Hierarchical Mobile IP
HMIPv6	Hierarchical Mobile IPv6
HSPA	High Speed Packet Access
HTTP	Hypertext Transfer Protocol
IDMP	Intra-Domain Mobility Management Protocol
IEEE	Institute of Electrical & Electronics Engineers
IETF	Internet Engineering Task Force

IMS	IP-Multimedia Subsystem
IP	Internet Protocol
LCA	Loosely Coupled Approach
LDAP	Lightweight Directory Access Protocol
LDP	Label Distribution Protocol
LER	Label Edge Routers
LSP	Label Switched Path
LSR	Label Switching Routers
MANET	Mobile Adhoc Network
MIME	Multipurpose Internet Mail Extensions
MIP	Mobile IP
MIP	Mobile IP
MM	Mobility Management
MM-MPLS	Mico Mobile MPLS
MMS	Multimedia Message Service
MN	Mobile Node
MPLS	Multiprotocol Label Switching
MobServ	Mobile Server
NA	Network Administrator
NAT	Network Address Translation
NAT-PT	Network Address Translation- Protocol Translation
OPNET	Optimized Network Engineering Tool
P2P	Peer to Peer
PC	Personal Computer
PDA	Personal Digital Assistant
PPP	Point-to –Point Protocol
QoS	Quality of Service
RFC	Request for Comments
RSVP-TE	Resource ReSerVation Protocol – Traffic Engineering
RTP	Real Time Protocol
S-MIP	Seamless Handoff architecture for Mobile IP
SDP	Session Description Protocol
SIP	Session Initiation Protocol
SIPRREP	SIP Response
SIPRREQ	SIP Request
SMS	Short Message Service
SMTP	Simple Mail Transfer Protocol
SPS	Synchronized Packet-based Simulcast
TCA	Tightly Coupled Approach
TCP	Transmission Control Protocol
TE	Traffic Engineering
TeleMIP	Telecommunication Enhanced Mobile IP

TTL	Time To Live
UA	User Agent
UAC	User Agent Client
UAS	User Agent Server
UDP	User Datagram Protocol
UMTS	Universal Mobile Telecommunication System
URI	Uniform Resource Identifiers
VH	Vertical Handoff
VoIP	Voice over IP
VPLS	Virtual Private LAN Services
VPN	Virtual Private Network
WCDMA	Wideband Code Division Multiple Access
WHN	Wireless Heterogeneous Networks
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Networks
WMAN	Wireless Metropolitan Area Networks
WPAN	Wireless Personal Area Networks
WWAN	Wireless Wide Area Networks

CHAPTER 1
INTRODUCTION

In recent times, advances in technology have made wireless networks very popular. With the ever-increasing use of mobile devices, the demand for node mobility and support for various mobile applications are also growing at an exponential rate. Apart from mobile phones and PDAs, people have started using portable laptops to access Internet and other resources using wireless networks while on the move. For any MN, if it has to take full advantage of their mobility feature provided, the MN should be able to move to any point in the network and still retain all its ongoing communications.

The movement of the MN can be between the access points where each access point may belong to the same subnet, same domain, it could also be different subnets or different domains altogether. In most cases the end-client would have access to both networks at the same time, but connectivity to the network would be determined by any local policy defined in the client itself such as signal strength or any other measurement based on QoS (Quality of Service) parameter of the traffic.

Mobility management across the network plays a very important role since the user could be moving between multiple types of access networks involving different service providers during a multimedia session. These access networks could be IEEE 802.11b, CDPD, CDMA or GPRS based network supporting DHCP or PPP servers in the networks. Third generation systems (3G) have been designed to provide Internet connectivity to the end user across different networks with support of various IETF protocols. Concurrently, as IP routing capability is being pushed to the edge of the networks, it is inefficient to manage mobility at layer 2 as cellular networks currently do. An integrated solution, where Mobility Management (MM) is an integral part of the IP layer routing protocol, would be more advantageous for both the operator and the end user.

1.1 Thesis Outline

- Chapter 1 presents the motivation, thesis organization and its contributions.
- Chapter 2 outlines the background and objective of the thesis work. It presents the typical requirements of an integrated seamless handoff mobility model for the 4th Generation (4G) network, Introduces the concepts of seamless mobility over heterogeneous network; provides an insight into the related research work that has been carried out. At the end a bird's eye view is provided on the design solution proposed in this thesis.
- Chapter 3 is a short tutorial on the technologies used for developing the proposed model.
- Chapter 4 provides an introduction on Session Initiation Protocol. And also discusses the possibility of deploying SIP to support mobility
- Chapter 5 elaborates on the design and implementation of the model proposed in this dissertation.
- Chapter 6 describes in brief the related work published toward deployment of SIP for next generation network mobility and also provides results for the validation of the model presented in this dissertation.
- Chapter 7 presents the simulation results and the performance evaluation of the prototype.
- Chapter 8 presents the conclusion and an insight into possible area that could be explored as a part future work proposed on this thesis.

1.2 Thesis Contribution

The major contributions of this thesis are the following:

- An in-depth study was undertaken on the major problem encountered during the mobility across heterogeneous networks.
- Multilayer mobility management architecture to support mobility across next generation heterogeneous network has been proposed.
- The proposed architecture was developed to support both real time and TCP connections
- Performance of the proposed model is studied through simulation developed with OPNET simulation tool

CHAPTER 2
BACKGROUND AND MOTIVATION

As wireless communication technologies evolve at a mind-boggling pace, the recent focus of research has shifted focus to the development of fourth-generation (4G) mobile systems. Instead of developing a new uniform standard for all wireless communications systems, 4G communication networks strive to integrate various existing wireless communication technologies seamlessly. The structure of future wireless networks is a coexistence of current access networks, which will be integrated into a single network using the Internet Protocol (IP) as the glue. To fulfill this vision, one of the most challenging steps is to develop a model for seamless mobility across the network.

In Section 2.1 we put forth the typical requirements for supporting mobility over heterogeneous networks; Section 2.2 provides a background study, which involves an insight into the evolution of wireless technology, an overview of heterogeneous networks and a study of handovers; Section 2.3 provides a definition of seamless mobility; Section 2.4 we study to some detail, the challenges involved in supporting mobility across heterogeneous network. In Section 2.5 we perform a study on the related mobility supporting protocols and Section 2.5 describes the problem domain. Finally in section 2.6 we present the problem domain which would be addressed in this thesis

2.1 Typical Requirements of Mobility across Heterogeneous Networks

- When an interface loses contact, re-establishment attempts should be performed automatically.
- A handover between different network interfaces could occur due to various reasons, such as poor signal quality or a user wanting more facilities and thus moving to a different network. In such cases, the handover must be transparent to the user's
- The inter-domain handoff should be fast so that the user moving across the network will not be aware of the change. Handover delay and packet loss should be minimized, especially in real time applications.

- If a situation occurs wherein the session is not maintained throughout mobility or is interrupted by varying network conditions, the session will have to be terminated and be re-initiated if the user is interested in continuing the session through a different interface. This termination and re-establishment would defeat the whole objective of the next generation of seamless handoffs across heterogeneous networks. Logically, the aforementioned process needs to take place when the MN moves from one network to another and the requirement is to keep this process transparent to the user.

2.2 Background

2.2.1 Wireless Generations at a Glance

Telecommunication and Internet technologies have undergone rapid developments during the past decade. Focus of the first and second generations of cellular networks [1] was on providing voice communications, essentially for wireless and mobile telephones. The original analog cellular systems developed in the early 1980s are considered the first generation of mobile telephony (1G). Around the same period, the cellular industry also began developing the second generation of mobile voice communication (2G), which used digital signaling instead of the analog signaling used by 1G. It took about 10 years for 2G mobile phone systems to be developed and this generation was deployed in the early to mid 1990s. The most popular second-generation cellular system became the Global System for Mobile communication (GSM) [1] and continues to have great popularity on a global basis.

The growing interest in non-vocal communications slowly introduced services such as the Short Message Service (SMS) [17] and Multimedia Message Service (MMS) [2]. The evolution of General Packet Radio Service (GPRS) and Enhanced Data rates for Global Evolution (EDGE) under GSM technology was the other major developments in the telecom world during the nineties. GPRS and EDGE were classified as “2.5 G Wireless”. However, both relied on GSM’s infrastructure, which was initially designed for voice communication. This resulted in poor performance of data transmission and also relatively inefficient use of a new technology capable of multimedia communications over cellular devices (e.g. watching television shows or video conferencing).

The technology developed as result became the corner stone of 3G technology and was called Wideband Code Division Multiple Access (WCDMA). It consists of 2 competing standards. WCDMA and CDMA2000 championed by QUALCOM. WCDMA is the most popular one and appears to be the winner. Work on the development of 3G concepts started around 1991; just about the time when 2G systems were getting deployed. The most out-standing feature of 3G is that, both voice and data are supported intrinsically. 3G systems have been enhanced further through advances such as High Speed Packet Access (HSPA), the outcome of which was the evolution of "3.5G Wireless" technology.

The advances in technology after the third generation aim to provide extended mobility with optimized data rates and services. Mobile users should be able to enjoy higher levels of flexibility while using multi-service networks that provide services such as seamless connection to the Internet via heterogeneous networks, advanced spatial location and navigation services, and true IP based real-time multimedia. Providing end-to-end optimization is one of the key challenges in future network management. This takes into account variables such as throughput optimization, routing optimization, delay profiles for heterogeneous wireless environments and economical profitability. The time is ripe for the emergence of the next generation networks and services beyond 3G. These systems are called B3G (beyond 3G) or 4G. They will make heavy use of heterogeneous networking technologies. Based on the previous experiences of requiring about 10 years to develop new mobile systems; 4G could become operational by around 2011. This technology is expected to build on the second phase of 3G, with all networks embracing Internet protocol (IP) technology.

2.2.2 Heterogeneous networks

The term *heterogeneous* is defined as referring to element that are not of the same kind or nature [3]. Applying this definition to the term *network*, it can be said that a heterogeneous network is one that consists of multiple network technologies. A homogeneous network, in contrast, is a network that consists of a single network technology [4].

General visions of 4G wireless networks are essentially the future of Wireless Heterogeneous Networks (WHN). A WHN is made up of multiple wireless access technologies. Each of these technologies has its own characteristics with respect to coverage, QoS assurance, implementation, operational costs, supported features, etc. Presently, heterogeneous environments are expanding and mobile devices often have built-in support for multiple network interfaces. Based on the size of the network, Heterogeneous wireless access networks can generally be classified into the following categories.

- **Wireless Personal Area Networks (WPAN):** WPANs are wireless networks that cover a limited range of geographical area. For example, Bluetooth can provide an ad-hoc wireless network [5] over a limited area of up to 10 meters.

- **Wireless Local Area Networks (WLAN):** WLAN Network covers a range of 50 to 300 meters. This type of network is typically used in buildings such as offices, schools etc. WLAN can provide wireless Ethernet access without expensive infrastructure, e.g. IEEE 802.11b [8].

- **Wireless Metropolitan Area Networks (WMAN):** WMAN's are wireless networks that cover a geographic area such as a city, e.g. WiMAX [8].

- **Wireless Wide Area Networks (WWAN):** WWANs are wireless networks that extend over a large geographical area, e.g. UMTS.

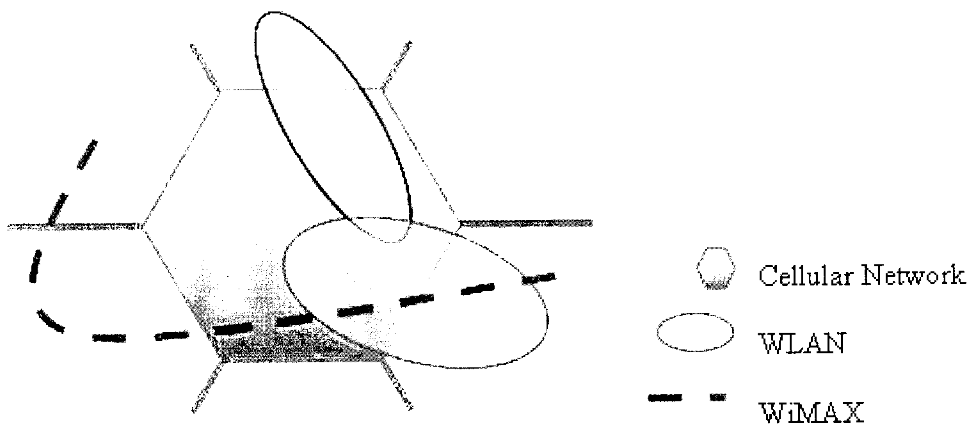

Figure 2-1: Overlapping of 3 Wireless Technologies [8]

WHN coverage comprises of wireless overlays, i.e. coverage overlaps different technologies. An example of a wireless overlay is shown in Figure 2-1. WHN enables users to select the access technology that is most appropriate for their terminal capabilities, application requirements and service cost. Users will have the flexibility to connect simultaneously to more than one access technology.

With this capability, a MN can initiate connectivity through the technology that most closely matches the user's or the application's requirements. However, if varying network conditions interrupt the session, the session will have to be terminated and re-initiated if the user wants to continue through a different interface. The following section discusses further details regarding mobility across different networks in WHN.

Handover (or handoff) refers to the event that occurs when a MN changes its point of attachment from one access network to another. Handovers can be either horizontal or vertical (Figure 2-2). Handovers are said to be seamless if the handover is transparent to the user of the MN [8]. This capability is called inter-technology (inter-domain) handoff or Vertical Handoff (VH). The term "horizontal handover" refers to a handover-taking place in a homogeneous network. For example, a terminal changing associations between two 802.16 Base Stations (BS) or two 802.11 Access Points (AP) is said to undergo a Horizontal Handoff (HH), while one changing associations between a BS and an AP, or vice versa, is said to undergo a vertical handoff. More details regarding these handoffs will be discussed in the following section. Once VHs are made possible, the full advantages of WHNs can be materialized.

2.2.3 Handover

In this thesis, we deal with the issues of mobility over heterogeneous networks. As mentioned in the previous section, handovers can be either horizontal or vertical. The movement of a MN within the network is called Horizontal Handoff (HH), and it will remain under the same administration, where as, when it moves across different administrations it is called Vertical Handoff (VH). A more critical difference between the Handoffs is that the Horizontal Handoffs (HHs) are most commonly triggered due to variations in the signal level received at the terminal side, and result mostly from the user's mobility. On the other hand, while VHs can certainly be triggered by user mobility, they can also be applied for or requested by users with little or no mobility. For example, the applications run by a certain user in a single place office or home can vary from demanding Voice over IP (VoIP) and video streaming applications to general web browsing and basic file transfers. Another aspect to be kept in mind is the variation in security requirements, for example in the case of users accessing the Virtual Private LAN Services (VPLS) like accessing bank accounts. Based on the requirements of the applications used, different networks may be engaged in service delivery at different times, even simultaneously if a feature such as multi-homing is employed. For operators, admission management need not be restricted to a single network. Coverage permitting, a Network Administrator (NA) receiving a call request can direct the request to the network that best suits the user's requirements and/or that complies with the status of different networks. Furthermore, the NA can balance the admission load between different networks with less concern that a mobile user will be disconnected if moved beyond the vicinity of the accommodating network.

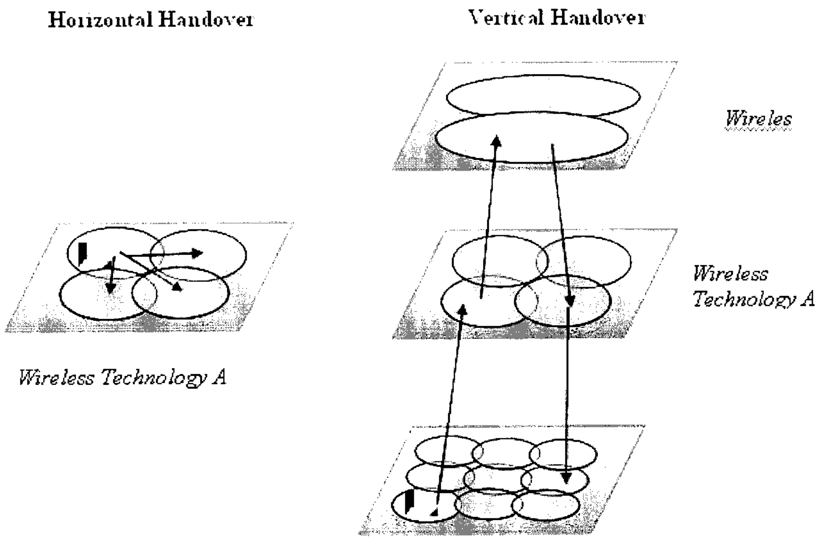

Figure 2-2: Vertical/Horizontal Handoff [8]

2.3 Seamless Mobility

Seamless mobility can be defined as "a handover scheme that maintains the connectivity of all applications on the mobile device when the handover occurs". Seamless mobility is the result that occurs when all sessions of a MN maintain their connections even as the MN changes its point of attachment. If seamless mobility is supported, then a MN can roam across heterogeneous networks and keep its connections active. The underlying heterogeneous network is transparent to the user. Assume a scenario in which a MN runs sessions with TCP and UDP connections. If the MN performs a handover to another network, the MN will receive a new IP-address. However, the corresponding sides of the active TCP/UDP connections do not know the MN's new IP-address; the consequence is that all ongoing TCP/UDP connections will break.

2.4 Mobility Challenges

Mobility management in heterogeneous networks is quite important since a user can move between multiple types of access networks involving several service providers during a multimedia session. The mobility of stations in wireless networks results in issues such as route changes, link failures, and the need for changes of IP address. Handling the seamless migration of ongoing data-flows from one access point to another is one of the key issues in the deployment of so-called next generation networks. Migration is a challenge, mainly due to the limitations of mobile devices, the complexities in upcoming highly integrated wireless networks, and the highly dynamic behavior of nomadic environments. These constraints lead to several challenges.

- Delays due to inter-domain roaming should be eliminated to support real-time services in heterogeneous wireless network environments. The cell handoff delay is the period of time elapsing between the moment the MN detects the subnet change or domain change and the time it receives the first packet of its ongoing communication in the new subnet.
- During inter-domain mobility, disruptions can occur while trying to establish a connection to the new network. Occasional disruptions, delays occurring while in the process changing over to a different network can have an enormous impact on applications performance.
- The integration of various wireless overlays increases the complexity of the handover process. It can aggravate the effects of mobility on the protocol stack and the user experience in a number of ways.
- The addition of new technologies, devices and services to networking environments is increasing the complexity of resource management. Furthermore, the demand for seamless operation and permanent connectivity increases the need for solutions that can handle these complexities and hide them from the user.
- Some real time applications have stringent QoS requirements such as minimum bandwidth, delay, jitter and loss rate.

2.5 Various Approaches

Several schemes have been proposed to support node mobility. These solutions can be classified into micro-mobility and macro-mobility according to the type of roaming they support. When a MN executes a Vertical Handover, its domain usually changes and the appropriate macro-mobility support is required. In the following subsections we provide an insight into popular schemes for micro- and macro-mobility.

Unlike traditional link layer handovers (e.g. those in cellular networks), Vertical Handovers take place in different layers according to the level of integration between the different access technologies. Thus, Vertical Handovers can occur at the network layer, e.g. Mobile IP or even at higher layers e.g. using TCP Migrate or SIP.

2.5.1 Micro-mobility Solutions

The implementation of handover considers the transfer of the ongoing session to the new wireless link. This requires the network to transfer routing information about the new target router in order to establish a new session. Due to differences between access technologies, the transfer of additional contextual information may be required. This contextual information could include QoS or Authentication Authorization and Accounting (AAA) among others. The aim of contextual transference is to minimize the impact of different access technologies and their policies in order to transfer different types of data on applications and services.

Intra-Domain Mobility Management Protocol (IDMP) [14] is a stand-alone approach that provides a multiple Care-of-Address (CoA) intra-domain mobility solution, and is one of many IP-based hierarchical mobility management solutions that aims to minimize handover latency. However, unlike some of the popular micro-mobility solutions as HAWAII and Cellular-IP, IDMP is capable of working completely independent with Mobile IP for enabling global host mobility. IDMP can enhance Mobile IP in micro-mobility environments with frequent handovers and also decreases signaling load. Furthermore, it offers paging services to locate MN within a particular domain and thus save power, which is an important consideration for resource-limited or energy constrained devices.

However, this protocol is not sufficient for managing mobility in 4G communication systems because it cannot offer macro-mobility support by itself, despite the fact it optimizes intra-domain roaming. Heterogeneous roaming demands not only a global solution such as Mobile IP, but also one that is optimized to perform well in inter and intra domain roaming. IDMP was originally deployed using the Linux Mobile IP code produced during of the Stanford University Mosquito Project. A study of vertical roaming was not conducted at that time [12].

Similar approaches have been originally applied to cellular systems to optimize network-level handovers. Micro-mobility protocols such as Cellular IP [16] and HAWAII [14] optimize handover latency and reduce signaling load by distinguishing between the movement of the MN within the domain and outside the domain (i.e. hierarchical handovers). Also, other optimizations are based on the availability of cross-layer information, particularly from the link layer, to anticipate handovers. These are called "fast handovers". [22]

The Cellular IP protocol [16] from Columbia University and Ericsson Research supports paging as well as a number of handover techniques and optimizations. Cellular IP combines the cellular networks' ability to provide smooth, fast handoff and the efficient location management of active and idle mobile users with IP networks' flexibility, robustness, and scalability. To minimize signaling and reduce power consumption, regular sending of packets to update host location information and paging is avoided. However, Cellular IP is also limited in supporting heterogeneous roaming between different domains, and relies on an inter-domain mobility management protocol such as Mobile IP for the support of global mobility.

HAWAII [83], from Lucent Technologies, proposes a separate routing protocol for micro-mobility. HAWAII relies on Mobile IP for inter-domain roaming. One important aspect of HAWAII is that it is not a standalone solution, but rather extends Mobile IP to provide intra-domain mobility with QoS support. When Mobile IP is used for micro-mobility, it results in high control overhead due to the frequent notifications sent to the Home Agent, and in high latency causing disruptions during handover. Also, in the case of a QoS-enabled host, acquiring a new Care-of-Address (CoA) on every handover would trigger the establishment of new QoS reservations along the complete path to the

Correspondent Node (CN). To account for these issues, HAWAII leverages Mobile IP for the support of QoS-aware micro-mobility.

2.5.2 Macro-mobility Solutions

Telecommunication Enhanced Mobile IP (TeleMIP) [16] is a scalable and hierarchical IP-based architecture that provides lower handoff latency and signaling overhead compared to Mobile IP. TeleMIP is also designed to address additional factors such as address space limitations in IPv4 and dynamic load balancing.

To overcome the limitations of global mobility, TeleMIP combines IDMP and Mobile IP for intra-domain and inter-domain mobility support respectively in order to provide a scalable mobility management solution for All-IP networks. Although hierarchical extensions to Mobile IP clearly improve high-frequency updates, the authors of TeleMIP argue that introducing multiple levels of hierarchy in a commercial multi-level provider environment can lead frequently to network management and security issues. Thus, instead of employing a multi-level hierarchy, TeleMIP attempts to achieve a balance between the problems of high update latency and complex management architectures by introducing a structure that uses only a two-level hierarchy.

Despite the fact that TeleMIP improves Mobile IP performance in the intra-domain scenario, it employs MIP for inter-domain roaming. This makes the deployment of TeleMIP unsuitable for heterogeneous environments where nodes roam frequently between different domains.

Seamless Handoff architecture for Mobile IP (S-MIP) [23] varies slightly from all of the above-mentioned techniques in that it explores the intersection of the features of mobile device tracking techniques, handoff algorithms, and hierarchical Mobile IP architecture and combines their key advantages in order to achieve seamless handoff architecture. S-MIP introduces the concept of SPS (Synchronized Packet-based Simulcast) [21] which send packets simultaneously to both the current and new networks, thus minimizing packet loss during handovers. S-MIP [23] builds on the structure of Hierarchical Mobile IPv6 (HMIPv6) with fast handovers and operates in a way similar to the MN Initiated Fast handover scheme. S-MIP could be identified as a technique inherited from HMIPv6.

However, unlike the HMIPv6 [84] S-MIP has fast-handover approach that uses layer-two triggers. In S-MIP, the network uses the MN's location and movement patterns to instruct the MN as to when to handoff. S-MIP uses physical context data to enable context-aware handovers. In [8] the authors perform a comparative study between S-MIP and HMIPv6 and fast handovers for homogeneous micro-and macro-mobility.

Hierarchical Mobile IPv6 (HMIPv6) [19] was developed to reduce the amount of signaling traffic required, which affects the handoff latency of the MN's communications. Unlike MIPv6, HMIPv6 addresses the issues of local mobility and global mobility separately, which means that local handoffs are managed locally without notifying the Home Agent, while global mobility is managed using the MIPv6 protocol. Hierarchical MIPv6 focuses the process required for the support of user mobility locally thus reducing the signaling load on the network. The idea behind HMIPv6 is to divide the global Internet into logical regions defining domains that are independent from subnets. This reduces the signaling load on the network.

TCP Migrate [8], [24] provides a way to achieve session-layer host mobility. Here, TCP is modified at both the mobile and correspondent nodes such that it can withstand changes in IP address during a connection. Using DNS, the correspondent node learns the current address of the MN, with the DNS being updated every time the host moves. However, TCP Migrate lacks support for location privacy and cannot have two MN's communicating simultaneously, making it suitable only for client-server types of applications (e.g. email, web downloads, etc.) but not appropriate for peer-to-peer topologies.

Finally, SIP (Session Initiation Protocol) is an application layer mobility protocol. SIP exploits knowledge about the traffic at a higher layer to benefit real-time flows. This scheme is quite similar to MIPv6 (or MIP with route optimizations). SIP is known today as one of the best existing schemes in that it is advantageous for real-time traffic, both voice and video, as it reduces end-to-end latency by allowing a CN to communicate directly with the MN's CoA, without requiring direct traffic tunneling through the home agent.

2.6 Statement of the Problem

One of the major challenges in deploying a heterogeneous network is achieving seamless handoff among various communications systems with small handoff latency and packet loss. Traditionally, handoff management means that the system maintains communication connection(s) with a MN when it moves from the current serving area to a new serving area served by the same network technology and administrative domain. However, in a heterogeneous environment, handoff management is more complex to deal with, as it has to cover not only Horizontal Handoff but also Vertical Handoff. Horizontal Vertical Handoff deals with the inter-system handoff i.e. when a MN moves from one wireless communication network of a certain technology and administrative domain to a network of different technology and/or administrative domain. For example, a MN moves from a GSM based cellular network to a Wireless LAN network or a Rogers GSM network to a Telus GSM network. It is difficult to realize the vertical handoff among different wireless communication network while meeting the various Quality of Service (QoS) requirements, AAA (Authorization, Authentication and Accounting) mechanism, etc. The physical channel of wireless networks is unpredictable and their frequency spectrum is also limited. For different network technologies their management of resources needs to be able to support the QoS requirements from applications running on a MN. To be able to support applications such as streaming video, wireless network application traffic are usually divided into different service classes. For example, in UMTS traffic is categorized into 4 service classes: conversational, streaming, interactive and background classes. Examples of conversational traffic are video telephony, which has the highest priority and the background classes, e.g. email, and has the lowest. Similarly the IEEE 802.16 standard and WiMAX also has four different service classes: Unsolicited Grant Service, Real-time Polling Service, Non-real-time Polling Service and Best Effort [21]. In general, the IEEE 802.11 offers poor support of QoS capabilities.

If handoff latency (i.e., the time spent during handoff process) is too long, packets may get lost or disconnection may occur. This degrades the QoS in such an environment. Therefore, seamless handover is a big challenge for 4G networks that are supposed to support real-time high-speed multimedia applications. They require small handoff delay and sometimes low packet loss rate and large bandwidth. Another major challenge is that the user has to undergo the AAA process every time the MN moves across different heterogeneous networks. The AAA policy might be different for each network. Authentication provides a way of identifying a user, typically by having the user enter a valid user name and valid password before access is granted. The process of authentication is based on each user having a unique set of criteria for gaining access. The AAA server compares a user's authentication credentials with other user credentials stored in a database. If the credentials match, the user is granted access to the network. If the credentials are at variance, authentication fails and network access is denied.

CHAPTER 3
RELATED TECHNOLOGIES 1

This chapter describes briefly some of the technologies used in the architecture proposed in this dissertation, for supporting seamless handoffs over heterogeneous networks. In Section 3.1 we introduce the Mobile IP technology, while Section 3.2 deals with the Multiprotocol Label Switching (MPLS) technology. In Section 3.3 we discuss Mobile MPLS and, finally Section 3.4 deals with Hierarchical Mobile MPLS.

3.1 Mobile IP

Mobile IP is the most popular approach used at the IP layer to support user mobility. An IP address indicates a user's point of attachment to the Internet in the traditional Internet Protocol. In MIP, regardless of the current location of the MN in the network, every MN is identified by their Home Address (HA). When the MN moves away from its home, it is assigned a Care-of Address (CoA) that provides information about its current point of attachment to the Internet. The MN registers the assigned Care-of Address with the Home Agent. Thereafter, the Home Agent begins redirecting the traffic to the MN's Care-of Address through a tunnel. At the end of the tunnel, each datagram is delivered to the MN.

3.1.1 Mobile IP Terminologies

- Agent Advertisement
- Care-of Address
- Tunnel:
- **Agent Advertisement**: An advertisement message is constructed by attaching a special extension to a router advertisement [2] message.

Care-of Address: The care-of address is the new address assigned to the MN when it moves to a Foreign Network. The protocol can use the following two types of care-of addresses:

- A "Foreign Agent care-of address" is an address given by the Foreign Agent with which the MN is registered [24].
- A "Co-located care-of address" is an externally obtained local address, in which the MN is associated with one of its own network interfaces [27].

Tunnel [2]: A Tunnel is the path established between the Home Agent and the MN. Tunneling has two primary functions: encapsulation of the data packet in order to reach the tunnel end point and de-capsulation when the packet is delivered at that endpoint. The default Tunnel mode is IP Encapsulation within IP Encapsulation. The concept is that, while it is encapsulated, a datagram is routed to a knowledgeable de-capsulating agent, which de-capsulates the datagram and then correctly delivers it to its ultimate destination.

3.1.2 Working of MIP

- The three major processes involved in MIP are the following:
- Agent Discovery
- Registration
- Tunneling

Agent Discovery: Home agents and Foreign Agents may advertise their availability on each network for which they want to offer service. A newly arrived MN can send a message to find out whether any prospective agents are present. Agent Discovery is the method through which a MN first establishes contact with an agent in the local network to which it is attached. Messages containing important information about the agent are sent from the agent to the node; a message can also be sent from the node to the agent asking for this information to be sent.

Registration: When the MN is away from home, it registers its Care-of-Address with its Home Agent. Depending on it's method of attachment, the MN will register either directly with its Home Agent or through a Foreign Agent that would then forward the registration to the Home Agent. Successful registration establishes a mobility binding between a Home Agent and a MN. The Home Agent will encapsulate and forward datagram's addressed to the Home Address over to the Care-of-Address. For the duration of the registration, the MN's' regular home address is tied to its current Care-of-Address.

Tunneling/Packet Forwarding:

- Tunneling is the technique used by Mobile IP, when away from home, to hide a MN's' Home Address from intervening routers between its Home Network and its current location. The tunnel terminates at the MN Care-of-Address. The original datagram is removed from the tunnel at the node corresponding to the Care-of-Address and is delivered to the MN. Thus the location of the MN is hidden from the CN by this mechanism.

- Foreign Agents and Home Agents advertise their presence via Agent Advertisement messages. Based on these Agent Advertisements a MN determines whether its location is on its Home Network or a Foreign Network.

- If the location of the MN is in its Home Network (HN), no special mode of operation is needed. The packets are exchanged (using normal IP routing) between the MN and any other node with which it is communicating with the CN.

- When the location of the MN is in a Foreign Network, the MN is assigned a Care-of-Address in the Foreign Network. The Care-of-Address can be determined either from a Foreign Agent's advertisements (a Foreign Agent Care-of-Address), or by some external assignment mechanism such as DHCP [4].

- The MN then registers its new Care-of-Address with its Home Agent (HA).

- While the MN is away from its Home Network, the HA intercepts the packets sent to the MN's Home Address and establishes a bidirectional Tunnel terminating at the MN's CoA. This Tunnel is used to redirect intercepted packets to any current location of the MN.

- In the reverse direction, datagrams sent by the MN need not pass through the Home Agent. They are delivered to their destination using standard IP routing mechanisms.

3.2 Multi Protocol Label Switching

Multi Protocol Label Switching (MPLS) is an architecture developed for fast packet switching and routing as well as the designation, routing, forwarding and switching of traffic flows through the network. MPLS integrates the performance and traffic management capabilities of Data Link Layer 2 with the scalability and flexibility of Network Layer 3 routing [29].

Routing is based on the exchange of network reachability information in conventional IP. Each router in a network maintains its own IP routing information. As a packet traverses the network, each router extracts all the information relevant to forwarding the packet from the packet header. This information is used as an index for a routing table lookup to determine the packet's next hop. This procedure is repeated at each router across the network. At each hop in the network, the optimal forwarding of a packet will be determined. MPLS's ability to support constraint-based routing and traffic engineering delivers the QoS that is required to support conversational, streaming, and interactive traffic, something previously possible only with ATM. Hence, Multi Protocol Label Switching (MPLS) is emerging as the technology of choice for facilitating traffic engineering and internetworking.

3.2.1 Issues with Existing Packet Forwarding

Conventional IP packet forwarding has several limitations [27].

- It has limited capability for dealing with addressing information beyond just the destination IP address carried in the packet.
- Because all traffic to the same IP destination-prefix is usually treated similarly, various difficulties arise. For example, it becomes difficult to perform traffic engineering on IP networks.
- Also, IP packet forwarding does not easily take into account extra addressing-related information such as Virtual Private Network (VPN) [28] membership.

3.2.2 MPLS Network Structure

The key networking elements, devices and protocols that participate in the MPLS protocol mechanisms can be classified into:

- Label Edge Routers (LERs)
- Label Switching Routers (LSRs)
- Label Switched Path (LSP)
- Forwarding Equivalence Class (FEC)
- Label Distribution Protocol (LDP)

- MPLS Label

Label Edge Router (LER): A LER is a device that operates at the edge of the Access Network and MPLS network. Usually LERs have the capability to support multiple ports connected to dissimilar networks (such as Frame Relay, ATM, and Ethernet) and forward this traffic on to the MPLS network after establishing Label Switched Paths (LSPs). One major function of LERs is assigning and removing labels as traffic enters or exits an MPLS network.

Label Switch Router (LSR): A LSR is a high-speed router device in the core of an MPLS network that participates in the establishment of LSPs [27] by using the chosen label signaling protocol and implements the high-speed switching of the packets over the established paths.

Label Switched Path (LSP): In MPLS networking, a LSP is an end-to-end defined route connecting two LERs (one acting as the traffic's incoming point the other one then outgoing point). An LSP is established using a certain signaling protocol such as Label Distribution Protocol [29], RSVP-TE [32], or CR-LDP [33]. The path is established according to the criteria mapped to a certain forwarding equivalence class (FEC).

A "Label Switched Path (LSP) of level m" for a particular packet P is a sequence of routers [13].

<R1, ..., Rn>

R1, the "LSP Ingress", is a LSR that pushes a label onto the packet label stack, resulting in a label stack of depth m.

Forwarding Equivalence Class (FEC): MPLS ensures scalability by supporting traffic aggregation through the use of FEC. The ingress LER is the place where aggregation is implemented. LER is responsible for classifying incoming packets and relating them to the FEC's. Within a router, each FEC is associated with an appropriate label and forwarding path. LER uses several modes to classify traffic, such as using the packet destination address and port, priorities, etc. All of the packets of a certain FEC are treated in the same manner (e.g. they go over the same path, with the same forwarding treatment at each router associated with the path).

Label Distribution Protocol (LDP): A label distribution protocol is a set of procedures by which one LSR informs another of the label/FEC bindings it has made. Two LSRs that use a label distribution protocol to exchange label/FEC-binding information are known as "label distribution peers" with respect to the binding information they exchange. If two LSRs are label distribution peers, we say there is a "label distribution adjacency" between them. The label distribution protocol also encompasses any negotiations in which two-label distribution peers need to engage in order to learn of each other's MPLS capabilities.

MPLS Label

A label is a short, fixed length, locally significant identifier that is used to identify a FEC. The label, which is put on a particular packet, represents the Forwarding Equivalence Class to which that packet is assigned.

MPLS label structure

Table 1: MPLS Label Structure [27]

20	23	24	32Bit
Label	Exp	S	TTL

The Label Value carries the actual value of the Label. The Label stack consists of the following information about the packet:

- The next hop to which the packet is to be forwarded.
- The operation to be performed on the label stack before forwarding. This operation may be to pop an entry off the label stack, or to replace the top label stack entry and then push one or more additional entries onto the label stack. The label summarizes essential information about routing the packet i.e. Destination, Precedence.
- Virtual Private Network membership.
- Quality of Service (QoS) information from RSVP.

- The route for the packet, as chosen by traffic engineering (TE).
- Exp - Experimental Use: reserved for experimental use.
- S - Bottom of Stack: This bit is set to one for the last entry in the label stack, and zero for all other label stack entries.

Time-to-Live (TTL): Each packet carries a "Time to Live" (TTL) value in its header in conventional IP forwarding. When a packet passes through a router, its TTL gets decremented by 1. If the TTL value reaches 0 before the packet has reached its destination, the packet gets discarded [28].

3.2.3 Working of MPLS

The main concept of MPLS is to include a *label* in each packet. Packets or cells are assigned short, fixed-length labels. Switching entities perform table lookups based on these labels in order to determine where data should be forwarded.

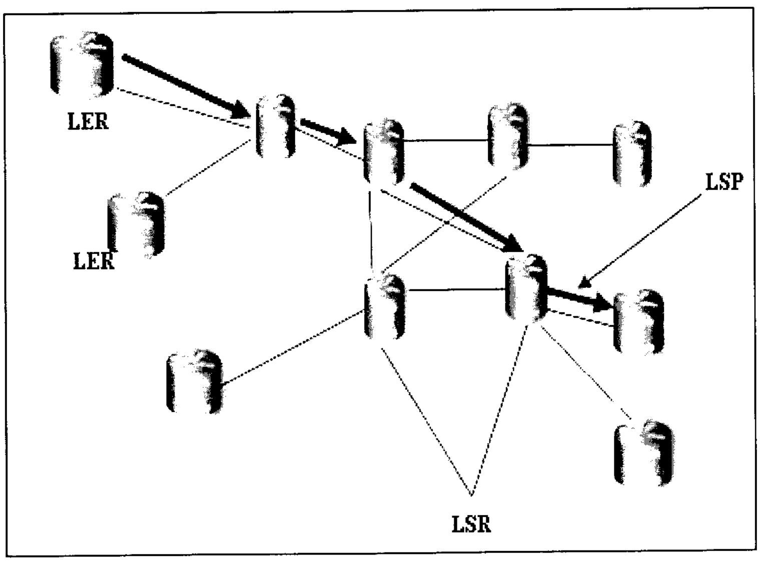

Figure 3-1: Typical MPLS Network

- The complete look up of the label is performed only once at the edge LSR, which is located at each edge router of the network.
- When an IP packet enters a LSP, the ingress router examines the packet and assigns a label to it based on its destination. Then the label is placed in the packet's header.

- The label enables the packet to be forwarded based on the information associated with the label. This is in contrast to the packet being forwarded based on its IP routing information. The packet is then forwarded to the next router in the LSP. At each router across the network, the label of the incoming cell or packet is examined in order to send it on its way across the network.

- Each router along the LSP uses the label to look up information in its label-forwarding table. The old label is then replaced with a new label. The packet is then forwarded to the next router in the path.

- An edge LSR at the other end of the path swaps the label out for the appropriate header data linked to that label.

- The egress router removes the label when the packet reaches it and the packet again becomes a native IP packet. It is again forwarded based on its IP routing information.

The main advantage of this technique is its simplicity and scalability. The forwarding decisions based on some or all of these different sources of information can be achieved by means of a single table lookup using a fixed-length label. It is the significance of the label concept that is very important property since it is the main reason for making MPLS a very scalable technology.

3.3 Mobile MPLS

As discussed in the previous section, the major procedures involved in the Mobile IP technology are the Agent Advertisement process, Registration process, and data packets forwarding process. In the Data Forwarding process of Mobile IP, the HA checks every IP packet that it receives to see whether the destination IP address of the packet matches any MN that are currently registered in a Foreign Network. Then the HA will perform IP Tunneling on the packet by adding an IP header to the packet and then sending it out to the routing process for forwarding. If the matching is not found, the HA will send the packets out to the routing process for further forwarding.

Scalability becomes a major issue when the number of MN registered with Foreign Networks increases. As the number of MN's increases, the forwarding process also becomes longer. If we assume that every packet forwarded by the HA has to undergo this forwarding process, the overhead of this process may become too high. This is a matter of serious concern and is likely to affect the use of the Mobile IP protocol in future wireless mobile systems.

In order to handle this concern with Mobile IP, a new approach integrating MIP and MPLS was introduced and was studied quite extensively [27], [28]. The integration improves the scalability of the Mobile IP data forwarding process by leveraging on such features as fast switching, small state maintenance, and the high scalability of MPLS.

3.3.1 Working of Mobile MLPS

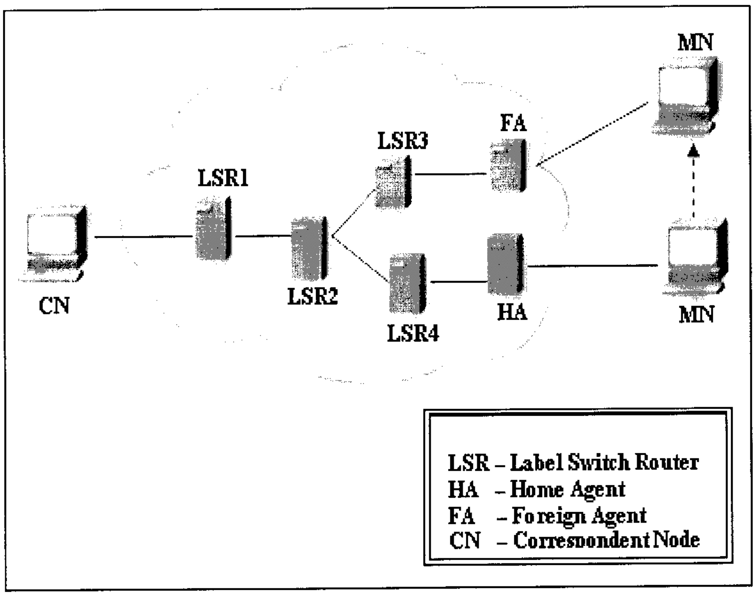

Figure 3-2: Working of Mobile MPLS [28]

As shown in Figure.3-2, the HA and FA are edge LSR's and belong to the same MPLS domain. They support both MPLS and Mobile IP functionality. We assume that the MN's home address is *x.x.x.x* and the HA's address is *y.y.y.y*. In addition, we assume that the FA's CoA is *z.z.z.z*.

Registration Procedure

- The MN checks for its current location when it receives Agent Advertisement messages broadcast by the Foreign Agent (FA).
- If the MN identifies itself as being located in a Foreign Network, it acquires a temporary CoA from the FA and sends a registration request to the FA.

- The FA, which acts as an edge LSR, will analyze the incoming registration request message and assign it a destination address.

- The FA updates its routing table and adds a host specific row with the value of MN's home address. In addition, it sets the outgoing port value of this entry to be the incoming port number from which it received the registration request.

- Based on the IP routing table, the FA forwards the registration request message to the HA.

- When the HA receives the registration request message and checks the CoA, it searches its label table to find the row with the MN Home Address as FEC. Refer to the second row in Table-2.

- The label request is then sent to the FA using LDP, and using the FA's CoA as FEC. The FA replies with a LDP label-mapping message to the HA. When this label-mapping message arrives at the HA, the LSP will have been established (the first row in Table-1).

- The HA changes the row in its label table that corresponds to the MN home address as FEC. It sets the empty out label and outgoing port entries with the values of the out label and outgoing port respectively of the LSP from HA to FA. This way, the HA can relay the packets destined to the MN's home address to its current location in the Foreign Network.

- Later, HA sends a registration reply to the FA along the LSP from HA to FA.

- When the FA receives the registration reply, it records the incoming port number and label value of the reply message. Then it adds a new row in its label table.

3.3.2 Advantages of Mobile MPLS Concept

- Following are some of the main advantages of the Mobile MPLS concept.

- As noted above, integrating MPLS and Mobile IP makes IP-in-IP Tunneling obsolete in the data forwarding process. Instead, we make use of the advantage of MPLS to switch the packet to the Foreign Network. Switching is much faster than conventional IP forwarding. The whole forwarding process is done at the MPLS layer and the HA does not need to involve the IP layer. This improves the scalability of the Mobile IP protocol.

- Since the label header is much smaller than the IP header, the traffic overhead from HA to FA is reduced.

- Another significant advantage is that we could use QoS aware label distribution protocol such as the protocols like the Constraint-Based Label

Distribution Protocol (CR-LDP) [28] to setup an LSP that satisfies the QoS requirements of the traffic as well as to perform traffic engineering

- IP packets can be assigned to different label switched paths based on their destination IP address and other criteria, such as QoS attributes.

3.4 Hierarchical Mobile MPLS

Recent wireless networks have successfully implemented small-size cells (Micro-cells) in order to achieve higher system capacity. Micro-cell implementation is able to increase the handoff rate of the mobile hosts substantially. In such a scenario, normal mobile MPLS is no longer the best solution. To overcome the limitations of the Mobile IP protocol, the author of [37] proposed a new protocol called Hierarchical Mobile MPLS.

In conventional Mobile MPLS, when a mobile host moves to a new MPLS network a registration message has to be sent to the HA though the FA. A new LSP will then be established between the HA and FA. This approach has a number of serious drawbacks for being deployed in a wireless network that is supporting uses of moderate to high mobility, such as the following:

- Every time the MN changes the point of contact to a new FA, a registration message has to be sent all the way to the HA. This is quite a long process.
- In the meantime, if a CN is in communication with the MN and is sending packets to MN at the time MN is in the handoff process, there will be a disruption in the traffic flow.

In order to reduce the handoff latency, a new mobile MPLS concept called Hierarchical Mobile MPLS was introduced to reduce the handoff latency.

In that work, the author introduced a new component called the Foreign Domain Agent (FDA), which is affiliated with each MPLS domain. Figure.3-3 shows the architecture for hierarchical mobile MPLS. As in the conventional Mobile MPLS, there is a FA in each sub-network.

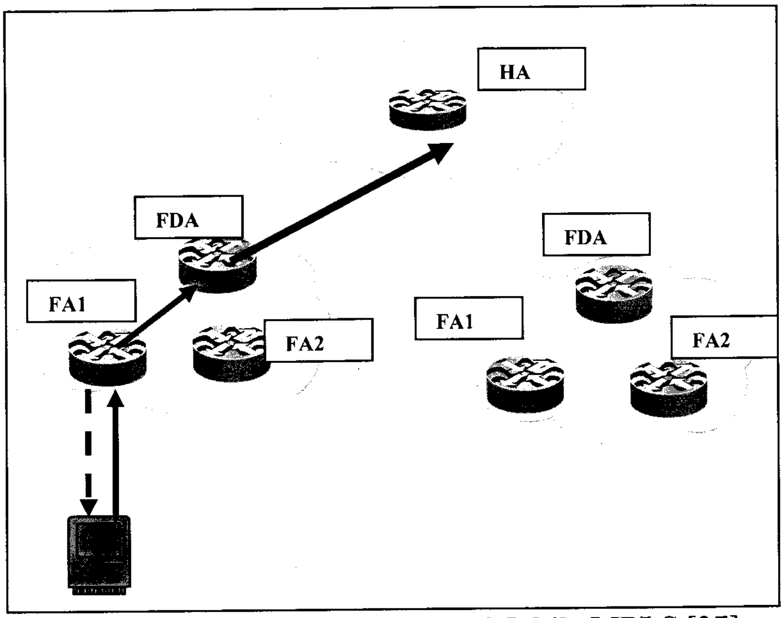

Figure 3-3: Hierarchical Mobile MPLS [37]

3.4.1 Registration Procedure

- In a hierarchical mobile MPLS architecture, the MN determines whether it is at home or in a Foreign Domain by receiving Agent Advertisement messages broadcast by the FA. If the MN determines it is in a Foreign Domain, then it acquires a temporary CoA from the FA and sends a registration request to the FA.

- The FA forwards this Registration Message to the FDA instead of sending it directly to the HA of the MN. The Registration Message is forwarded to the HA by the FDA where it includes its own IP address as IP address. When the HA receives the Registration Message and becomes aware of the IP address of the FDA, it uses the LDP protocol to send a label request to the FDA with the IP address of FDA as FEC.

- The FDA replies back to HA with a LDP label mapping message and sends a label request to the FA of the sub-network in which the MN is currently located.

- When the label mapping message arrives at the HA, the LSP from the HA to the FDA is established. Similarly, the FA replies with a LDP label-mapping message back to the FDA, and when this message arrives at the FDA the LSP from the FDA to the FA is established.

- Then, the HA will search the label table and find the row with the MH's Home Address as the FEC and change the outgoing port and out label to the same values as the ones for the LSP from the HA to the FDA.

- Finally, the HA sends a registration reply to the FDA along the LSP from the HA to FDA, and the FDA forwards this registration reply to the FA along the LSP from the FDA to the FA.

- When the FA receives the registration reply, it adds a new row in its label table and enters the incoming registration reply's label value and port into the table.

3.4.2 Datagram Delivery

- When a CN sends some packets to a MN that is currently located in a Foreign Network, the HA will intercept these packets. In hierarchical mobile MPLS, the HA will use the incoming label value as an index to look up its label table and find the outgoing label and outgoing port for this packet.

- The packets are then delivered from the HA to the FDA along the LSP via label swapping. The FDA receives the packets and forwards them through the LSP from the FDA to the FA.

- Once the FA receives the packets, it will look up its label table. Since it is the egress router for the LSP from the FDA to FA and the out label and outgoing port fields are empty, the FA removes the label and sends the packets to the MN through the IP layer. Finally, the MN receives the packets sent by the CN.

3.4.3 Mobility Support by Hierarchical Mobile MPLS

Hierarchical Mobile IP (HMIP) is an ideal solution for micro-mobility management. It reduces the amount of signaling to Correspondent Nodes and the Home Agent and improves the handoff speed performance of Mobile IP. The paragraphs below describe briefly the handoff procedure using hierarchal Mobile MPLS.When the MN hands off from one sub-network to another sub-network within the same domain (Figure 3-3), it will send a registration request to the new FA. Then the new FA (FA2 in Figure 3-3) will forward the Registration Message to the FDA, which will send a label request message back to the new FA. The new FA receives the label request and responds to the FDA with a label-mapping message. A new LSP will then be established from the FDA to the FA2. The greatest advantage of this approach is that the LSP from the HA to the FDA will remain unchanged. Thus, the advantage of hierarchical mobile MPLS over the conventional mobile MPLS is that the former does not need a totally new LSP to be set up from the HA to the new FA. Instead, there is a need only to set a new LSP only from the FDA to the new FA.

CHAPTER 4
RELATED TECHNOLOGIES 11

4.1 Introduction

The Session Initiation Protocol (SIP) is an application-layer solution for handling mobility in Communications Technology. This protocol is used for establishing, modifying, and tearing down multimedia communication sessions over telecommunication networks. SIP is a control-signaling protocol that has been standardized by the Internet Engineering Task Force (IETF) for initiating, modifying, and terminating interactive user sessions that involve multimedia elements such as video, voice, instant messaging, online games, and virtual reality. Since SIP is designed to be independent of the underlying transport layer of the Open Systems Interconnection (OSI) communications model, it can run on TCP, UDP, or SCTP environments.

SIP was also accepted as a 3rd Generation Partnership Project (3GPP) signaling protocol and permanent element of the IP-Multimedia Subsystem (IMS) architecture. It is one of the leading signaling protocols for Voice over IP, along with H.323 in communications systems. Borrowing from ubiquitous Internet protocols, such as Hypertext Transfer Protocol (HTTP) and Simple Mail Transfer Protocol (SMTP), SIP is a text-encoded and highly extensible technology. SIP could be extended to accommodate features and services such as call control services, mobility, interoperability with existing telephony systems, and much more. The protocol is published as IETF RFC 2543 and currently has the status of a proposed standard. SIP is able to sit comfortably alongside Internet applications, enabling telephony to become another web application and becoming easily integrated into other Internet services. The protocol is a simple toolkit that service providers can use to build converged voice and multimedia services.

SIP is not the only protocol that communication devices require for establishing and carrying out the communication task. The two protocols most often used along with SIP are Real Time Protocol (RTP) and Session Description Protocol (SDP). While RTP is used to carry real-time multimedia data, SDP is used to describe and encode the capabilities of the session participants.

SIP is a powerful yet simple protocol. Unquestionably, SIP is an important technology that is becoming widely deployed. It is a catalytic protocol that delivers key signaling elements, which can turn a Voice over IP network into a true IP communications network; one that is capable of delivering next generation converged services.

4.2 SIP Architecture

The SIP technology is examined in detail in this section.

4.2.1 SIP Components

- The following are five components that constitute the SIP architecture.
- User Agent Client (UAC)
- User Agent Server (UAS)
- Proxy Server
- Redirect Server
- Registrar Server

4.2.1.1 User agent client (UAC)

UAC is one of the two client–side components. The other component is the User Agent Server (UAS), which is described in subsection 4.2.1.2. The UAC is a client application that can initiate a SIP request. When the SIP session is initiated by the UAC SIP component with a request, it determines the information essential for the request, specifically the protocol, the communication port, and the IP address of the UAS to which the request is being sent. This information can be dynamic in nature.

4.2.1.2 User agent server (UAS)

UAS is a server application that contacts the user when a SIP request is received and then returns a response on behalf of the user. The UAS is located at the receiver end, and hosts the application responsible for receiving SIP requests from a UAC. Upon receipt of an

INVITE message, it sends a response to the call initiation request (INVITE) back to the UAC. The UAS may issue multiple responses to the UAC (not necessarily a single response). The communication between the UAC and UAS is in client/server mode and is also peer-to–peer.

4.2.1.3 Proxy server

The SIP Proxy Server is a device that receives session initiation requests from the client and then forwards them on the client's behalf. Proxy Servers receive SIP messages and forward them to the next SIP server in the network. SIP proxies are very helpful in offloading tasks and simplifying the implementation of end station devices. They provide functions such as authentication, authorization, network access control, routing, reliable request retransmission, and security. It may be noted that the SIP proxy participates in SIP messaging only when the call is set up or terminated. The devices send their traffic directly to each other without involving the proxy during the session. A Proxy Server mimics the behavior of both the SIP server and client. A Proxy is involved only in the set-up and teardown of the communication session. Once a session is established, communications occur directly between the caller and the callee. The Proxy Server also serves as a mediator that services requests or forwards them to other UASs or UACs for processing.

4.2.1.4 Redirect server

The Redirect server is an optional SIP component that does not route SIP messages. The Redirect Server provides a client with information about the next hop or hops that a message should take. It also returns a redirect (change in routing) to the UA [User Agent] for direct routing. It must be remembered that a server enables SIP for end-to-end signaling without intervention. Redirection allows servers to push routing information for a message request back to the clients as a response, thereby taking themselves out of the loop of further messaging in the ongoing session while still aiding in locating the target of the request. When the originator of the request receives the redirection, it will send a new request based on the Uniform Resource Identifiers (URIs) it has received. By propagating URIs from the core of the network to its edges, redirection allows for considerable network scalability. The client is able to contact the next hop server or UAS directly. Also, the

Redirect Server enables users to temporarily change geographic location and still be able to be reached using the same SIP identity. A Redirect Server accepts a SIP request, sets the SIP address of the called party to zero (if there is no known address) or with the address known to it, and then responds back to the client. Unlike Proxy Servers, Redirect Servers do not pass requests to other servers.

4.2.1.5 Registrar server

The Registrar Server processes requests from UACs for the registration of their current location. They are often co-located with a Redirect or Proxy Server. A Registrar Server updates the location of the MN. It then accepts REGISTER requests from the MN for the purpose of updating a location database with the contact information of the user (user's new address) specified in the request. The Registrar Server makes it possible for users to alter the address at which the MN can be contacted.

4.2.1.6 SIP B2BUA server

SIP Back to Back User Agents (B2BUA) Servers are the most powerful types of SIP servers. B2BUA servers are used to provide value added features for point-to-point calls and manage multi-point calls. IETF standard (RFC 3261) defines a back-to-back user agent as "a logical entity that receives a request and processes it as a user agent server (UAS). In order to determine how the request should be answered, it acts as a user agent client (UAC) and generates requests. Unlike a proxy server, it maintains a dialogue state and must participate in all requests sent on the dialogues it has established."

4.3 SIP Messages

This section provides a brief overview of SIP messages and methods. All SIP messages are either requests from a server or client or responses to a request. The two major classifications of SIP messages are as follows:

- Requests: Messages of a request type that are sent by the client, usually for the purpose of initiating a session.
- Responses: Messages sent by the server in response to the request message type sent by the client.

4.3.1 Request Methods

SIP Requests are messages sent by the session initiator, also called the SIP User Agent Client. There are six basic request types. Table-3 below shows the different types of SIP requests and the purpose of each method.

Table 2: SIP Request Methods

METHOD	DESCRIPTION
INVITE	Initiates a call, changes call parameters (Re-INVITE).
ACK	Confirms a final response for INVITE.
BYE	Terminates a call.
CANCEL	Cancel Search and start "RINGING".
OPTIONS	Queries the capabilities of the called side.
REGISTER	Registers with the location services.
INFO	Sends mid session information that does not modify the session state.

4.3.2 Responses

A SIP Response is a message generated by a UAS or a SIP Server to reply to a request generated by a UAC. Response messages contain numeric response codes. The SIP response code is partly based on the Hypertext Transfer Protocol's (HTTP) response codes. There are two broad classifications of responses and each message type has a total of six methods.

Response messages can be classified into two major categories:

- **Provisional Response**: The server uses provisional responses to indicate progress, but these messages are not capable of terminating SIP transactions.
- **Final Response**: (2xx, 3xx, 4xx, 5xx, 6xx classes)—final responses.

Table-4 shows a mapping between SIP Response Methods and the description of each method.

Table 3: SIP Response Methods

SIP RESPONSE METHOD	DESCRIPTION
SIP 1xx	Provisional, searching, ringing, queuing etc.
SIP 2xx	Success.
SIP 3xx	Redirection, forwarding.
SIP 4xx	Request failure.
SIP 5xx	Server Failure.
SIP 6xx	Global Failure (busy, refusal, not available anywhere).

4.3.3 SIP Message Structure

In the previous section we discussed in detail the different types of SIP messages. In this section we will investigate the structure of each message.

Communication using SIP (often called SIP signaling) is comprised of a series of ***messages***. The network has the capability to transport messages independently; usually each message is transported in a separate User Datagram Protocol (UDP) datagram. Each message consists of a "Start Line", "Header", and "Message Body". The first line identifies the type of message, such as ***requests*** and ***responses***. Requests are generally used to initiate some action or inform the message's recipient that something is required. Responses are used to confirm that a request has been received and processed. They also contain the status of the processing.

SIP messages consist of 3 major parts [6]

Start Line: Every SIP message begins with a "*Start Line*". The Start Line specifies, the message type (the request message code or response code. A response also known as status code.) And the protocol version (SIP protocol version being used, e.g. the current SIP version is 2.0).

In short, a Start Line represents a Request-line (requests) or a Status-line (responses) as follows:

- The Request-line includes a Request URI (Uniform Resource Identifier) that indicates to the user to whom this request is being addressed to. Unlike the "To" field, this address can be re-written by proxies.
- The Status-line holds the numeric Status-code and its associated textual phrase. Status is a three-digit number that indicates the outcome or the response of the request. Refer to Figure-4-2 to see examples of a three-digit response code.

Header: SIP header fields are used to convey message attributes and to modify message meaning. They resemble HTTP header fields. The Header format is:

<name>:<value1>,<value2>.....<valueN>

Headers can span multiple lines. Some SIP headers contain information such as Via, Contact, Route, and Request-Route.

Message Body: A Message Body is used to indicate the features of the session to be initiated. For example, in a multimedia session it may include audio and video codec types, sampling rates, etc. Message bodies can appear in both request and response messages. SIP makes a clear distinction between the signaling information conveyed in the SIP Start Line and Headers and the session description information, the latter of which is outside the scope of SIP. The session setup agreement depends on the compatibility of the features mentioned in the message body by both, the caller and callee. Message body types include the following Session Description Protocol (SDP) [3], Multipurpose Internet Mail Extensions (MIME) [23], as well as others are in the process of been by IETF and protocols associated with specific implementations.

4.4 SIP Entity Interactions

4.4.1 Session Establishment and Termination

Figure 4-1 shows a typical example of a SIP message exchange between the two users used in our example, Andrew and Amy.

SIP messages are usually arranged into *transactions* between user agents (UAC & UAC). Thus, SIP is also called a *transactional protocol*. A transaction is a sequence of SIP messages exchanged between SIP network elements. A transaction consists of a request and all responses to that request. Each transaction consists of a request that invokes a particular method on the server and at least one response to the request. In this particular example, the transaction begins with Andrew sending an INVITE request addressed to Amy's SIP Uniform Resource Identifier (URI).

Figure 4-3 represents the basic SIP transactions for SIP Session Establishment and termination. Table-7 shows the structure of each SIP message.

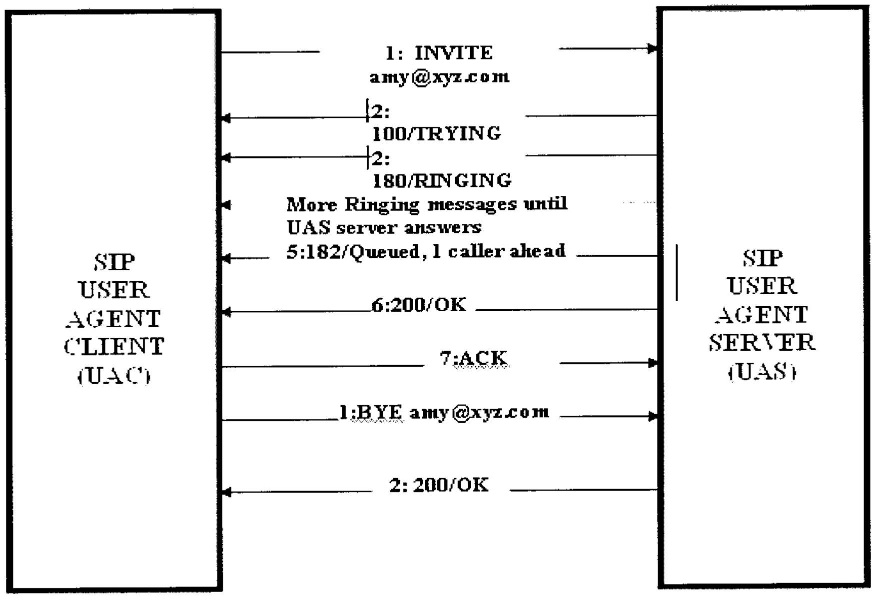

Figure 4-1: SIP Session Establishment and Termination

The following table shows the structure of each message exchanged between the UAC and UAS for a session's initiation and termination, as shown in Figure 4-1.

Session Initiation Protocol Call Flow

- A call is initiated when the UAC sends an INVITE message to Amy's SIP address: amy@xyz.com. This message also contains a SDP packet describing the media capabilities of the calling terminal.
- The UAS receives the request and immediately responds with a 100-Trying response message.
- The UAS starts "ringing" to inform Amy of the new call. At the same time, a 180 (RINGING) message is sent to the UAC.

- Amy picks up the call. The UAS checks the features of the session and sends a 200 (OK) message to the calling UA. This message also contains a SDP packet describing the media capabilities of Bob's terminal.

- The calling UAC sends an ACK request to confirm that the 200 (OK) response message has been received.

The Session Termination Call Flow

- When the caller decides to end the call, a BYE request is sent to Amy's UAS at SIP address sip:amy@xyz.com

- Amy's UAS responds with a 200 (OK) message and notifies Amy that the conversation has ended.

4.4.2 SIP Functionality when Working with a Redirect Server

If a Redirect Server is used in the SIP signaling path, and the caller UA sends an INVITE request, the Redirect Server contacts the Location Server to determine the path to the callee and then sends that information back to the caller. The caller then acknowledges the receipt of this information. The figure below (Figure 4-2) shows the working of SIP with the Redirect Server.

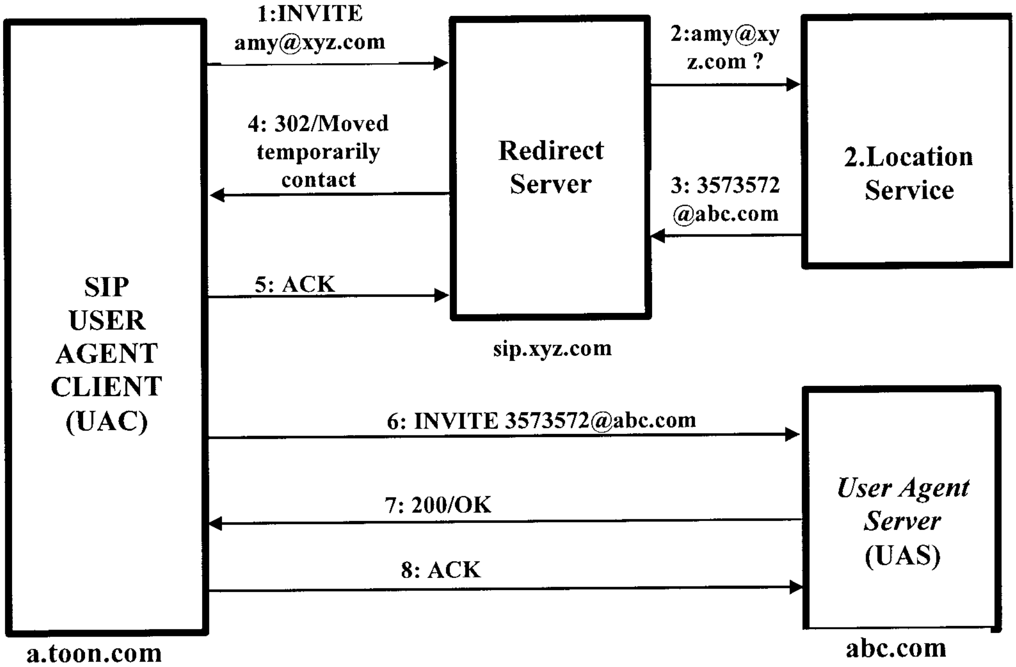

Figure 4-2: SIP Messaging Using a Redirect Server

Call Flow

- First a SIP INVITE message is sent to amy@xyz.com. If a SIP entity is placed in the SIP signaling path it could act as a SIP Redirect Server sip.xyz.com along the signaling path.

- The Redirect Server looks up Amy's current location in a Location Server using protocols such as the Lightweight Directory Access Protocol (LDAP) [2]

- The addresses registered to a Registrar are stored in a Location Server. The working of the Registrar Server is mentioned in Section 4.2.1.5. The Location Service returns Amy's current location: SIP address 3573572@abc.com

- The Redirect Server returns this information to the calling UAC using a 302 (Moved Temporarily) response. In the response message that is being sent back to UAC, it enters the field contact header and sets the value to Amy's current location, 3573572@abc.com.

- The calling UAC acknowledges the response by sending an ACK message.
- The calling UAC then continues the transaction directly with abc.com by sending a new INVITE.
- abc.com is able to notify Amy's terminal of the call. A 200 (OK) response is sent back to the calling UAC.
- The calling UAC acknowledges with an ACK message.

4.4.3 SIP Working with Proxy Server

Servers are generally part of any network. They possess a predefined set of rules for handling the requests sent by clients. As mentioned earlier, SIP makes use of elements called "*Proxy Servers*" to help route requests to the user's current location. The Proxy Server is an intermediate device that receives SIP requests from a client and forwards them on the client's behalf. Basically, Proxy Servers receive SIP messages and forward them to the next SIP server in the network. The functionality of the Proxy Server is not limited to just forwarding the messages to their destination, but also provides functions such as authentication, authorization, network access control, routing, reliable request retransmission, and security.

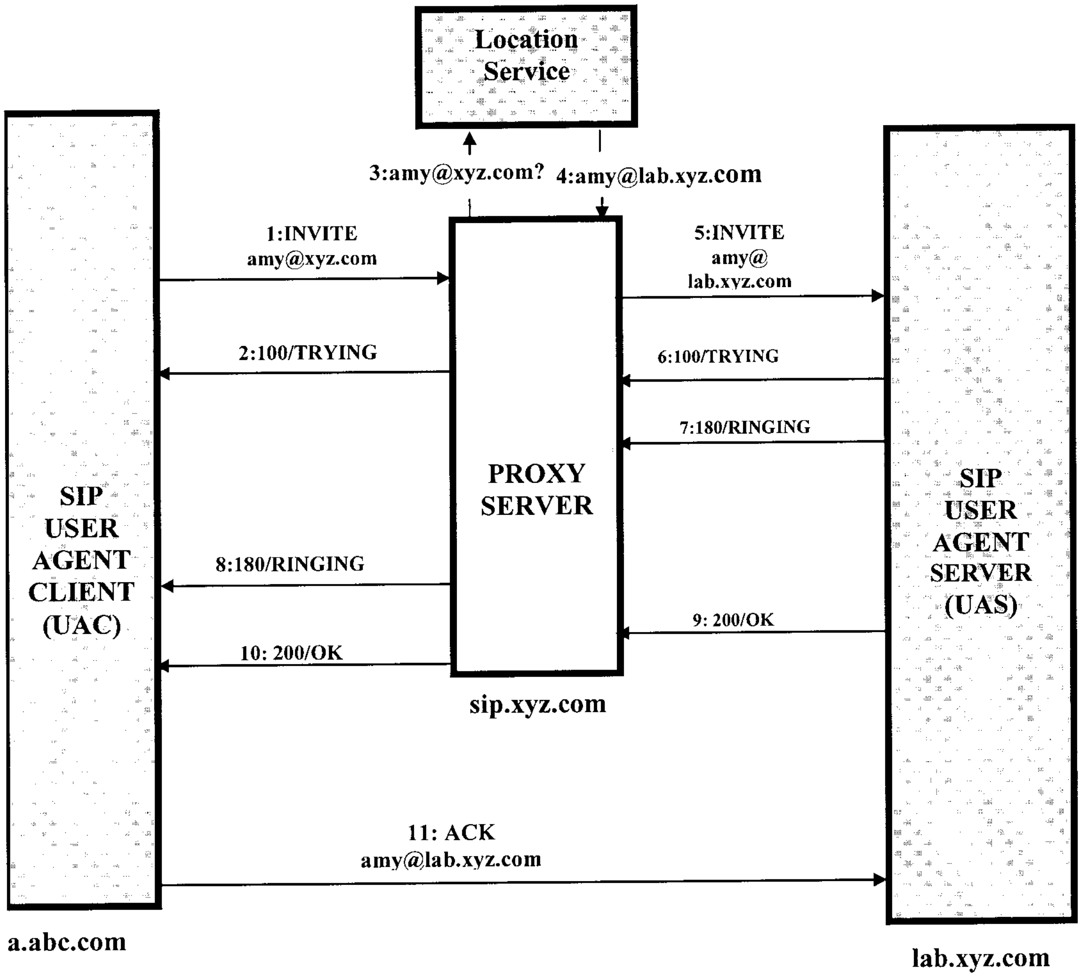

Figure 4-3: SIP Messaging Using a Proxy Server

Call Flow

- An INVITE message is sent to amy@xyz.com, but finds a SIP entity that responds like a Proxy Server sip.xyz.com along the signaling path.
- The Proxy Server immediately responds with a 100 Trying response.
- The Proxy Server looks-up Amy's current location in a Location Service using a non-SIP protocol (e.g. LDAP).
- The Location Service returns Amy's current location: SIP address amy@lab.abc.com.
- The Proxy Server decides to proxy the call and creates a new INVITE message based on the original INVITE message, but with the request URI modified to

lab.amy@abc.com. The Proxy Server sends this request to the UAS at lab.abc.com. This approach is different from that of the Redirect Server, which does not enter the communication path but only provides the location information to the caller for initiating the call.

- The UAS responds first with a 100 (Trying).
- The UAS responds with a 180 (Ringing) response.
- The Proxy server forwards the 180 Ringing response back to the calling UA.
- When the user has accepted the call, the UAS at abc.com sends a 200 (OK) response. In this example, Amy's UAS inserts a Contact header into the response with the value amy@lab.abc.com. Further SIP communication will be sent directly to the UA and not via the Proxy Server.
- The Proxy forwards the 200 (OK) response back to the calling UAC.
- The calling UA sends an ACK directly to Amy's UA (according to the Contact header it found in the 200 (OK) response).

4.5 Why SIP for IP Mobility & How?

In recent times, we have seen a rapid growth in cellular mobile telecommunications and Internet penetration. One technology that has emerged over the past few years is the Voice over IP (VoIP) service, which has shown rapid growth. The natural evolution of these technologies is toward a wireless Internet, which will provide access not only in real-time, but also to non-real time services from anywhere at any time.

Currently, there are two basic approaches to supporting mobility in VoIP services. The first approach is to solve the mobility in the network layer by using Mobile IP (MIP). Although Mobile IP is not directly related to VoIP applications, Mobile IP can address mobility support for VoIP service. The other approach is to solve the mobility problem in the application layer, e.g. SIP.

However, the MIP solution suffers from some serious drawbacks such as Triangular routing, tunneling, etc. In Triangular routing, a packet being sent to a MN travels via home agent, whereas a packet from the mobile host is routed directly to the destination. The route optimization [12] solves this issue by sending binding updates to inform the sending host of the actual location of the mobile host. This approach has several problems [11] including tunneling, which add overhead to the packet thus consuming more from the already bandwidth constrained wireless links. This overhead introduced due to the increase in the path the packets has to traverse through before reaching the final destination adds an extra delay to the traffic toward the mobile host, which is not acceptable for delay sensitive traffic because it might result in excessive delays.

For real-time traffic such as voice or video over IP, it is more common to use the Real-Time Transport Protocol (RTP) [13] over UDP (User Datagram Protocol). The important issue to be handled is fast handoff, since real-time traffic cannot accept big handoff delays. These issues inspired the introduction of mobility in the higher layer. Another issue with MIP is that the MN needs a permanent home IP address, which can be a problem due to the address exhaustion in IPv4.

With these existing issues in the network layer, the expected growth of Internet multimedia services, particularly the expected migration to voice telephony, have motivated a large amount of research on SIP based mobility.

Some of the advantages of handling SIP based mobility are the following:

- Mobility in the higher layer means that knowledge about the traffic is more transparent and thus helps to make decisions about how to handle mobility in various situations.
- It provides a means of route optimization and improved performance for real-time services via SIP signaling messages for address binding, registration, etc.
- It allows users to depend on their appliances rather than on the network for supporting mobility on an end-to-end basis without relying on knowledge about the abilities of network elements for packet interception and forwarding, i.e. mobile users can roam across SIP environments without any concern about whether the nodes of the network they are operating are within support network layer mobility or not.

In the following section we provide a brief introduction to various types of mobility that are supported by SIP.

4.5.1 Pre-Call Mobility

The easiest part of SIP mobility is pre-call mobility, in which the MN acquires a new address prior to receiving or making a call. The MN simply re-registers with its Registrar Server in its Home Network each time it obtains a new IP address. The only difficult part is to detect, at the application layer, when the IP address has changed.

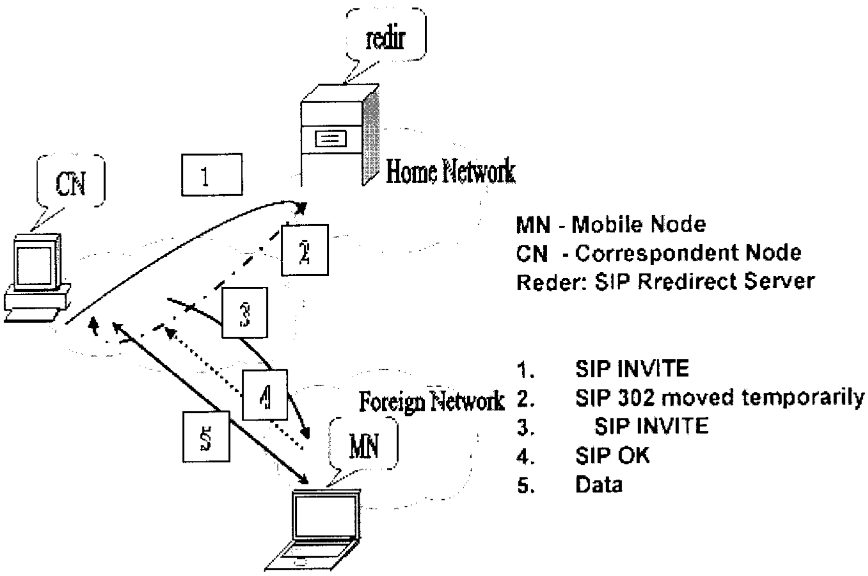

Figure 4-4: Pre-Call Mobility [38]

Steps for Pre-Call Mobility

- Once the MN moves to a new network, the MN registers its new IP address with the Redirect Server in the Home Network. The Correspondent Node (CN) then sends a SIP invite request to the Home Network.
- The SIP home proxy responds back to the CN with a SIP 302 "moved temporarily" message. This message provides to the CN information regarding the location of the MN with which it wishes to communicate.
- Using the information provided by the Redirect Server, the CN sends a new Invite message to the new location of the MN.
- The MN responds back with an OK message.
- A session is established between the CN and MN.

4.5.2 Mid-Call Mobility

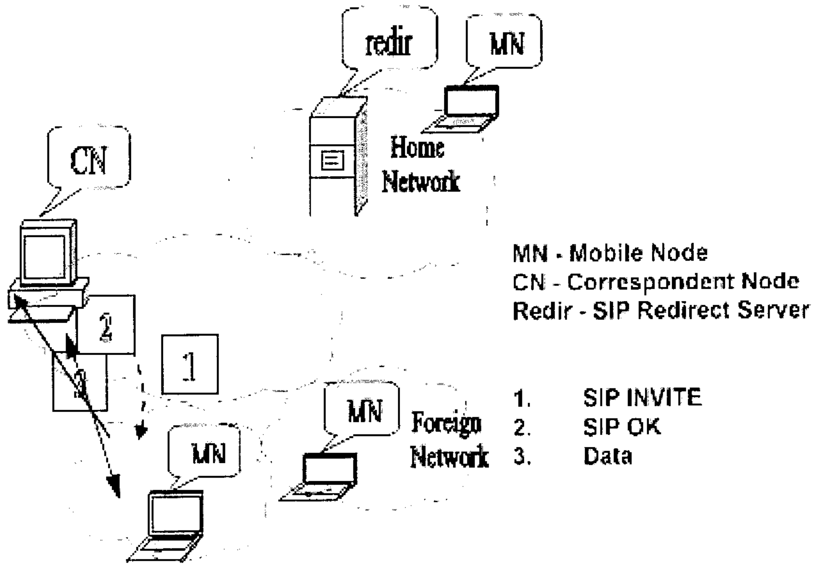

Figure 4-5: Mid-Call Mobility [38]

Call Flow for Mid-Call Mobility

Mid-Call Mobility means that the session is "on" at the time the MN moves out of the Home Network and into to a Foreign Network. Once the MN detects it is in a Foreign Domain, it sends a request to the DHCP Server to acquire a new IP address from the Foreign Network.

- Once the MN receives an IP address, it sends a Re-Invite Message to the CN in order to update its records with the new IP address.
- The CN responds with a SIP OK (200) message, accepting the establishment of session again.
- Once the CN receives the new IP address of the MN, a new call is set up between the CN and the MN with the new IP address.

4.5.3 Session Mobility

Session mobility allows a user to maintain a session while retaining the liberty to change terminals. For example, a caller starts a session on his personal mobile device while on the way to work. After reaching his office, the same user might want to continue the session using his desktop PC. A user may also want to move parts of a session to different devices. For example, the user may wish to spread a set of media across a series of devices that are each specialized for a particular medium, e.g. if he has specialized devices for audio and video such as a video projector, video wall, or speakerphone. SIP can successfully implement such session mobility (or mid-call mobility) by using the Re-INVITE method. This is an INVITE method used while a multimedia session is in progress. While keeping the existing session alive, new terminal(s) can be added to the session and existing ones can be removed. Changes could also be made to the parameters of the session in progress in order to match the capabilities of the newly added terminal(s).

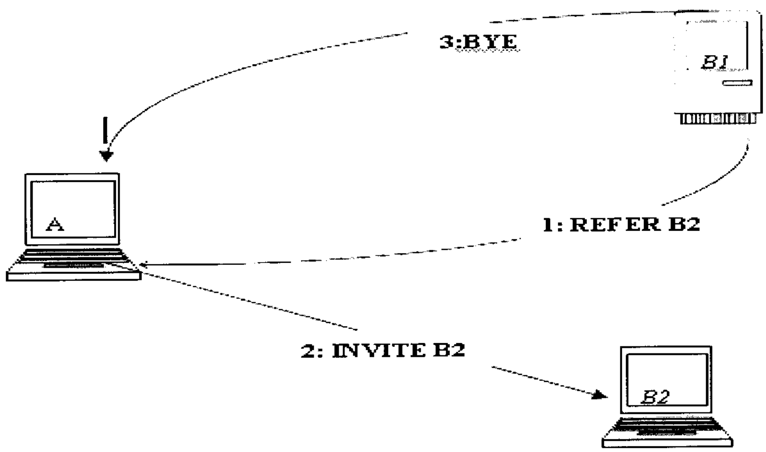

Figure 4-6: Session Mobility [38]

Session Transfer Steps

- While the session is in progress between A and B1, B1 informs A of a request to transfer the session to B2.
- Client A sends an INVITE message to B2 to set up the session.
- B1 sends a BYE message and terminates the session with that device.

4.5.4 Personal Mobility

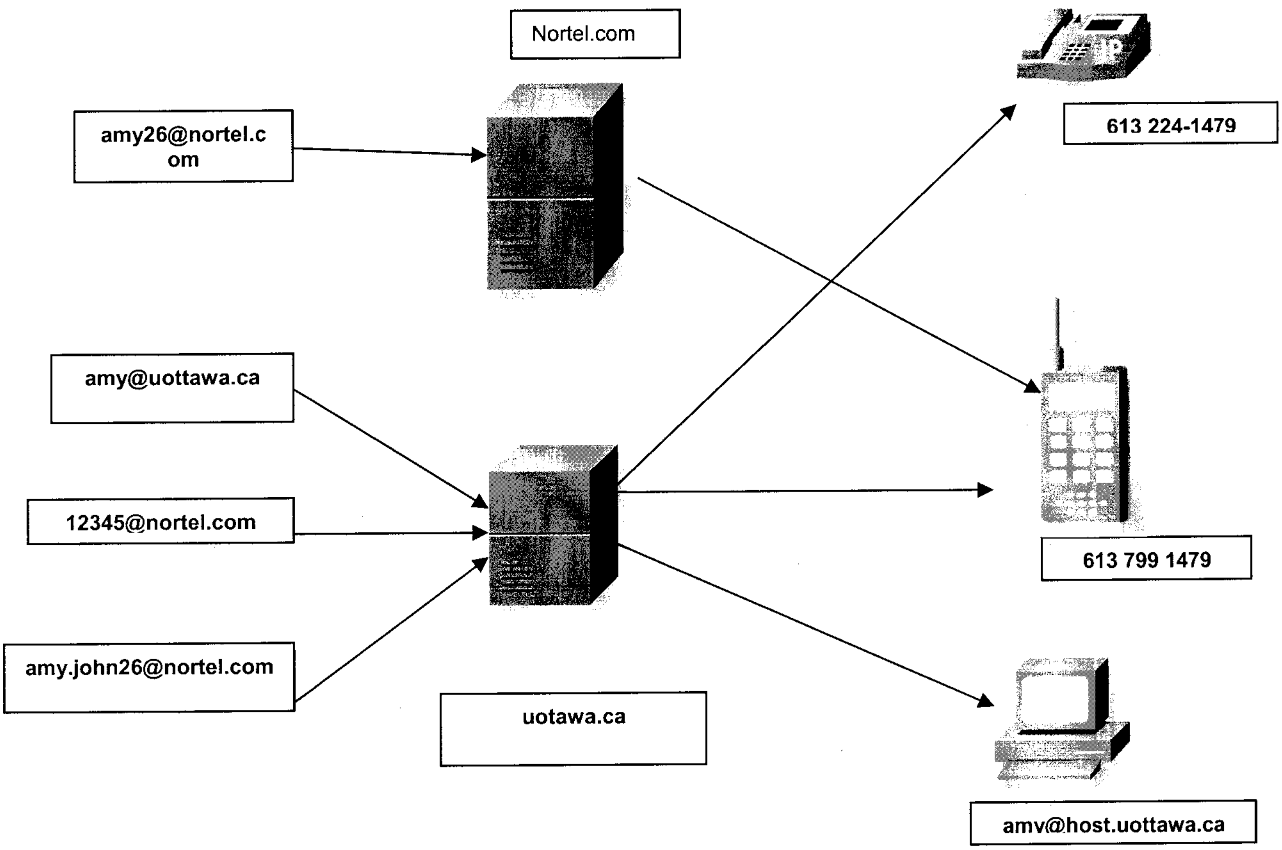

Figure 4-7: Personal Mobility [38]

Personal mobility allows a single logical address to be assigned to single user located at different terminals. Both 1-to-n (one address, many terminals) and m-to-1 (many addresses reaching one terminal) are possible as shown in the Figure 4-7 which illustrates some of the possible mapping examples. User Amy may want to be reachable via a traditional circuit switch phone, a PC, and a PDA. She may use these devices all at the same time or alternate between them. Using SIP forking proxies, Amy could be reached at any one of the devices via the same address, making her device choice transparent.

One practical issue is that Registrars should be capable of matching different devices as belonging to the same person.

4.8 Latency Involved in SIP Based Mobility

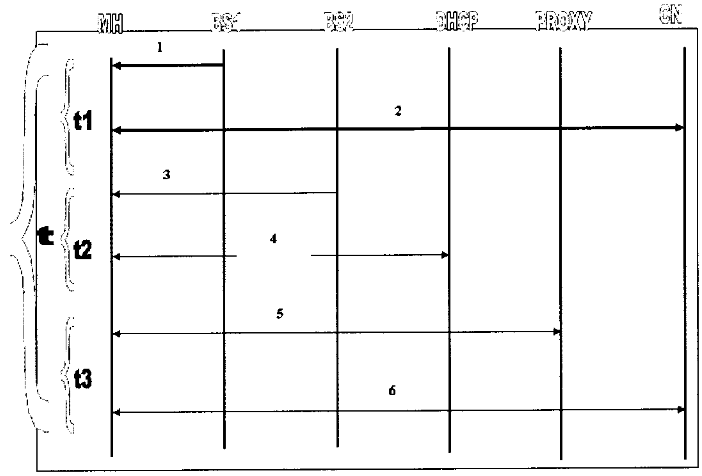

Figure 4-8: SIP Handoff Latency

Various delay factors involved in the SIP based handoff.

- **Base Station 1 (BS) binding:** Handover delay due to the movement of MN involves latency factors at different layers. For example, Layer 2 movement detection includes the identification of the new base station, IP address discovery and configuration and signaling to redirect the media to the MN's destination. This delay depends on the wireless technologies used. This delay could range from several 10 to 200 ms.

- **Handoff detection delay:** It is the time taken in detecting the movement of MN using Router Advertisement.

- **Address allocation delay:** Another major delay involved in the terminal handoff is the acquisition of the new IP address. And this delay depends mainly on the type of protocol being used, such as DHCP, PPP, etc.

- **Configuration delay:** It is the time taken in reconfiguring the network interface and setting some network parameters in order to communicate again.

- **SIP Re-Invite delay:** It consists of round trip time between the participants and processing time of a SIP Re-Invite message.

- **Session Establishment Delay:** Once the client has configured with the new IP address and its new domain, the signaling from the client to the CH to redirect the media to the new location of the MN will contribute the third trigger of the latency. It is the time delay between the arrivals of an OK message to the reception of first RTP packet by MN from CN.

t1 – Layer 2 Hand-over Latency

t2 – Delay due to IP address Acquisition

t3 – Registration and Media Redirection delay

t = t1 + t2 + t3

All the above mentioned delays are discussed in detail in Chapter 5. Out of the 6 delays discussed above the "Address allocation delay" and SIP Re-Invite delay are the major ones. Each takes about a few 100ms. And hence affects the quality of the real time communication adversely. So both these delays need to be studied in greater detail and minimized for better performance.

CHAPTER 5
PROPOSED MODEL

This chapter discusses the design of a new architecture for seamless mobility in converged IP centric networks. Section 5.1 provides an introduction to mobility management, while in Section 5.2 we elaborate on the issues that would be handled ivolved while handling mobility across heterogeneous network. Section 5.3 presents a brief comparative study between application layer and network layer mobility. Section 5.4 provides an insight into the design of the proposed model, whereas Section 5.5 discusses some of its advantages. Finally, Section 5.6 discusses the mathematical calculation of the handoff delay for the proposed model.

5.1 Introduction

Seamless mobility in converged IP centric networks provides an important paradigm for uninterrupted service in pervasive/ubiquitous computing environments. Seamless services require network and device independence, which allow users to move across different access networks and change computing devices as needed. IP convergence has led to the coexistence of several IP based wireless access technologies (e.g. GPRS [60], CDMA 2000 [88], Wireless LAN [61]) and the emergence of other next generation technologies (e.g., UMTS [62]), along with the diverse range of mobile devices that make seamless service provisioning an extremely nontrivial issue. The problem is even more difficult to tackle for multimedia streaming applications (e.g. Voice over IP (VoIP), video streaming, etc.), since these applications have stringent QoS requirements such as minimum bandwidth, delay, jitter and loss rate.

Mobility management protocols are in general responsible for supporting the seamless use of services offered to mobile users as these users change their point of attachment. This requires an on-time connection migration from one network to another. This action is known as vertical handoff. Thus, in addition to providing location transparency, the mobility management protocols used in this case also need to provide network transparency. Mobility management plays a very important role in heterogeneous networks, since a user could move between multiple types of access networks involving many service providers during a multimedia session. These access networks could be IEEE 802.11b, CDPD [63], CDMA, or GPRS based networks supporting DHCP or PPP servers. The movement of the MN can be between the access networks, with each access network belonging to the same subnet, the same domain but different subnets, or different domains altogether. In most cases the end-client would have access to both networks at the same time, but the connectivity to the networks would be determined by any local policy defined in the client itself such as signal strength or any other measurement based on the QoS parameters of the traffic.

5.2. Vertical and Horizontal Handoff Challenges

5.2.1 Vertical Handoff

The handoff should be as fast as possible. The goal is to provide a "seamless" handover with no perception of change from the user's point of view. Thus the duration of the handoff process should be made as short as possible.

Scenario Sketch [83]

Tom, a technology worker, is on his way to work, traveling in a vehicle. He is equipped with a WLAN and GSM/GPRS capable MN and he has set up a session through the GSM/GPRS radio interface with videoconference service that is offered by an application service provider (ASP). On his way to work he passes through several stations offering IP-connectivity through a WLAN access point. These traveler services offer a cheaper and higher bandwidth service than the public GPRS service his videophone call connection is maintained and handed over - without disruption between the GPRS and WLAN networks. The video quality is adapted to the available bandwidth: with the GPRS connectivity, the video is dropped, leaving only the audio signal. Tom is traveling in rush hour and heavy use of the network is made on the train stations. Therefore at these hot-spots his video call is only presented in a stamp-size format on his terminal. Once Tom arrives at his office premises, his company WLAN network is detected, followed by a log-on and again a seamless session handover of the video service from GSM/GPRS to WLAN. The video is now presented in full size with high resolution and high refresh rate.

Hence Mobile users are likely to move from coverage of on access network to another. These Access networks can have different characteristics, belong to different administrative domain, offer different features and services with different quality levels (eg: bandwidth), and cost differently.

Challenges handling vertical handoff in heterogeneous network

1. **Access Network Selection**: A user entering into a particular space wishes to select the appropriate network from the set of those available at that location at that time.

   The selection of the Access Network (AN) depends on

   - Application Service Provider's requirements,
   - Provider policies (e.g. agreements between user, access network provider and Service Platform provider about payment and usage rights),
   - Access network characteristics (such as available bandwidth),
   - User requirements (e.g. about quality of service, costs or service support),
   - User subscriptions (i.e. is the user allowed to use the network? Does the user have the right certificates and user credentials?)
   - Terminal capabilities (such as what kind of hardware interfaces are available and active in the terminal).

   Information regarding all these items is needed in order to seamlessly maintain sessions of the mobile user,with the connection optimally tailored to his (new) circumstances.

2. **Security and Authentication**

   Security and authentication are needed to prevent unauthorized use of resources and exposition of information to 3rd parties which are not supposed to have access to it. Not only does the user need to authenticated before becoming to make use of the network resources, also the operators must have a trust relationship with each other.

3. **Session Control And Mobility Management**

   A session, as an instantiation of a service, is controlled by Session Control functionality that cooperates with the Mobility Management functionality. Session control includes a set of non-service specific procedures to establish, maintain and terminate sessions and includes interactions between controlling parties and resources. The Mobility Management functionality is in charge of locating mobile users for session initiation and handing over their sessions between different network attachment points, as users roam across terminals and access networks. This section provides a high-level description of these functions and how they interact.

5.2.2 Horizontal Handoffs

Horizontal handoff is the changeover from one base station to a geographically neighboring base station supporting the same technology and belonging to same system autonomous domain i.e, same provider as a user roams about. Horizontal handoff is also referred to as intra-technology handoff. Every time a mobile cellular host crosses from one cell into a neighboring cell (that supports the same technology), the network routinely and automatically exchanges the coverage responsibility from one base station to another. Every base station change, as well as the exchange procedure or method employed, is known as a horizontal handoff. In a properly operating network, handoff takes place smoothly and efficiently, without gaps in communications and without uncertainty regarding which base station should be dealing with the MN. Mobile users need not get involved in order for horizontal handoff to take place, nor do they have to sense the handoff process or identify which base station is managing the signals at any given time. This thesis considers the handling of the horizontal as well as vertical handoff scenarios.

5.3 Proposed Solution

From the discussions on the above sections we could conclude that a single existing handover technology is not likely to be optimally suited for handling mobility over the heterogeneous network. As discussed in chapters 2,3,4 MIP and SIP are 2 popular and industrially accepted solutions for mobility. As network layer solution, MIP is suitable for mobility unaware applications. SIP on the other hand is suitable for situations where mobility awareness is required.

5.3.1 Intra-Domain Mobility

To handle fast mobility, the concept of intra-domain mobility was introduced. In this type of mobility, the MN hands off from one base station (BS) to another depending on the signal strength, but the nature of the network remains the same throughout. Whenever the routing path of a MN changes due to handoff, the data flows of the host need to be rerouted by the access network to the MN's new point of attachment. This change in the routing path should be transparent to the node. Since the nature of the network is homogeneous, the aspect that matters most is the amount of time it takes for the handoffs to take place.

MIP (Mobile IP) provides a transparent mobility management solution in the network layer for IP based applications. In general, application layer mobility management protocols such as SIP perform worse than lower layer protocols in terms of handoff delay, signaling overhead. This fact indicates that a pure SIP mobility approach would generate considerably higher signaling loads compared to MIP. SIP is much more generous in terms of message size, since its messages are text based. Thus, intra-domain mobility management is best handled by lower layer protocols like HMIP and MM-MPLS.

5.3.2 Inter Domain Mobility

When suitability for deployment in next-generation networks is considered, SIP is a better mobility management solution. This is because it obviates the need for protocol stack and infrastructure changes. The benefit of using SIP for (application layer) mobility management is that it allows applications to adapt their service behaviors, based on the mobility management strategy selected, in order to provide the best possible end user experience. SIP has the ability to modify the existing session by sending a Re-Invite message. This message would allow, for example, a session with both audio and video on a WLAN network to be changed to a session with audio only over a GPRS network. Note that after the final acknowledgment (ACK) of the Re-Invite message, the original media session is modified according to the newly agreed-upon session description parameters in accordance with RFC 3261 [8].

MIP (Mobile IP) and all of its modified versions proposed in the past operate in the network layer of the TCP/IP protocol stack and handle transparently the mobility of IP addresses in multiple access networks for applications. However, MIP cannot hide network characteristics, meaning that the behavior of applications can be unpredictable when the required network characteristics (e.g., a certain amount of bandwidth) cannot be met at the time a switching over to a new service area must be done. Unlike SIP, MIP does not have the capability of session parameters Re-negotiation.

5.4 Proposed Model

5.4.1 Overview

As mentioned earlier, this thesis deals with both intra-domain mobility and inter-domain mobility, in which a MN moves across a heterogeneous network. We propose an integrated scheme that combines both network layer and application layer mobility management to handle the seamless handoff, and a combination of Micro Cell Mobile MPLS (MM-MPLS) [7] and SIP. MM-MPLS handles the intra-domain handoffs while SIP technology is used to support inter-domain handoffs. This combination of network and application layer mobility management reduces the signaling load and provides fast handoffs for ongoing sessions.

5.4.2 Intra-Domain Mobility

Micro Cell Mobile-MPLS (MM-MPLS) to handling intra-domain mobility, is a new technique that integrates MIP and functionalities of MPLS to support intra-domain mobility. The concept of this model is that the MN registers with the nearest mobility agent at the lowest level of the hierarchy, which in turn registers with the agent nearest to it at the next level of the hierarchy.

When the MN needs to perform a handoff from one access network to another within the same foreign domain, it will send a Registration Request to the new Foreign Agent (FA). Then the new FA (i.e. node B) will relay the Registration Message towards the Foreign Domain Agent (FDA) of the domain. Each intermediate LSR from the new FA to the FDA will check whether there is an entry for that particular MN. The Registration Request message eventually reaches an upstream LSR (node A, Figure 5-6) the crossover LSR* (in the worst case, the cross over LSR is the FDA), that has a forwarding entry for the MN.

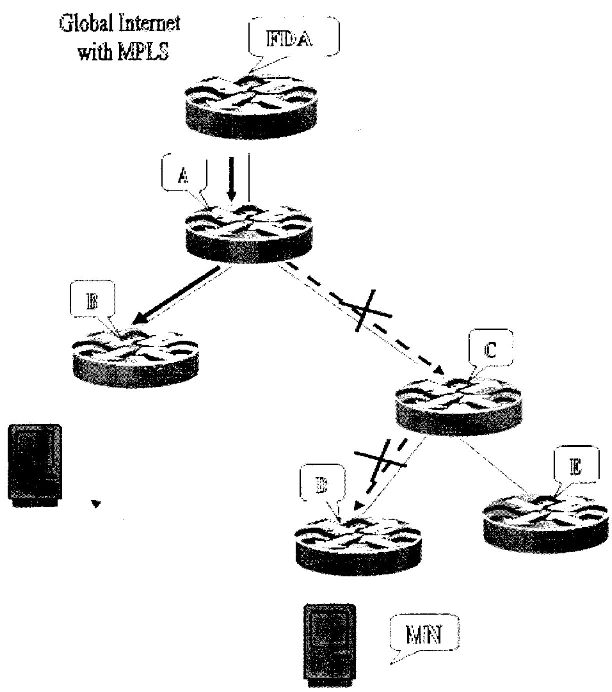

Figure 5-1: MM-MPLS Model [7]

Upon receiving the Registration Request message, the crossover LSR will set up a new LSP to the new FA using the mobile host's home address as FEC (Forwarding Equivalence Class). After the new LSP (from the crossover LSR to the new FA) has been established, the crossover LSR will change its label table and redirect the LSP to the new FA, with the MN's home address as FEC.

As depicted in Figure 5-1, when a MN first moves into a foreign domain, it will send a Registration Request message to the nearest Foreign Agent (FA) (node D). The FA relays this Registration Request message to the Foreign Domain Agent (FDA) of this MM-MPLS

* *We define crossover LSR as the LSR closest to the MN that is at the intersection of two paths. One between the (Foreign Domain Agent) FDA and the previous (Foreign Agent) FA, the other between the FDA and the new FA.*

domain. After the FDA receives this message, it negotiates with the Home Agent (HA) of the MN and establishes a LSP between the HA and the FDA. Then the FDA will set up a LSP between the FDA and the current FA.

In the above Figure 5-1, when a MN moves from FA-E to FA-D, a new registration message is sent to FDA. In this case, the intermediate LSR (node C) checks for the entry of the MN. Node C will find the entry for the MN and thus the registration message is not propagated further up the hierarchy. Instead, the packet for the MN is redirected by the nearest node (node C) to the new location. While the MN moves from FA-1 to FA-3, the registration message is sent by FA-2 to the FDA. In this case, node A becomes the crossover LSR, finding the entry corresponding to the MN. The packets destined to the MN will then be redirected by node A.

5.4.3 Inter-Domain Handoff

5.4.3.1 Handoff procedure

The procedure implementing a vertical handover can be divided into three steps:

- Network discovery
- Handoff decision
- Handoff implementation

5.4.3.1.1 Network Discovery

This is the initial process during which a MN searches for other reachable networks while moving across a network. A MN with multiple interfaces activates them to receive service advertisements broadcasted by different wireless service providers, using a variety of access technologies. The MN decides a wireless network is reachable if its service advertisements can be heard. The simplest way to discover reachable wireless networks is to keep all interfaces ON at all points in time. However, one of the major issues with keeping an interface active all the time is that it reduces the battery's life since it consumes energy even without receiving/sending signals. Therefore it is critical to avoid keeping the idle interfaces always in the ON state. Also, the time duration for discovery should be low so that the MN can benefit faster from the new wireless network.

Power efficiency and network discovery time are the most critical factors for evaluating the network discovery methods. Another option is that the interface may be activated periodically to receive service advertisements. The frequency of activities directly affects the network discovery time. It is outside the scope of this thesis to analyze the most optimal solution for the network discovery method.

5.4.3.1.2 Handoff Decision

A decision for vertical handoff depends on several parameters relating to the network to which the MN is already connected and the one to which it is moving to. For example, the decision to perform mobile controlled handoffs may be made by the mobile device based on policies such as network bandwidth, load, coverage, cost, security, QoS, or even user preferences. User preference is important when performing vertical handoffs. For instance, if the new network does not offer the required security, the user may still decide to use the old network.

Handoff Decision Criteria for heterogeneous networks

Handoff decision metrics are used to indicate whether or not a handoff is needed. In traditional handoffs techniques, only signal strength and channel availability are considered. In next generation heterogeneous wireless environments, new handoff metrics will have to be considered in addition to signal strength. We will discuss some of the common handoff decision metric parameters in this section. The choice of matrix parameter could depend largely on the type of network and the application to be supported by the network.

- **Cost**: Cost is a major consideration for users, since different network operators may employ different billing strategies. These variations in billing plans may affect the user's choice of handoff.
- **Network conditions**: Network-related parameters such as traffic, available bandwidth, network latency, and congestion (packet loss) may need to be considered for effective network usage. The use of such network information could be useful in handoff decision-making, in order to achieve load balancing across different networks. This also may avoid possible network congestion in certain systems.

- **Battery power**: Battery power may be yet another significant factor for handoff. For example, when the battery level is low, the user may choose to switch to a network with lower power requirements, e.g. an ad-hoc Bluetooth network.

- **Application types**: Different types of applications require different levels of reliability, latency, and data rate. The user applications running on a mobile device may also influence the handoff decision.

- **MN conditions**: MN conditions include dynamic factors such as velocity, moving pattern, and location information.

- **User preferences**: User preferences can be used to cater special requests for one type of system over another. Next generation handoffs will have to incorporate many of the above criteria rather than only signal strength in order to perform a handoff decision. Developing a handoff decision function that takes into account numerous factors remains a research challenge.

5.4.3.1.3 Handoff Implementation

The main contribution of this thesis is on the handover implementation. The analysis described here assumes that the MN is participating in an ongoing session when it intends to change domain by performing the handover

Before beginning the description of the proposal, we should perform a quick walkthrough of typical SIP based terminal mobility (mid-call mobility) as well as a study of the delay involved during a typical SIP handoff.

5.4.3.2 Typical SIP based inter-domain mobility

The main entities in SIP are user agents, proxy servers, and redirect servers, SIP Back-to-back user agent (B2BUA) server. Various methods are defined in SIP regarding INVITE, ACK, BYE, OPTIONS, CANCEL, and REGISTER. SIP terminal mobility requires SIP to establish a connection either during the start of a new session, when the terminal or the MN has already moved to a different location, or in the middle of a session. The former situation is referred to as pre-call mobility, while the latter is known as mid-call mobility. For pre-call mobility, the MN re-registers its new IP address with the redirect or SIP server in its home network by sending a REGISTER message; on the other hand, for mid-call mobility, the MN has to communicate with the DHCP server for that network and obtain a new IP address. Once the MN completes the AAA process, MN receives a new IP address and once a new IP address is assigned to the MN it needs to intimate the correspondent Node (CN) or the host, by sending a Re-Invite message containing the terminal's new IP address and updated session description parameters.

The CN starts sending data to the new location as soon as it receives the Re-Invite message. The MN needs to register with the redirect server of the home network for future calls. The high level messaging process of SIP-based mid-call mobility management is depicted in Figure-5-2.

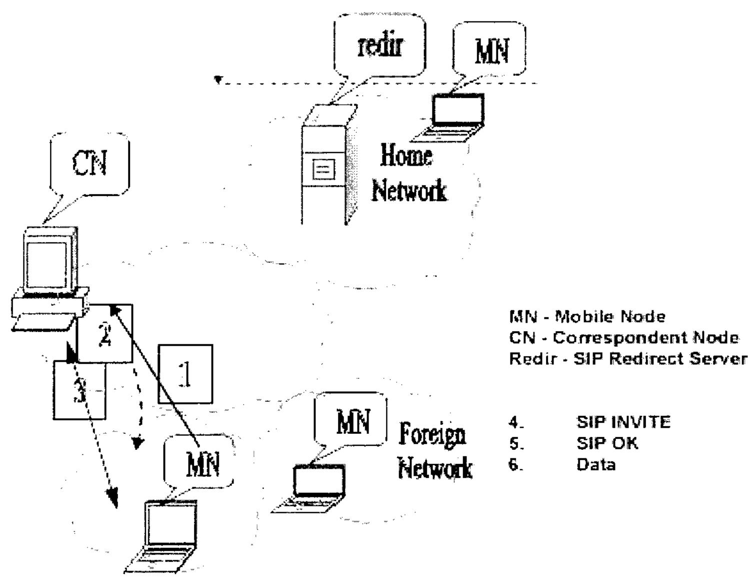

Figure 5-2: Typical SIP based Mobility

5.4.3.3 Handoff Delay generated when using SIP Based Mobility

In this section we analyze the major handoff delays involved during the mobility of a MN across a heterogeneous network.

1. Acquiring a new IP Address

When a MN moves into a new network, it needs to go through the Dynamic Host Configuration Protocol (DHCP) registration procedure to secure a new IP address for its new location/network. The message exchanged in the registration procedure is shown in

Figure 5-3. When the MN identifies itself as moving to a new network, the following steps occur.

- The MN broadcasts a DHCP DISCOVER message to discover the DHCP server in the network.
- The appropriate DHCP server sends out a DHCP OFFER message to offer service to the requesting MN.
- Upon receiving this OFFER message, the MN sends a DHCP REQUEST message to the DHCP server to confirm the offer made.
- The DHCP server then sends to the MN a DHCP ACK message containing such information as the new IP address assigned to the MN.
- The DHCP server also performs a Duplicate Address Detection (DAD) process which also contributes a considerable increase to duration of the DHCP process.

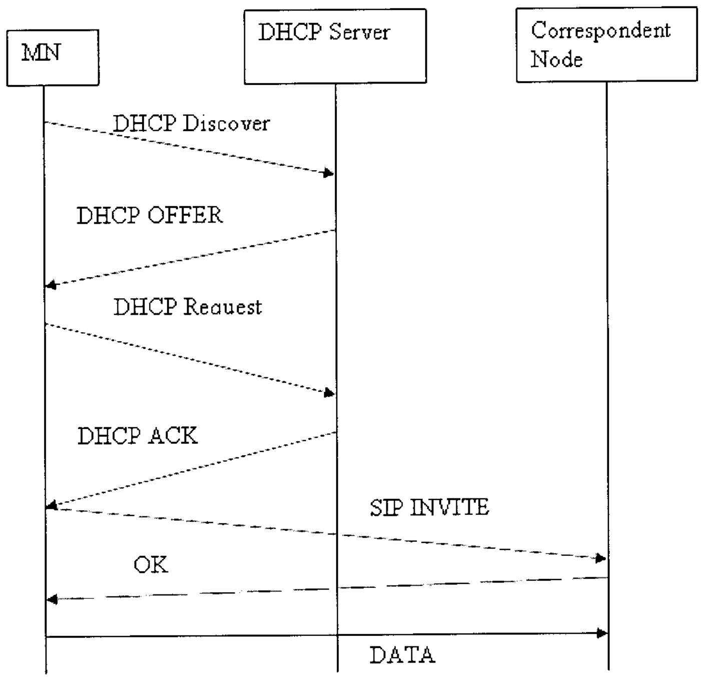

Figure 5-3: DHCP Registration Process

II AAA & QOS Process

When a MN requests for a handoff to the new network and tries to register with the new SIP proxy server, the SIP UAC generates the request, which may include credentials for Authentication, Authorization and Accounting (AAA) and for QoS.

The SIP proxy then sends a request to the AAA server asking if it is OK to provide a particular service to the MN making this request. It also contacts the Home network, gets the user profile, and negotiates the provision of QoS. The AAA server checks if the credentials are correct (authentication), and checks the user profile. If the user profile contains the service requested by the MN the AAA authorize provision for the service. The SIP proxy provides the service requested by the SIP UAC.

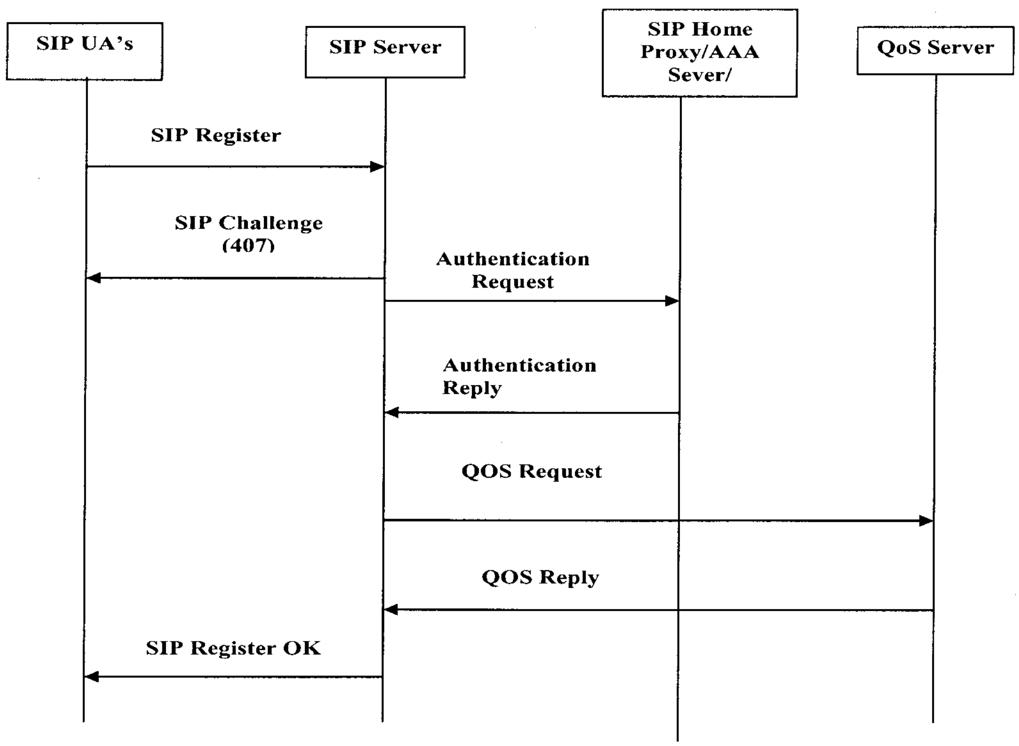

Figure 5-4: AAA & QoS process for inter-domain Handoff

III Session Re-Establishment

In a wireless environment, when a MN moves from one network to another and acquires a new IP address, other QoS, Server, and AAA credentials might also change. Thus, in order to keep a VoIP for example session going even after the participant acquires a new IP address, it is necessary to "modify" the session to reflect the new peer relationship and ensure proper multimedia delivery. In SIP, the MN sends a Re-invite message for modifying session parameters, including changing addresses or ports, adding media streams, and deleting media streams.

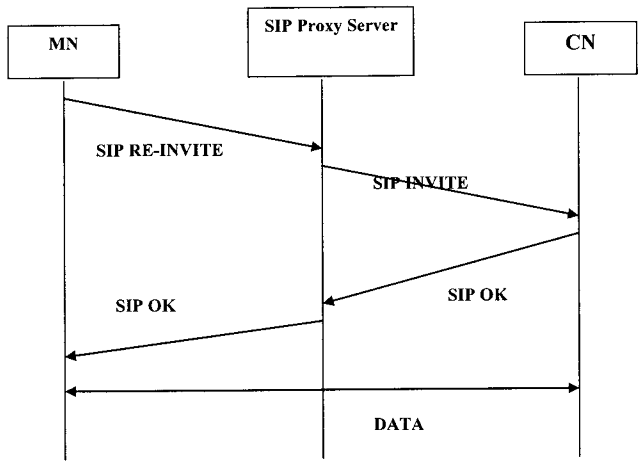

Figure 5-5: Sessions Re-Invite

5.4.3.4 Implementation of Inter-Domain Handoff

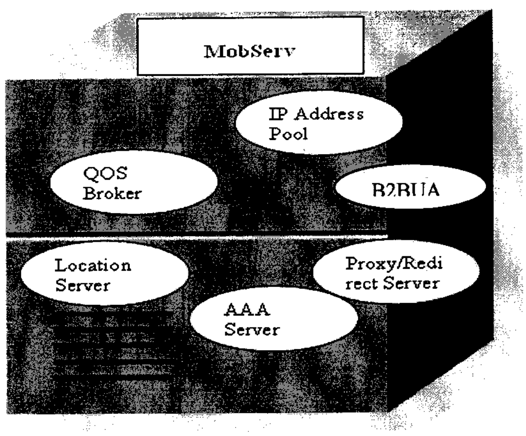

Figure 5-6: MobServ (Mobile Server)

In inter-domain handoffs, we make use of the capabilities of an enhanced SIP Proxy and B2BUA server having some additional capabilities which will be mentioned in the upcoming sections.

In this dissertation we introduce a new network element called Mobile Server (MobServ). From now on the SIP Proxy/B2BUA server will be referred to as MobServ. Figure 5-6 shows the capabilities of the Mobile Server. Each bubble represents a feature that is added to the functionality of a conventional SIP proxy and B2BUA server

Handoff allows the established call sessions to continue without interruption when a MN moves from one network to another. A successful handoff requires registration, configuration, dynamic address binding, and location management functions. The hand-off process must perform the necessary AAA to ensure the integrity, authenticity, privacy, and confidentiality of a user's location. Furthermore, handoffs should maintain the QoS requirements whenever possible which among other's means, striving to meet the end-to-end, delay jitter and packet loss requirements of real-time applications, as well as protect traffic application from experiencing sever degeneration and/or unintended termination.

5.4.3.4.1 Capabilities of Mobile Server (MobServ)

5.4.3.4.1.1 Dynamic Address Configuration

IP Address Pool: MobServ is configured with the ability to serve as a local IP address pool. MobServ holds a pool of IP addresses that belongs to the different independent Autonomous System (AS) domain operating with in its locality with which the service provider operating the specific MobServ has service Level Agreement (SLA). The MobServ basically maintains a separate pool of IP addresses for each network it manages. It allocates an IP address corresponding to the domain, the MN wishes to move.

Procedure for Configuring a Local IP Address Pool in MobServ:

MobServ is configured to obtain IP addresses from the DHCP server. The MobServ requests a set of IP addresses from the respective DHCP server of every network that that is under the management of the particular MobServ. The set of IP address requested from and provided by certain AS network domain is allocated to the specific MobServ for a defined period of time called the "lease time". Once the lease time is over, either the lease over this address/addresses has to be renewed or the address will be allocated back to the DHCP server IP address pool.

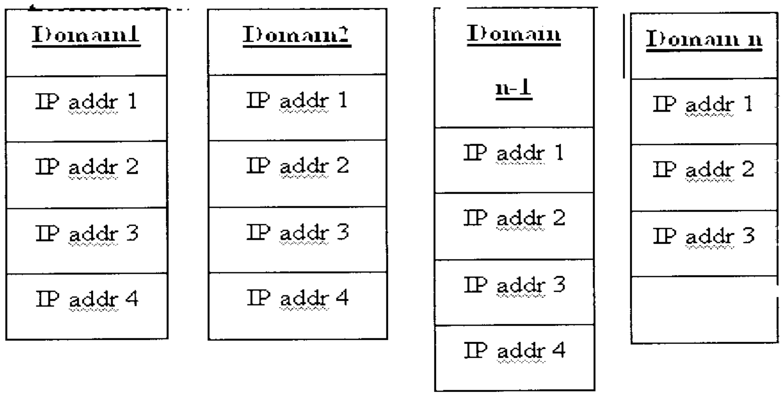

Figure 5-7: Mobile Server IP Address Pool

If a local address pool has a lower threshold value "x" (where x >0), at any point of time the number of usable IP addresses should not be less than this threshold value x. The value of x can be made dynamic by studying the number of MN in the network for each period of time.When the free IP address reaches x, a new range of IP addresses is requested from the DHCP server to extend or supplement the existing range of addresses. The IP addresses in each pool of the MobServ are allocated only once. Once an IP address is released by the MN it cannot be allocated to any other MN, untill those set of IP's are released back to the DHCP server, and requested again by the MobServ (this is done inorder to prevent the situation where same address is being allocated to more than 1 MN). Once the MN releases the IP address back to the MobServ, or if a local address pool is deleted, then those addresses are released back to the DHCP Server. Otherwise, the DHCP server has the power to takes back the IP address when the lease period is over, unless the MobServ have already renewed the lease. By maintaining a pool of IP addresses, a considerable reduction to the handoff delay is possible due to the elimination of the time duration associated with process of acquiring a new IP address from the DHCP server every time the MN moves into a new network, and the time consumed while waiting for the Duplicate Address Detection (DAD) process to complete.

5.4.3.4.1.2 Authentication, Authorization & Accounting (AAA)

Interaction between the SIP server and AAA server is intended to provide a mechanism that monitors the communication activities of the user for accounting and auditing purposes.

The hand-off process must perform the necessary AAA to ensure the integrity, authenticity, privacy, and confidentiality of a user's location. The AAA server is a network server used for access control. Authentication identifies the user and implements policies that determine which resources and services a valid user may access. The AAA Server in a foreign network grants or rejects an access request to a roaming user after communicating with the Home AAA of the MN. It is important to establish security associations between MN and the networks that serve them. There have been a few AAA protocols created to handle these security associations as a MN moves between domains. Generally, the Home AAA server or an intermediate broker agent (SIP Proxy) must be contacted when the user moves into a new domain for the first time in order to establish its credentials. It is important to complete the registration process in a timely manner during the handoff process, and it is thus critical to minimize the time spent creating new security associations.

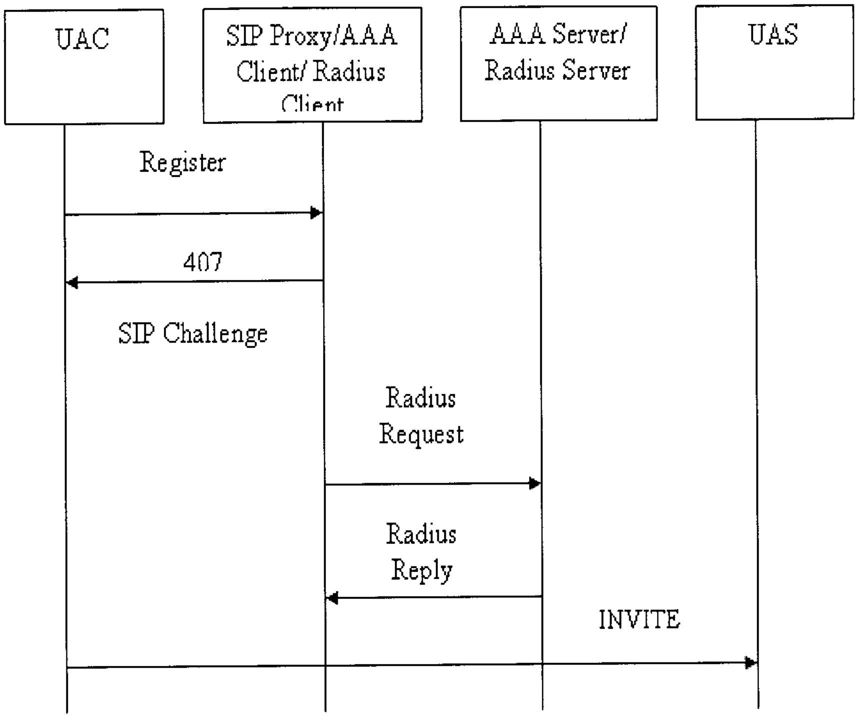

Figure 5-8: Mobile Server IP Address Pool

The introduction of AAA functionalities adds an undesired delay component when the user requests network access and hands off at the same time. AAA SHOULD NOT unduly burden call setup times where appropriate. It may be reasonable to support some delay during registration, but delay during on-going sessions (especially real-time) is problematic.

Therefore, one of the main objectives is to minimize and if possible eliminate the additional delay introduced by AAA procedures. One of the major delays involved in the AAA process while the MN is moving is the transfer of AAA state information from the home/old network to the new network. The major motivation behind this concept is avoiding the reestablishment of AAA communication every time the MN changes networks. This solution contributes to the seamless operation of application streams, minimizes packet loss, reduces delay, saves on bandwidth over the radio link, and reduces errors. The AAA information is transferred from the previous MobServ or Home MobServ to the New/Visited MobServ and maintained by the database in the current MobServ as long as the MN is within its administration. This AAA context transfer can be implemented using different protocols. One such example is EAP (Extensible Authentication Protocol). EAP provides a mechanism for supporting various authentication methods over wireline and wireless networks [69]. EAP [69] allows wireless client adapters to communicate with different back-end servers, in the, RADIUS (Remote Access Dial-In User Service).

One widely implemented platform for authenticating users and authorizing them for selected services is the RADIUS protocol [69, 70]. RADIUS allows the secure transfer of information between a Radius Client (often called the NAS – Network Access Server) and a Radius Server (which has access to the user's account information typically stored in a separate database). RADIUS does this and guarantees security of: (1) the Radius Client's authenticity, (2) the RADIUS Server's authenticity, (3) confidentiality of the user's password, and (4) integrity of any other account information passed between the Radius Server and Radius Client. This is accomplished by an out-of-band negotiation between the Radius Server and Radius Client of a "shared secret" that should be known only to the two entities.

The visiting MobServ acts as the RADIUS client and gets all the AAA information from the RADIUS server during the initial registration. Later, for all the handoff across the network under the administration of the MobServ, the MobServ will not have to communicate with the RADIUS server for the purpose of granting access to the MN in the new networks. The messaging involved in the implementation of AAA procedure is presented in Figure 5-8.

5.4.3.4.1.3 QoS Broker

Next generation wireless systems will consist of various access technologies with differing parameters interconnected with a common IP-backbone. Mobile terminals will run various types of multimedia applications. These applications will have different requirements some of which may not be satisfied by the best-effort IP framework. Providing proper QoS in terms of bandwidth, throughput, reliability, perceived quality, and cost in a heterogeneous mobile computing environment presents a major challenge. The major problem in providing QoS in a mobile environment is the coupling of QoS and mobility management, i.e. determining how to keep providing the same level of quality to the packet flow both, during and after the handoff.

In typical SIP based mobility, the SIP Proxy/Server is responsible for receiving SIP requests, interpreting the SDP part, and, based on the retrieved user profile, validating whether the user is authorized to use the service, the required codecs, and the correspondent maximum bandwidth.

In the model proposed in this thesis, the MobServ (advanced SIP proxy) maintains a policy decision engine, a network context processor, a session context database, a user context database, a presence database, and a network database. At first, the user sends a registration request "REGISTER" to the MobServ. The MobServ gathers context information through SIP messages and stores them in a specific database. The MobServ monitors periodically the situation of the networks that are under its administration with the SIP SUBSCRIPTION, NOTIFY messages [66].

When the MN sends a handoff request to a new network, the MobServ makes decisions depending on the policies and current network state, as well as appropriate negotiations are performed and resource allocations are made. For example, higher bandwidth and lower jitter and delay thresholds are set in order to maintain the quality of a voice session. By maintaining the user's information, network resources, etc., the MobServ will not have to contact the home server every time the MN requests a handoff. Figure 5-9 depicts the big picture of the QoS broker.

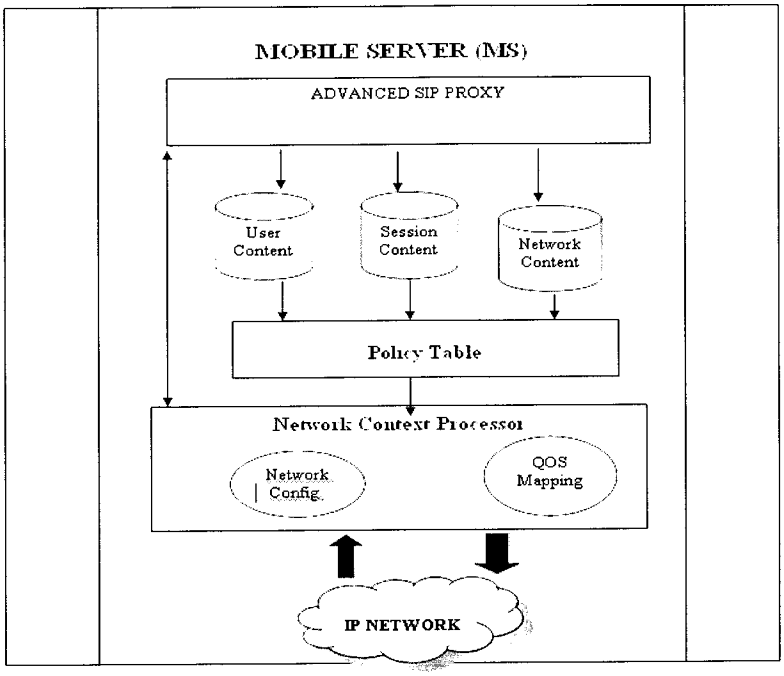

Figure 5-9: QoS Broker [66]

User Content: This database maintains the information including the terminal capabilities, e.g. device category, screen size, the device acceptable content types, the available service information such as service level agreement etc. This information can be acquired from the SIP REGISTER message.

Network Content: This database plays a major role in the adaptive service provisioning. QoS largely depends on the network's state. The services must be provided according to the pre-determined network condition. If the network cannot handle the service request, the service, after handoff, will fail to provide the resource. MobServ continuously monitors the network and keeps itself informed about the status of the network and network resources. Hence the MobServ itself can determine whether access satisfying the requirements can be provided.

Session Content: The session context database maintains the current status of session, flow characterization information of the session (SDP), the source/destination address and port and other information associated with the session. Generally this information is usually collected from the corresponding fields of the SIP messages (INVITE, REGISTER).

SIP QOS Message Flow

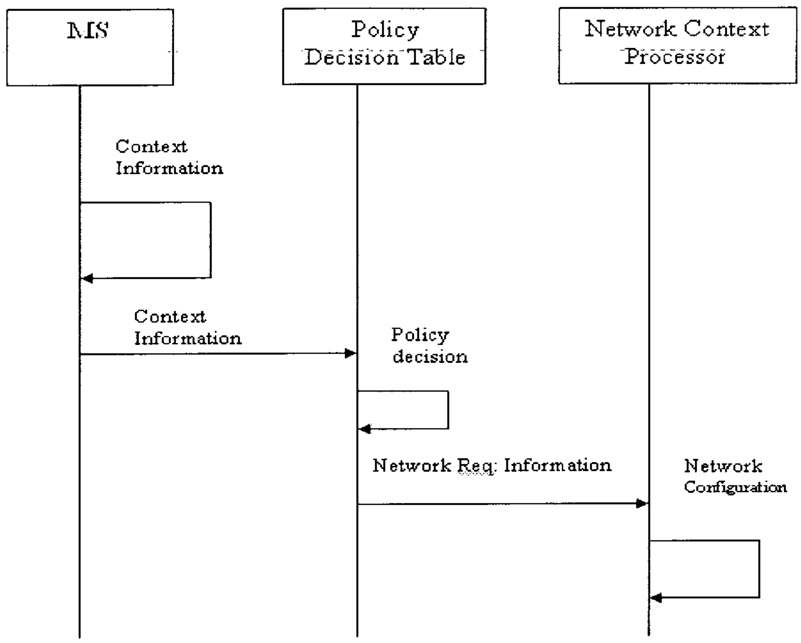

Figure 5-10: MobServ QoS Broker Network Resource Adaptation [66]

MobServ collects and frames the context information from the user context database, session database or from the Re-Invite message of the MN, forwards this information to the policy table. Where it frames the network requirement and negotiates with the network context database/processor to reserve the network resources. So that the required network resources could be reserved for the session.

B2BUA Server: IETF standard (RFC 3261) defines a back-to-back user agent as "a logical entity that receives a request and processes it as a user agent server (UAS). The B2BUA acts as a user agent client (UAC) and generates requests. The B2BUA server maintains a dialogue state and participates in all requests sent on the dialogues it has established. The standard definition of B2BUA server is that it is a concatenation of a UAC and UAS.

A B2BUA server is inherently "session stateful" since it acts as a terminating user agent end-point and holds ad-hoc mapping between the different sessions.

The B2BUA entity of the MobServ intercepts the RE-Invite message send by the MN and sends the reply back to the MN on behalf of the CN.

5.4.3.4.1.4 Location Server

The Location Server keeps track of the location of the MN

5.4.3.4.2 Inter-Domain Handoff Steps

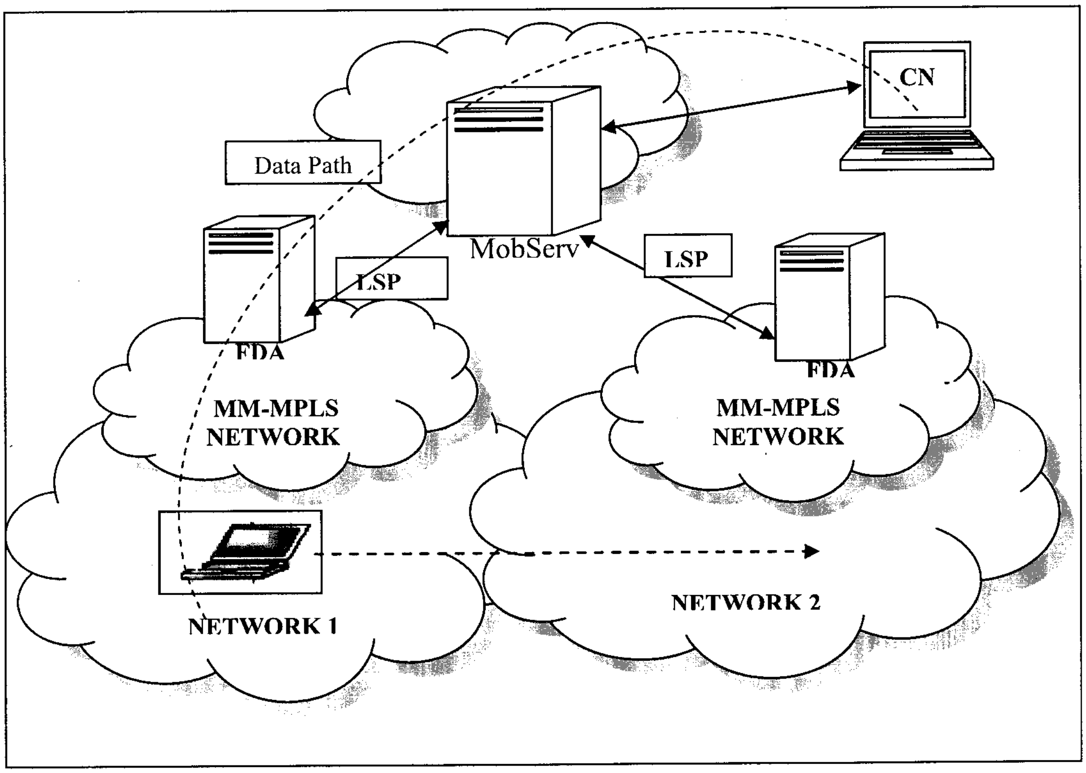

Figure 5-11: Inter Domain Mobility Model

The above diagram (Figure 5-11) depicts the procedures involved in inter-domain handoffs. In inter-domain handoffs, the MN moves from one network to another. The movement is between networks under different administrations that might support different features, offer different types of QoS and AAA, etc.

When a MN first enters the network, it registers itself with the Mobility Server (MobServ). The MobServ searches whether this particular MN is registered with the MobServ. If the MN is being registered for the first time, the MobServ server assigns an IP address to the MN from the IP pool. It also transfers the AAA and QoS information to the visited MobServ. The details of the context transfer are discussed in detail in the above section.

When the MN moves into and is jointly covered with another network, it begins receiving signals corresponding to network advertisement identifying another network, if it identifies the network by "reading" network advertisements sent out by its router, the MN reports the receiving messages and the signal strength. Once the MobServ receives the handoff request message from the MN, it goes through a process, implementing the criteria discussed in detail in the previous section, and decides whether the MN is eligible to be connected to the network net_{fd}. Once the MobServ receives the handoff requests of the MN, it looks up the AAA policies it maintains for both the MN and the new network and authenticates the MN for accessing the new network. Once the MN authentication to access the new network is successfully completed, Mobserv will assign a new IP address for the MN. The MN sends a RE-Invite message to the CN defining the new set of capabilities supported by the new network into which the MN has moved. This message is intercepted by the B2BUA server in the MobServ and a 200 OK message is sent back to the MN. Hence, a new session will be established. Now the original media session is modified according to the new agreed-upon session description parameters. An LSP between MobServ and the FDA of the new domain is established, even before the handoff to the new network is completed.

5.4.3.4.2.1 Inter-Domain Handoff Procedure

Figure 5 -12 shows the Handoff features in inter-domain handoffs

- A handoff request by the MN is sent to the SIP Proxy.
- The MobServ determines whether the MN can be granted the requested handoff depending on the various parameters like the visiting network resource availability status the MN demands, AAA policies of the MN and the visiting network etc. MobServ reserve the resources required and sets up the LSP with the

new FDA of the network into which the MN is moving to even before the MN actually moves into the new network.

- Once the MobServ has decided to approve the requested handoff, it searches the IP address pool (corresponding to the visiting network), selects and assigns a new IP address to the MN.
- A new LSP is established between the MobServ and the FDA of the new network.
- Then the MN sends a “Re-Invite” message to the CN.
- The SIP B2BUA entity of the MobServ intercepts this message and terminates it by responding to the MH with a SIP 200. And hence establish a new session.

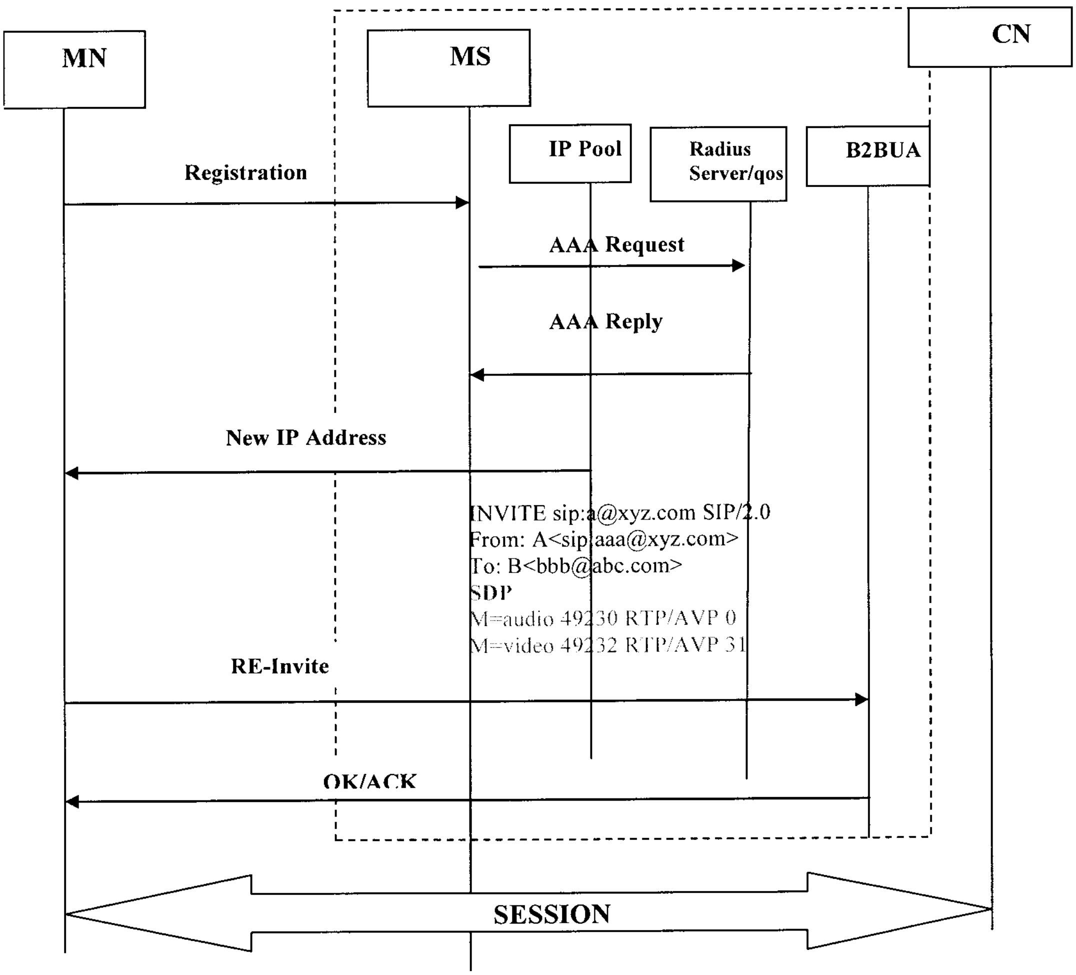

Figure 5-12: Message Flow

TCP Session Mobility

One of the known draw backs of SIP is SIP's incapability of handling mobility management of long term, connection-oriented TCP connections. The traditional SIP based Re-Invite and re establishing to handle mobility of the session would result in elimination of the ongoing TCP session, and this situation is highly unfavorable.

SIP identifies its call/session using a SIP session ID, while a TCP connection is identified by using the IP address of the MN and CN, and the TCP port. However, as a MN moves, its IP address changes. In order to support TCP applications, the proposed approach should spoof constant TCP endpoints (IP address and port address) despite change of their IP addresses.

One advantage of the model presented in this dissertation is that this model could be extended to handle the mobility of TCP session. The B2BUA entity of the MobServ has the capability of maintaining a record and status of ongoing TCP connections (SIP session ID). MobServ also maintains the information regarding the current (New) IP address, and original IP address (IP address at the beginning of the TCP session).

TCP Session Handoff Procedure

- Once the MN has successfully registered with the new network it has moved into, it receives a new IP address from the MobServ.
- Now the MobServ would bind the MN's old address to its new one.
- Then it would make use of encapsulation [89] and start forwarding the TCP packets bearing original destination addresses to the new address of the MN.

The key advantage of this approach is that TCP stays unchanged even though the IP address and the location of the MN changes. Also the B2BUA deployed in the MobServ would support multiple TCP sessions handled on behalf of multiple CN and MN at the same time.

5.5 Advantages of the proposed model

- By integrating MM-MPLS, SIP and MPLS technologies, we have developed a technique that handles the two broad varieties of mobility; intra-domain mobility and inter-domain mobility.
- Through this integrated model, triangular routing one of the major issues of MIP is avoided.
- The technique is taking advantage of the key property of SIP, its ability to hide network characteristics while handling mobility across domains.
- One of the major delays involved in SIP based mobility is that when the MN moves into a new network; it has to obtain a new IP address from the DHCP server. In our model, we completely eliminate the delay involved during acquiring a new IP address. The MobServ maintains separate pool of IP addresses for each network that it manages. Thus, as soon it receives a request from a MN, it assigns a unique IP address from its pool of addresses. The delay occurring during execution of the signaling process for getting a new IP address, the delay due to DAD, and the database delay occurring by the DHCP process is therefore avoided.
- The MobServ has established the LSP with the FDA of the new network into which the MN is moving to, even before the MN actually moves into the network.
- The exchange of Re-Invitation messages between the MN and CN, which helps to modify and thus update the session, is one of the major advantages of the SIP protocol. This is one of the essential requirements of handling mobility across next generation network.
- Introduction of the MobServ also helps to handle the mobility of the TCP sessions.

5.6 Calculation of Inter-Domain Delay

Table 4: Inter Domain Handoff Delays

Handoff detection time	Accepting the handoff + Assigning a new IP address + AAA process + Qos negotiation	RE-INVITATION Message (**$T_{re\text{-}invite}$**)

- **T_{hand}** → The time MN send a handoff request to the MobServ (when it start getting becons from the new network) The connection from the previous network is not lost completely.

- $\mathbf{T_{req}}$ → Time consumed for carrying the following process:
- to accept the handoff + Assigning an IP address to the MN + AAA process + Negotiation of QoS for the MN.
- $\mathbf{T_{re\text{-}invite}}$ → Duration for the new session to be set up between MN and CN.
- $\mathbf{T_{re\text{-}invite}}$ depends on the distance between CN and MN. In n our model the Re-Invite message gets intercepted by the MosServ and the reply is sent back by the Mobserv, thus in our model it depends on the distance between MobServ and the MN.
- $\mathbf{T_{req}}$ is the time duration for having SIP proxy (MobServ) to assign a new IP address and complete the AAA process to the mobile for handingoff to a new network.

As mentioned earlier according to this model the proxy (MobServ) holds a pool of IP addresses. The delay involved in searching for an IP address from a pool of addresses as well as the time delay for assigning an IP address was considered zero. In reality it can be negligibly small. Since every Mobserv stores the AAA information regarding the MN which is moving in the network under the administration of a particular MobServ the delay involved due to the AAA process is also negligible.

So for our scheme the major delay occurring during the inter-domain handoff is the time delay for setting up a new session.

CHAPTER 6
RELATED WORK & MODEL VALIDATION

This chapter describes briefly the evolution of the SIP protocol for supporting mobility. In section 6.1 we provide a brief introduction In 6.2 we discuss some of the work already published towards SIP and MIP supporting mobility across a heterogeneous network. Section 6.3 validates the proposed model that is used in this dissertation by comparing it a related work, which is discussed

6.1 Introduction

In previous chapters 4 & 5 we have discussed in detail the SIP based terminal mobility proposed model supporting mobility respectively. In the upcoming sections we quickly glace through some of the works already performed to deploy SIP based heterogeneous mobility models and some of their shortcomings. And also presents the output for our OPNET model validation.

6.2 Related Work

This section presents some of the solutions proposed in the past for the enabling SIP to support mobility across heterogeneous network.

Research Work 1 [85] Wang Miao, Zhang Yu-Jun and Li Jun, proposed the model "Fast Handover Solution for SIP-based Mobility (FMSIP). In FMSIP they have made use of the concept "movement anticipation". The proposed technique assumes that a MN is capable of discovering the new router and its prefix; it then formulates a prospective IP address according to the new router's prefix while being still connected to the previous access router (PAR). The technique proposes that when a handoff is initiated PAR transmits a Handover Initiate (HI) message to the new access router (NAR). During the handoff initiation process, PAR sends the IP address to NAR for verification, i.e. to check if NAR is aware of any the new IP address is duplicate (another MN generates the same address at an earlier time). If there is no duplication, NAR validates the IP address and the MN does not have to perform the IP address renewal process, avoiding the long delay associated with this process. In the opposite case (the address is invalidated) the MN has to go through the address renewal process. Once the IP address matter is settled (with or without having to go through the IP address renewal process), FMSIP sets up a tunnel between PAR and NAR to forward packets. This would help MN to continue sending packets to CN using the previous IP address via this tunnel during the application handoff. The tunnel remains active until the MN completes the SIP session re-establishment with its CN. Through tunneling, application handover takes place in parallel with its communication process, so the application later handover time is reduced from handover time. The model also has addressed the delay involved in the AAA procedure during vertical handoff. In order to reduce the number of AAA message exchanges involved, FMSIP takes advantage of the AAA context transfer to forward the AAA pre-established information from PAR to NAR as a means of re-establishing AAA service on a new link.

Some of the major issues of the proposed model are the following. The model dose not completely eliminate the delay associated with the IP address renewal since there is a chance the MN will have to go through the IP address renewal process, meaning it will have to contact the DHCP server of the new network and go through the process described in section 5.4.3.3 (see also figure 5-3). Our model has introduced a fool proof mechanism by which we are able to completely discard the IP address renewal delay, due to the use of the MobServ and its involvement in providing the MN with a unique IP address.

Also the authors of [85] have limited their work in examining only the handling of real time applications., Issues related to the handling of ongoing TCP connections are not discussed. Our work proposes distinctive methodologies in order to handle real time as well as TCP traffic

Research work 2 [86, 78]: Stefano Salsano and Andrea Polidoro, in their work "SIP-Based Mobility Management in Next Generation Networks" have proposed the Mobility Management Using SIP Extension (MMUSE) mobility model. The basic idea of this model is to extend the signaling and media functionality of the Session-Boarder-Controller (SBC) to manage mobility. To achieve this aim, they have introduced a new entity, called the Mobility Management Server (MMS), located within the SBC. Another entity called the mobility management client (MMC) is also introduced and is located within the MN. When the MN needs to change the access network while it is engaged in a call, the MMC sends to the MMS a SIP message that contains the additional information required to identify the call to be shifted to the new interface. When MMC starts the handover procedures, it sends the handover request (SIP REGISTER) to the MMS and at the same time, it starts duplicating the RTP packets over both interfaces. In this way as soon as the MMS receives the handover message, the packets coming from the new interface are already made available. The MMS performs the switching and then send the reply back to MMC. When the MMC receives the reply message, it stops duplicating the packets.

A similar concept was proposed by Nilanjan Banerjee, Sajal K. Das and Arup Acharya in [78]. The handoff procedures are executed and handled at the base stations. Each base station is equipped with a SIP B2BUA and a SIP proxy server. The media gateway has dual functionality, one being a RTP *packet replicator* and the other a RTP *packet filter*. The MN is equipped with a packet filter. During the handoff period when a new network interface gets activated, the SIP UAC at the MN sends an INVITE message to the SIP B2BUA proxy server. It is considered that during the transient period, the MN has both interfaces operational and is capable of receiving packets independently. The B2BUA configures the packet replicator at the media gateway to send a copy of all packets directed towards the old interface of the MN to the newly activated interface. During the transient handoff period, the MN sends and receives the packets through both interfaces. The packet filter at the media gateway and the MN discards the duplicate RTP packets. As soon as the packets reach the MN through the newly activated interface, a re- INVITE message is sent to the CN with the IP address of the newly active interface and the corresponding contact information. As a result, the call parameters are re-negotiated on an end-to-end basis, and a new intermediate SIP proxy server and B2BUA belonging to the new base station are selected. Once the call re-negotiation is complete and as soon as a duplicate packet reaches the newly activated interface a BYE message is sent to terminate the call-leg through the initial interface. Finally, the MN registers its new location information with the home network's registrar service by using REGISTER message.

The author's of [78, 86] have considered handling of only vertical handoff, whereas mobility in a heterogeneous network involves more than just handling vertical handoff, as horizontal handoff is yet another major issue and it occurs more frequently than vericail handoff. In both the cases they have considered the MN being equipped with multiple interfaces, meaning, only those MN's equipped with such capabilities like Dual interface, a packet filter are supported for the solution mentioned above.

According to the architecture proposed in the model [78] the base station is equipped with all the SIP elements and the BS is responsible to manage all the data in order to execute the handoff procedure. The decision to enter the handoff process is made by monitoring channels and taking and comparing based on a quality measurement and for the execution of the handoff process. Similarly, in the work reported in [86] the authors have introduced an additional feature that needs to be embedded into the mobile device. Placing all required processing on the BS or additional capabilities to the existing MN is not a favorable approach and is an expensive also solution, since these BS's are under the control of certain service provider or is implemented for a particular technology. The two discussed models have tried to reduce the packet loss by sending duplicate packets during the handoff process, which increases the traffic and congestion in the network.

Research work 3 [75]**:** Jin-Woo Jung, Hyun-Kook Kahng, Ranganathan Mudumbai and Doug Montgomery came up with a model to address the vertical handoff issue integrating SIP and MIP.

The basic concept of the proposed architecture is that instead of completely replacing the existing Mobile IP approach, they proposed a model that integrates SIP and MIP to complement each other based on the kind of application (TCP or UDP) [75] i.e, if TCP traffic is being handled the MIP technique is used and vice versa uses a SIP network server, and a Mobility Agent (MA) with SIP Registrar. MA facilitates the location management. While the SIP network server handles call/session delivery, the Mobility Agent with SIP Registrar is used for handling location registration, location updates, and location queries. The authors have proposed two kinds of registrations. Registration process in mobile IP to inform the mobility agent of a MN's new IP address and update the binding information between home address of MN and the care-of address. SIP session re-establishment, to inform location of a MN's new IP address to correspondent host (i.e., SIP UA). This allows correspondent hosts to communicate with MN directly.

When a MN moves into a new domain while a call is in progress, the MN obtains an initial care-of address and also acquires globally unique Care-of Address (i.e. co-located care of address). When a mobile moves into a new network domain, it sends through the foreign agent a registration request to the home agent, requesting mobility binding update for a period of time. The home agent updates its mobility binding table and sends a registration reply back to the MN through the foreign agent allowing or denying that registration. At the same time, the MN must contact the DHCP server to obtain a new IP address. After receiving a new IP address, MN sends a SIP re-REGISTER message to the home agent (this process does not factor into the handoff delay). Once the re-Registration is completed, a MN does not send any more registration update message to home agent. Instead, the location information for the MN is updated by SIP's re-REGISTER and all packets for old call are tunneled via home agent. The tunneled route in mobile IP is only used until a new SIP re-Registration is completed. This model ignores any AAA (Authorization, Authentication, Accounting) delays that would incure during inter-domain hand-offs. Also it does not consider the IP address renewal delay.

In this work, the authors proposed an integrated SIP-MIP concept, instead of completely replacing the existing Mobile IP approach. The basic concept behind this work is that MIP takes over during that period of handoff in-order to complement the delay due to the SIP. This is not an ideal situation as, various research works and literature analysis [83, 80] has proved that, when suitability for deployment in next-generation networks is considered, SIP is a better mobility management solution. Also the triangular routing problem associated with MIP remains impacting the performance of the system. They have not considered providing a solution for the delay incurring because of IP address renewal and the AAA procedure as well as they have not discussed how the QoS for the applications could be handled.

The handoff delay in SIP based mobility may be substantial, causing considerable packet loss, which affects the quality of the voice or video streams seriously. Call disruption could occur if the new SIP session is not established while the MN is in the overlapped area. A MN using SIP-mobility always needs to acquire an IP address via DHCP. Depending on implementation, this requirement can be a major contributor to the overall handoff delay. In [75] empirical results show that some common DHCP implementations result in IP address renewal times of more than 2 seconds.

6.3 Model Validation

Since the model proposed in this dissertation has similarities to the architectural approach reported in the Research work 3 [75] we used our simulator to implement the architecture of [75] and compared the two simulation results for validation purpose.

The proposed model was constructed to resemble as close as possible the one described in the paper [75]. The validation model consists of MN, BS, SIP Home Agent with proxy server, Redirect Server, Foreign Agent, and their communications ie, SIP messages.

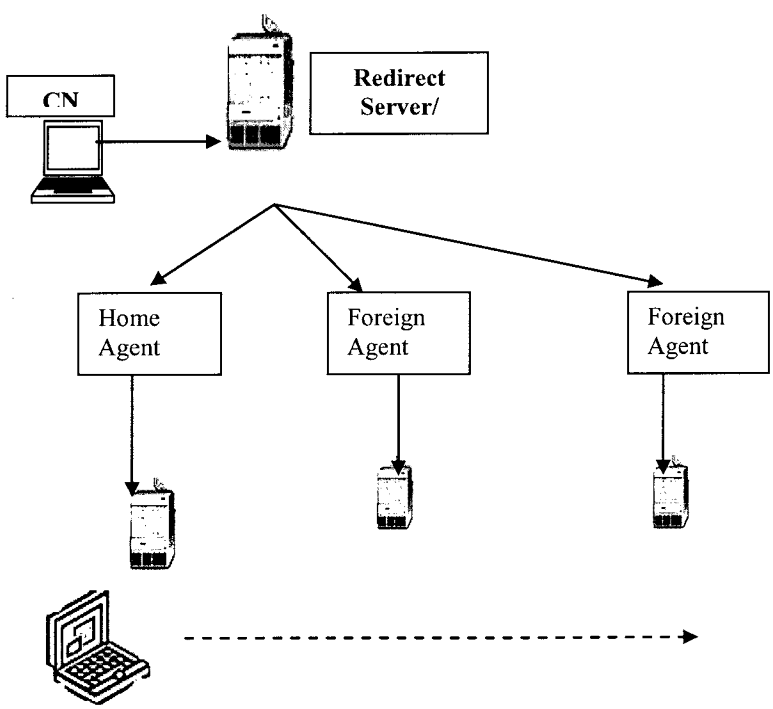

Figure 6-1: Validation Model

The proposed model is validated by comparing the simulation results presented in the paper [80] to those acquired by using our OPNET based simulation tool. The functionality of this model was explained in detail in previous sections.

All simulations are performed using the network topology shown in Figure 6-1. The simulated topology consists of a Correspondent Node (CN) that is streaming audio (or VoIP) data to a MN, and a Home Agent with Registrar and Redirect Server, The CN acts as a CBR source, producing fixed length packets {200 bytes: payload of 160 bytes and a header (RTP+UDP+IP) of 40 bytes} at the rate of 64Kbps All routers in the simulated topology utilize a drop-tail queuing strategy. Also, a MN connects to access points (BSs).

Reverse engineering had to be performed in the model in order to extract the values of parameters the author used in their model, eg: queuing delay, service time etc.

6.3.1 Validation Results

This section presents the model validation results. Validation mentodology and the simulation set up environment are explained in detailed in the above section.

In the work reported in [75], the authors consider only three factors that affect the handoff delay and packet loss:

- ***Toverlap***: the dwell time a MN is located in the overlapped area between old cell and new cell after detecting a requirement for movement.
- ***DoTOn***: the time it takes for a message sent by New Foreign Agent to reach the Old Foreign Agent or Home Agent.
- ***DmTOc***: the one-way delay from MN to Correspondent Node.
- In the research work 3 described above the renewal delay of IP address (consisting of the CoA delay of mobile IP and the DHCP delay of SIP-mobility) is ignored, because the study was focused on the disruption time and packet loss caused from location update delay.

In the work [80] simulation where carried out in 2 sets, making 2 different assumptions.

- Case 1: Handoff disruption time as the delay *DoTOn* increases while *DmTOc is* set to a constant value (30ms)
- Case 2: Disruption time as the delay *DmTOc* increases. *DoTOn* are assumed to be constant (30msec).

Case 1: Handoff disruption time as the delay *DoTOn* increases while *DmTOc is* set to a constant value

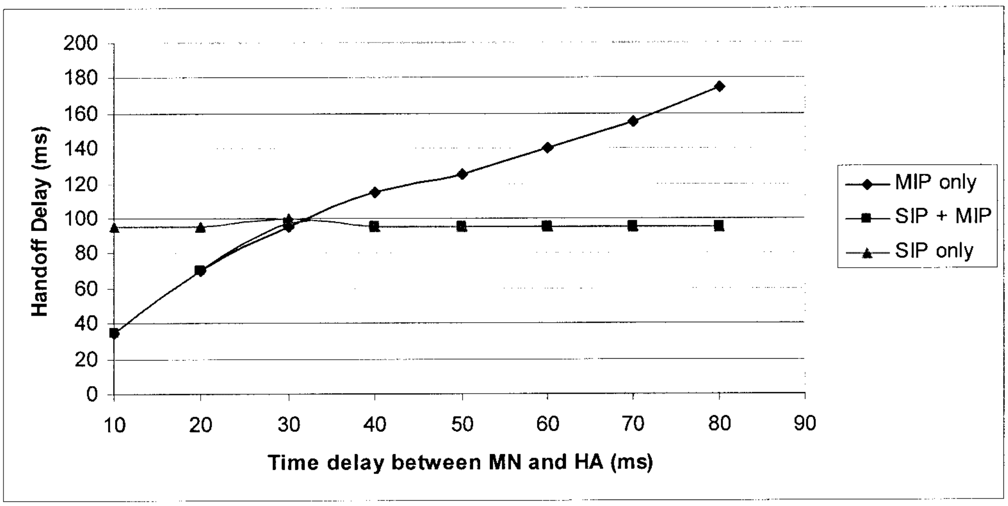

Figure 6-2: Handoff Delay vs Delay between MN and HA

The DmTOc value is set to 30 msec. Figure 6-2 shows the disruption time for vertical handoff results acquired through simulation, for the case of Mobile IP, SIP. The derived curves show that the disruption time when Mobile IP is used increases as the distance between the MN and its home network increases. However the disruption time for the SIP based mobility case is constant, about 95ms. This happens because the handoff delay when the SIP approach is used depends only on *DmTOc.* In the case of SIP-MIP integrated approach the MN and HA is below 30msec and then becomes constant. This is because the handoff delay is dominated by the component representing the time consume during the SIP Re-Invite process between MN and CN.

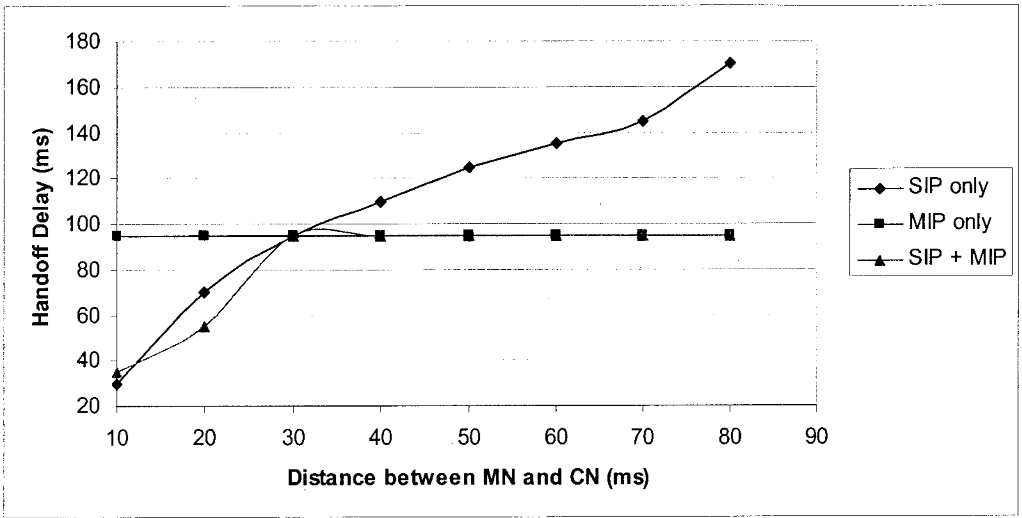

Figure 6-3: Handoff delay vs Distance between MN and CN

Figure 6-3 shows the disruption time versus delay *DmTOc*. The value of DoTOn is set t 30ms. Since the handoff duration when using SIP depends only on the distance between the MN and CN, the disruption time of SIP-mobility increases according to the value of *DmTOc*. The disruption time for MIP remains constant since the distance between MN and HA remains constant. For the integrated model the disruption time increases till the delay between MN and CN node is 30ms.

Basically according to their model the conclusion is that the disruption time is the minimum time for either of the approach (SIP or MIP**)**.

The results we have obtained are comparable to the result published in the paper [75].

CHAPTER 7
EXPERIMENTAL RESULTS & PERFORMANCE EVALUATION

This chapter presents the simulation results and performance evaluation of the model presented in this dissertation. Section 7.1 presets the simulation background and methodologies, it also discusses the various mobility models and traffic sources which are considered for simulation. Sections 7.2 present the simulation environment, and acquired results as well as discussing them.

7.1 Simulation Background & Methodology

To assess the performance of SIP over MM-MPLS method, we built a simulation model using OPNET. This section discusses the simulation methodology and simulation environment setup in detail. The simulation model contains important entities shown in figure 7-1.

1. All simulations are performed using a pair of CN and MN. For every simulation MN is considered to be moving across 3 different networks, performing 2 inter domain handoffs, and 2 intra-domain handoffs within each subnet.

2. The proposed model is measured for its performance by running Data, Voice and video traffic. Packet Delivery Ratio, Average Delay and Throughput are the 3 performance metrics used in this thesis. Their definition is given below:

   - **Packet Delivery Ratio** – It is the ratio of data packets received at the destinations to those generated by the sources.

   - **End-to-End Delay** – This is the time required for a packet to travel from source to destination.

Throughput – It is defined as the amount of data successfully delivered from the source to the destination in a given period of time.

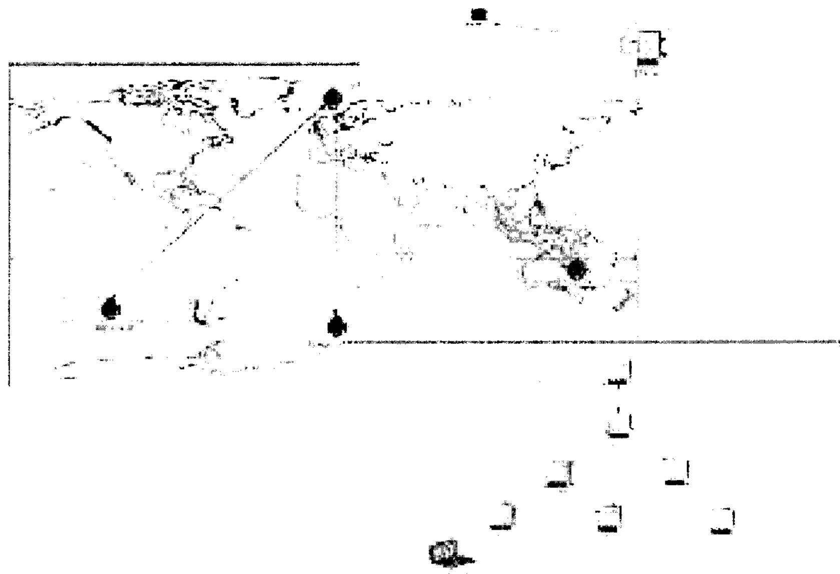

Figure 7-1: OPNET simulation model

3. **Mobility Configuration:** The two mobility models considered during the simulation with which the mobile node moves are 1) Random Walk 2) Random way point. Both are explained in detailed in section 7.2.

4. **Traffic Source:** To evaluate our model, simulation runs were performed using video, voice and data traffic sources. For non-real-time traffic, we are especially interested in the response time of web traffic because it is the dominant traffic in data packet networks. Apart from Web traffic we have also used CBR and Pareto traffic. The traffic model used in this thesis in explained in detail in section 7.1.1

7.1.1 Traffic Source Models

Data Traffic

3 kinds of traffic patters have been considered for data traffic source

- Constant Bit rate Traffic (CBR)
- Pareto Traffic
- WWW Web Traffic

Constant Bit Rate Traffic (CBR)

The simulation environment consists of a Correspondent Node (CN) streaming audio (or VoIP) data over UDP to a MN, with CN acting as a CBR source, producing fixed length packets at the rate of 64Kbps.

Pareto Traffic

In Pareto traffic distribution, packets are sent at a fixed rate during ON periods, and no packets are sent during OFF periods. Both ON and OFF periods are taken from a Pareto distribution with constant size packets. These sources can be used to generate aggregate traffic that exhibits long range dependency. The Pareto traffic distribution is considered to show Web traffic, assuming that Web object sizes form a Pareto distribution.

If X is a random variable following the Pareto distribution, then the probability that X is greater than some number x is given by

for all $x \geq x_m$, where x_m is the (necessarily positive) minimum possible value of X, and k is a positive parameter. The family of Pareto distributions is parameterized by two quantities, x_m and k.

It follows from the above that the cumulative distribution function of a Pareto random variable with parameters k and x_m is

WWW Web Traffic

Taking into consideration that WWW traffic has a significant presence in today's networks and it is expected that its volume will increase further in the future, it is important to have an accurate knowledge about how the proposed technologies work in the presence of such traffic. In our work we have used the traffic model introduced in [87]. In this paper, the authors have developed a traffic model which imitates closely a stream of HTTP requests originating from a population of Web users.

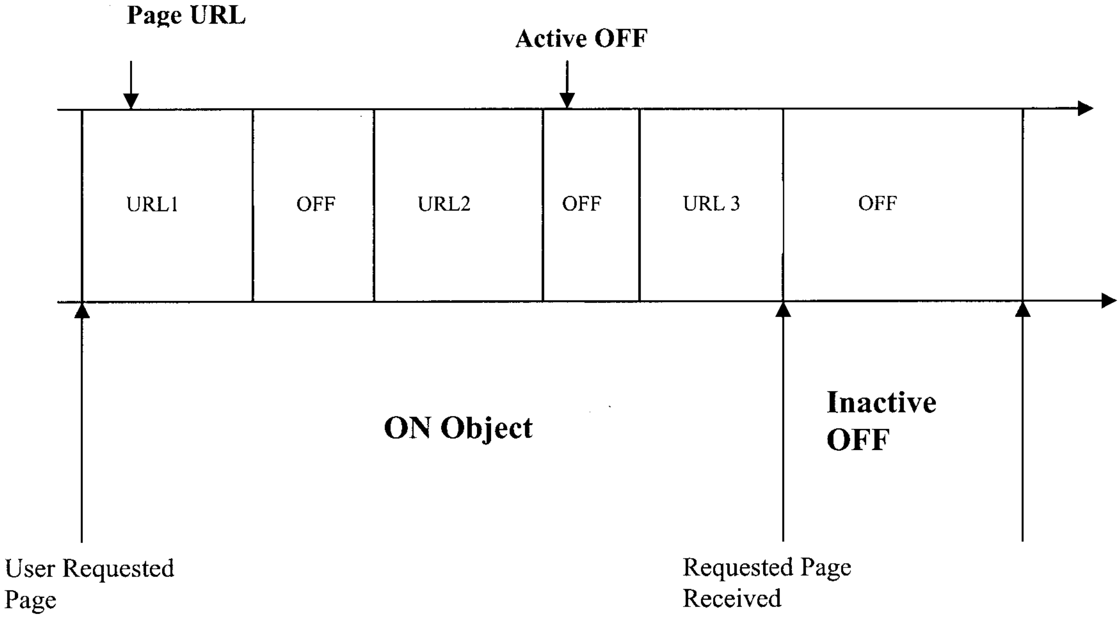

Figure 7-2: ON/OFF WWW traffic model (one TCP connection at a time) [87]

In this model the parameter that needs to be configured in the simulation is the "Number of User Equivalents" (UEs). Each UE is defined as a single process in an endless loop that alternates between making requests for Web files, and lying idle. Each UE has an ON period and an OFF period. The time segments during which files are transferred is identified as ON, and idle time segments as OFF.

For the transfer of Web objects, there are two kinds of OFF time, as shown in Figure 7-2. An inactive OFF time corresponds to the time segment between transfers of Web objects and corresponds to a user's "think time." An active OFF time corresponds to the time segment occurring between transfers of components of a single Web object; it corresponds to the processing time spent by the browser parsing Web files and preparing to start new TCP connections. OFF time is considered to be "Active" if it is less than a threshold time, which we chose to be 1 second based on inspecting the data. File size model follows the distribution as described in [89, 90]. Distribution model is developed as a hybrid consisting of a distribution described in [90] for the body, combined with a Pareto distribution for the upper tail [87].

Voice Traffic

During a connection, Voice follows a pattern of voice activity (talkspurts) and silence periods. The average of silent periods corresponds to 57% to 65% of the total duration of the call.

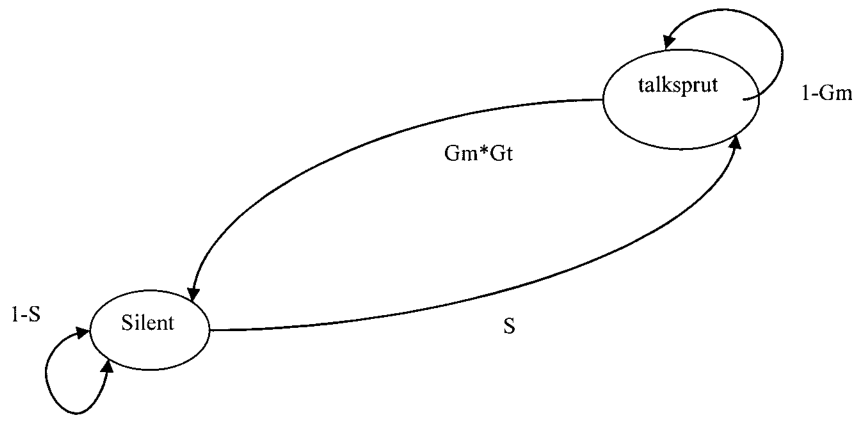

Figure 7-3: State Markovian Model of Fast Speech Activity Detector

When active, the voice users produce traffic (packets) at the rate of 64 KBPS. The average duration of the silence and talkspurt periods have been set to 1.35 sec. and 1 sec. respectively. Since digital voice is highly sensitive to delays, packet should satisfy a certain delay bound in inter packet arrival distance. If a packet violates this limit, it becomes obsolete and the transmitter should discard it in favor of a more recently produced packet. In this work, a voice packet becomes obsolete (and has to be rejected) if it is not accommodated within 32 msec. (which corresponds to 1.5 of the SFP) and the maximum acceptable packet-dropping rate is 1%.

Video Traffic

For video traffic modeling, we use real compressed video sequence traces for our simulations. This is produced with a H.263 video encoder.

In our simulations, we have used H.263 video sequences. The frame rate of the encoder is configured to be 15 frames per second. A video packet becomes obsolete if it is not accommodated with in 66.5 msec. The purpose of using video traffic in our model is to understand as to how well the video traffic would be handled during the mobility of the MN across network.

7.1.2 Mobility Model

The "Mobility" of users is a major advantage of wireless technology over fixed telecommunications systems. The signaling traffic and database processing to support the mobility of users were always key concerns in the design and performance of wireless networks. Mobility models play a key role in studying different mobility-management features such as registration, paging, handoff and database approaches. A mobility model with minimum assumptions and simple to analyze, will be very useful under such circumstances. A mobility model can be applied to represent and simulate the trajectory of moving subscribers in wireless cellular networks. The model can help direct the proper design of networks. It allows us to determine such system parameters as cell dwell time, channel occupancy time and handoff rate needed for performance evaluation (e.g., call dropping probability). In the past, designers have resorted to simulation models to simulate the mobile units either in vehicles or a hand held, carried by pedestrians moving across an area served by a multiplicity of micro-cells.

The 2 types of mobility models used for the analysis of this model are:

- The Random Waypoint
- Random Walk

Random Waypoint

In Random Waypoint Model, each node is assigned an initial location ($x0; y0$), a destination ($x1; y1$), and a speed S. The points ($x0; y0$) and ($x1; y1$) are chosen independently and uniformly within the region the nodes move. In this thesis we have restricted the movement of the MN only in the forward direction. The speed is chosen uniformly from the interval ($v0; v1$), without consideration of the initial location or the destination. After reaching the destination, a new destination and a new speed is chosen uniformly on ($v0;v1$), independently of all previous destinations and speed are chosen following the process described above.

Nodes may pause upon reaching each destination, or they may immediately begin traveling to the next destination without pausing. If they pause, the pause times are chosen independently of speed and location.

Random Walk

In this mobility model, an MN moves from its current location to a new location by randomly choosing a direction and speed in which to travel. The new speed are chosen from pre-defined ranges, [*speedmin*; *speedmax*]. Each movement in the Random Walk Mobility Model occurs in either a constant time interval *t* at the end of which a new direction and speed are calculated. If a MN which moves according to this model reaches a simulation boundary, it "bounces" off the boundary.

Therefore at each time interval *t* the node randomly and uniformly chooses its new direction (which in our case have restricted to move forward). Similarly the new speed $v(t)$ follows a uniform distribution.

7.1.3 Assumptions & Simulation Parameters

For our simulations, we make the following assumptions and choice of parameters:

a. Each MN is made to passes through 6 different base stations and hence 8 handoffs in total for a single simulation (2 inter-domain & 6 intra-domains).

b. In this simulation we have restricted the movement of MN only in the forward direction for both Random walk and Random Way point mobility models. This is mainly to simulate a scenario of a vehicle moving on the road or a highway and of a human being walking on the streets.

c. For each simulation scenario, we make 10 runs, each with different seed and compute confidence interval for each point presented in the graphs. All the intervals provided correspond to a 95% confidence level.

d. In all simulations runs the sources are configured to send a total of 64 Kbps of traffic into the network. This was done in order to study the performance of the proposed model when a large chunk of traffic is being pushed into the network. Especially in scenarios where there could be competing traffic sent along with the actual data source

7.2 Performance Analysis

CBR Traffic

Table 5: CBR Traffic

Source Rate	64 Kb/sec
Packet size	6000 bits
Packet Inter-arrival time	0.1

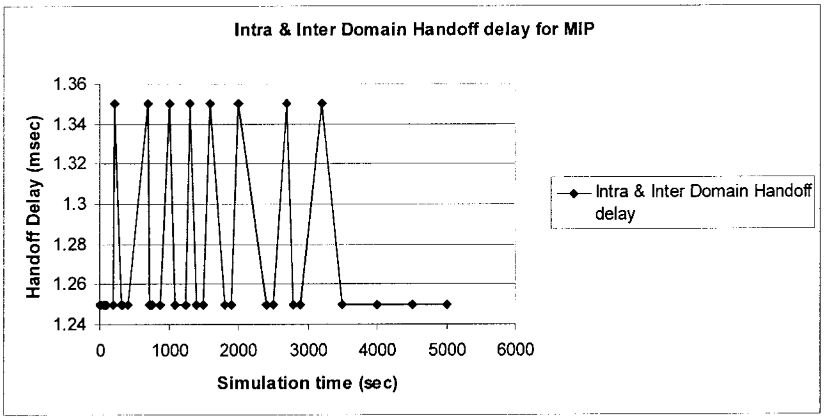

Figure 7-4: Intra & Inter domain handoff delay for MIP

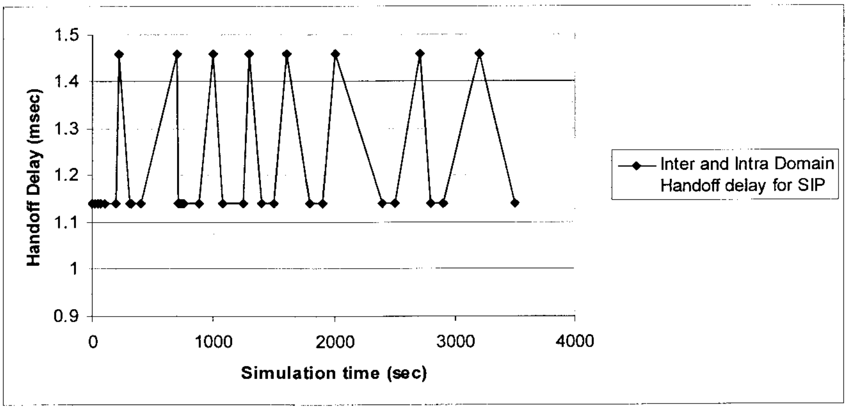

Figure 7-5: Intra & Inter domain handoff for SIP

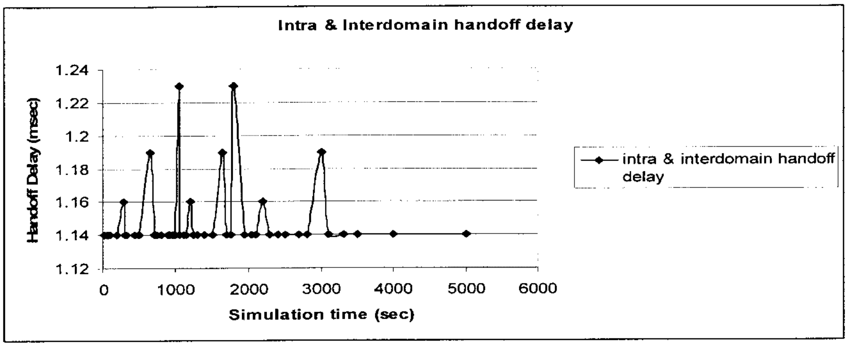

Figure 7-6: Intra & Inter domain Handoff Delay for SIP + MM-MPLS

Figure 7-4; Figure 7-5 and Figure 7-6 show the intra domain and inter-domain handoff delay for Mobile IP, SIP model and the SIP + MM-MPLS that introduced in this dissertation respectively.

The above results show that MIP experiences a delay of 100 msec for inter and intra domain handoff and SIP shows a slightly higher handoff delay with an end to end handoff delay of 320 msec for both intra domain and inter-domain handoffs. This is due to the fact that SIP when compared to MIP has more number of signaling involved for a single hand off, but studies [79, 80] have proved that SIP is a better option for handling inter-domain handoff

In the model proposed in this dissertation we has considered all the major delays introduced in the traditional SIP and MIP model (chapter 5 have explained it in detail). Fig 7-5 shows that the intra domain handoffs is 20 and 40ms respectively and the inter-domain handoff delay is 80.msec.

It is also important to point out that the End-to-End packet delay for MIP is 1.25 msec, becoming 1.14 msec for SIP and SIP over MM-MPLS model. This reduction in End-to End packet delay is due to the fact that SIP and SIP over MM-MPLS completely removes the delay incurred due to the triangular routing in MIP. Hence this is an ideal case when it comes to real time traffic, where mild delays could prove fatal.

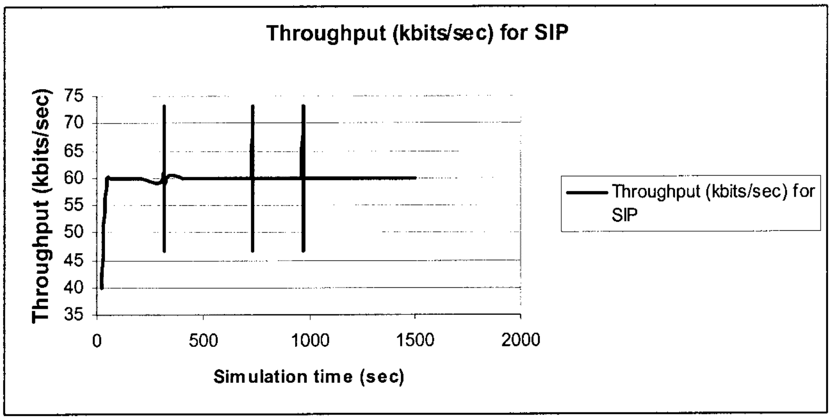

Figure 7-7 Throughput (kbits/sec) for SIP based mobility

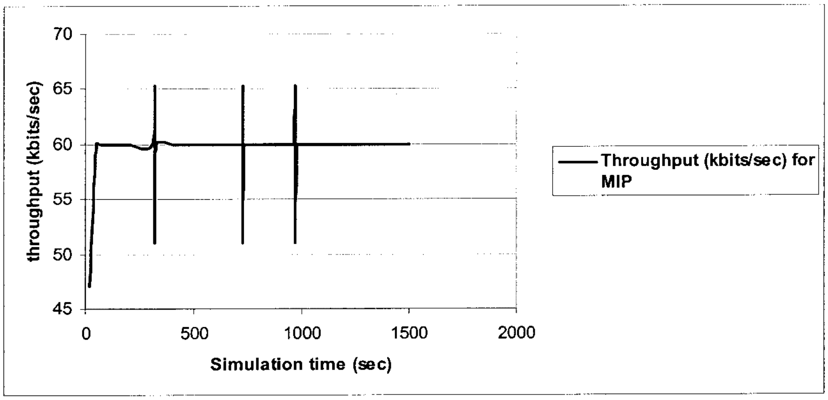

Figure 7-8 Throughput (kbits/sec) for MIP

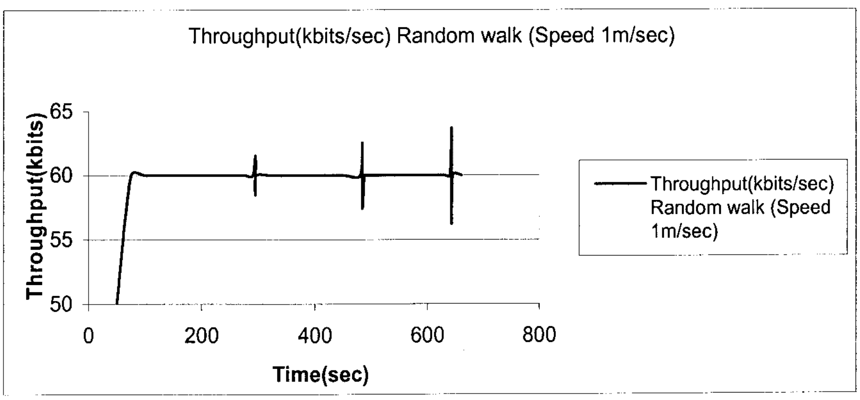

Figure 7-9: Throughput (kbits/sec)

Figure 7-7, Figure 7-8 and Figure 7-9 show the throughput in kilobits / second (Kbits/sec) for SIP, MIP and SIP over MM-MPLS models.

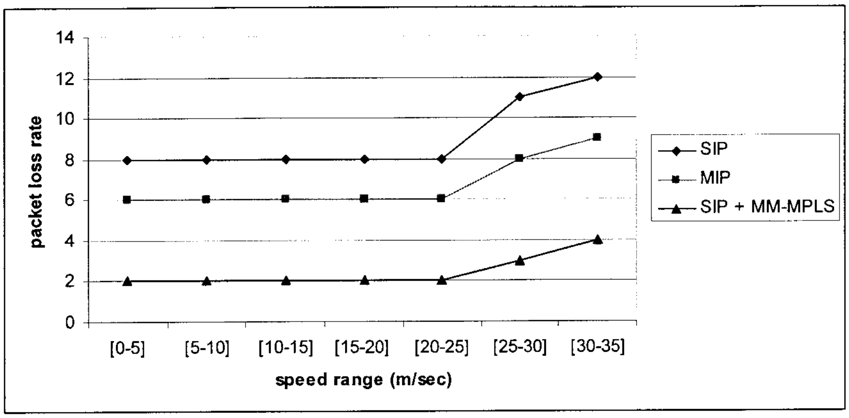

Figure 7-10: Packet loss (number of packets) for buffer size =6

Figure 7-10 shows the packet loss for CBR traffic for different speed of MN when the buffer size of the server (MobServ) is 6. Purpose of this result is to provide a clear comparison on how the MIP, SIP and the SIP over MM-MPLS schemes we are examining behave in terms of packet loss. User mobility follows the mobility model mentioned earlier.

The graphs initially shows a constant behavior up to a speed range of 15-20 m/sec, later on shows a slight linear increase as the speed range increases. As the speed of the MN increases, the number of time the MN has to change the point of contact increases, consequently increasing the handoff rates and hence the number of dropped packets.

When MIP is used the number of packet dropped increases from 7 packets for the speed range between 0-5 (m/sec) to 9 packets for the speed range 25-30 (m/sec).

For SIP the corresponding numbers are 9 and 11 packets as the speed increases. While for SIP over MM-MPLS it is from 2 and 3. It is evident that SIP and MPLS keeps packet loss at considerably lower level.

This result helps us to illustrate the fact that the mobility model in this dissertation considerably reduces the number of packet loss during the handoff when compared to conventional MIP and SIP.

Terminologies

1) **Random Factor "r": "random factor (r)"** is the duration in seconds with which the MN changes its speed.
2) **Pause time "p":** "p" is the duration for which the MN node stops after it moves a certain distance. After the MN pauses for "p" seconds it chooses a random speed from the speed range (1-5, 5-10 etc. and makes the next move.
3) **Speed range**: In the simulation, the speed ranges are categorized into low speed (0-5, 5-10) range, medium speed range (10-15, 15-20) and high speed range (20-25, 25-30).

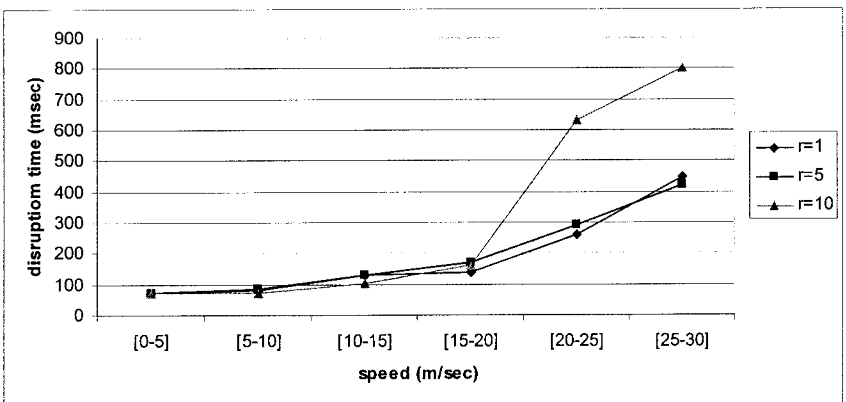

Figure 7-11: Call disruption time for random walk for CBR traffic

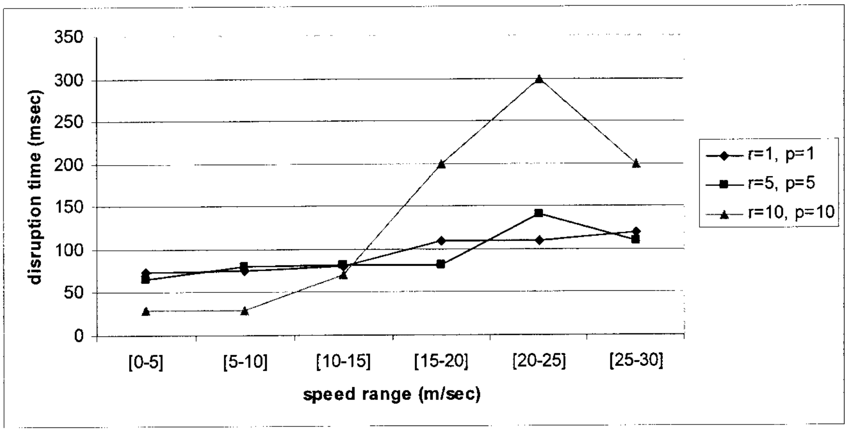

Figure 7-12: Call disruption Time for Random Way Point Mobility

Figure 7-11 and Figure 7-12 shows the end to end delay or the maximum call disruption time plotted for random walk and random waypoint mobility model respectively when CBR traffic is used.

Figure 7-11 shows the call disruption time for random walk for different values of random factor (eg: r=1,5,10) and Figure 7-12 shows the call disruption time for random waypoint mobility for various different combinations of random factor r and pause time p (Eg: r=1/p=1, r=5/p=5, r=10/p=10). This combination was selected to study how the model behave under high level of randomness in mobility (r=1/p=1), as well as under low level of randomness (r=10, p=10)

The simulations were performed for the MN moving at speed ranges ie; (eg: 0-5, 5-10, 10-15, 15-20, 20-25 m/sec). In both figures it shows that the disruption time for CBR traffic has a steady increase with the increase in speed and it happens for all considered random factor "r" values. This is because by increasing the average speed the average number of handoff occurring increases which in turn increases the disruption time.

When it comes to random walk mobility model (Fig 7-12) we see a slight deviation of the graph from its linear nature. This is due to the fact that in random walk there is another factor pause time "p". If the MN is in the out of coverage area when it pauses then it will have to wait till the next movement it is due to move.

The results prove that our model perform well under both, random walk and random waypoint mobility models for CBR traffic with a slight improvement in results for random waypoint when compared to random walk mobility model. Results show a slightly better behavior for less random mobility model. In the case of Random walk mobility model which has more random mobility behavior, the disruption time stays below 200msec until it reaches a speed in the range of 20-25 m/sec and then shows a spike. For Random walk mobility model the call disruption time remain bellow 120 msec, except some exceptional range.

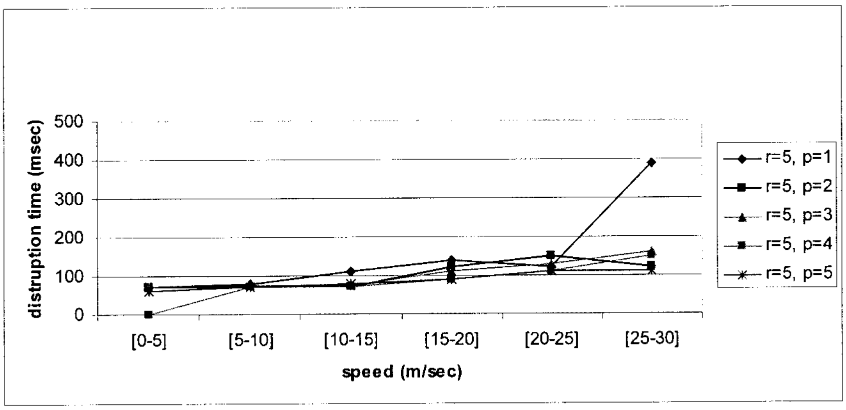

Figure 7-13: Call disruption time for Random Way Point Mobility Model for r=5 and Different Pause Time,for CBR traffic

Figure 7-13 shows the behavior of disruption time when the MN moves according to the random waypoint model and the value of the random factor is kept at 5. The figure indicates the impact pause time and for different speed range have on the disruption time.

The increase in pause time means the randomness degree in the MN's movement is moving decreases. From the above graph it can be concluded that call disruption time remains almost steady throughout and the delay remain under 150 msec, the only exception being for r=5, p=1. This is because, for a constant random factor r, as the pause time decreases, the randomness by which the MN moves increases which in turn increases, the average number of required handoff. This causes the slight increase in the call disruption.

Note: This behavior of call disruption and packet loss being directly proportional to the speed/the randomness with which the MN moves, is evident only for CBR traffic. In the upcoming section we will examine what is happening for other types of traffic (Pareto, voice, video)

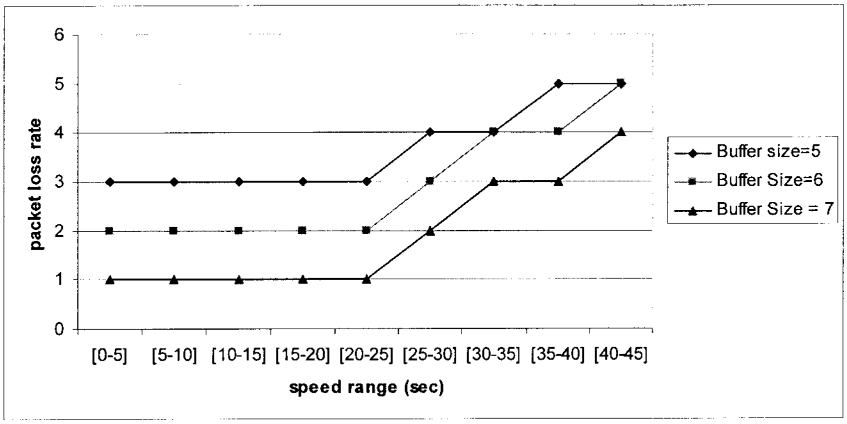

Figure 7-14: Packet loss for Random Walk Mobility using r=1

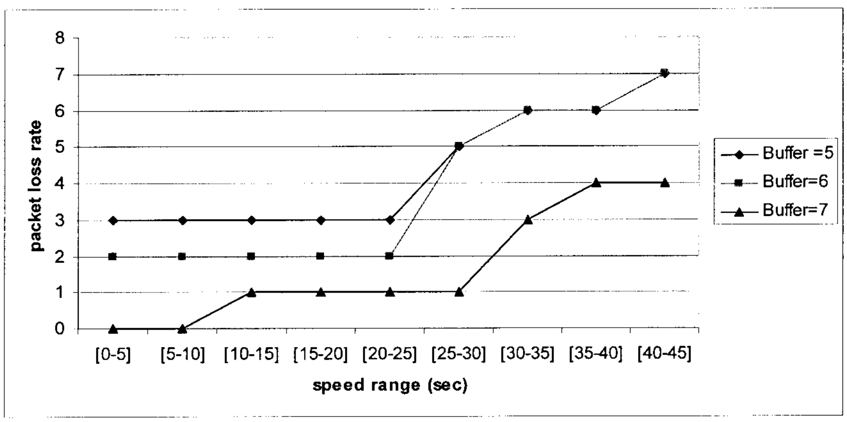

Figure 7-15: Packet loss for Random Walk Mobility, using r =5

Figures 7-14 and 7-15 show the rate of packet loss experienced by CBR traffic for under random walk mobility model with random factor r = 1 and 5 respectively, and buffer size at MobServ 5, 6 and 7.

In both cases, the results show a sudden increase in the packet loss when the speed reaches high value. This is due to the fact that, as the speed increases, the average number of times the MN has to change its point of contact increases the average number of handoffs to be performed. This causes an increase in packet loss. The results shows a maximum packet loss of 7 (for buffer size=5,6). This proves that SIP over MM-MPLS model handles mobility for CBR traffic up to a fairly good speed in the range of 25-30 m/sec. when the traffic is of CBR nature.

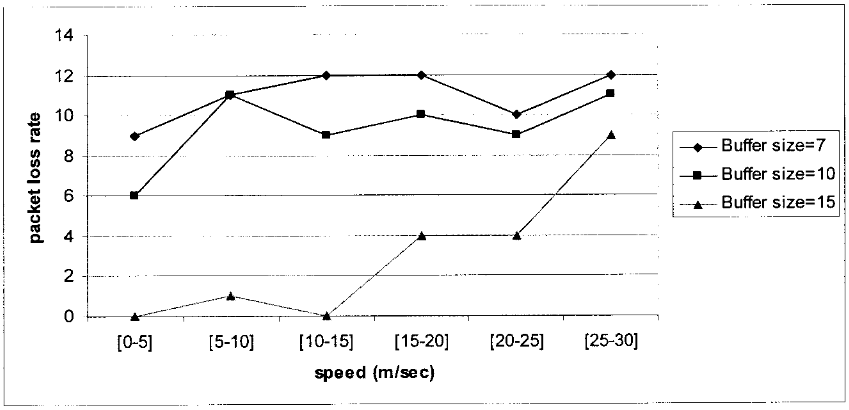

Figure 7-16 Packet loss for Pareto distribution where the random walk model with r=1 is used

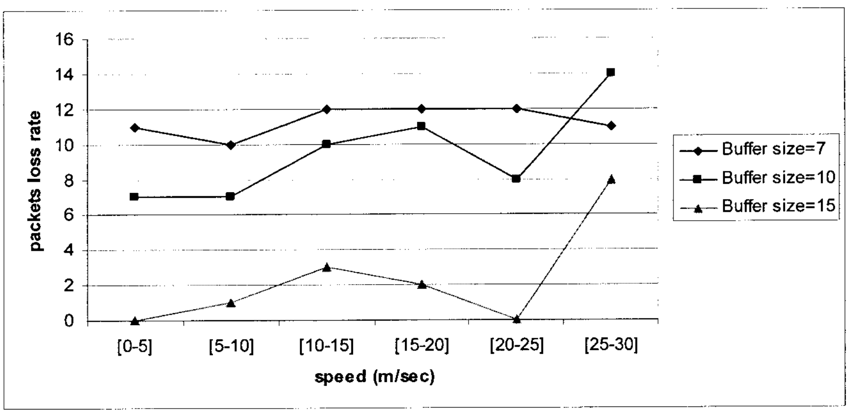

Figure 7-17: Packet loss for Pareto distributed when the random walk model is used (r=5)

Figure 7-16 and Figure 7-17 illustrate the packet loss for data traffic generated according to Pareto distribution, and for random factor r=1 and r=5 respectively. Unlike the graph corresponding to CBR traffic, those for Pareto traffic do not show a linear behavior for packet loss graph. This is because of the difference in traffic behavior. Pareto distribution has ON and OFF states, with packets being sent when it is in the ON state and no packets being generated during the OFF state.

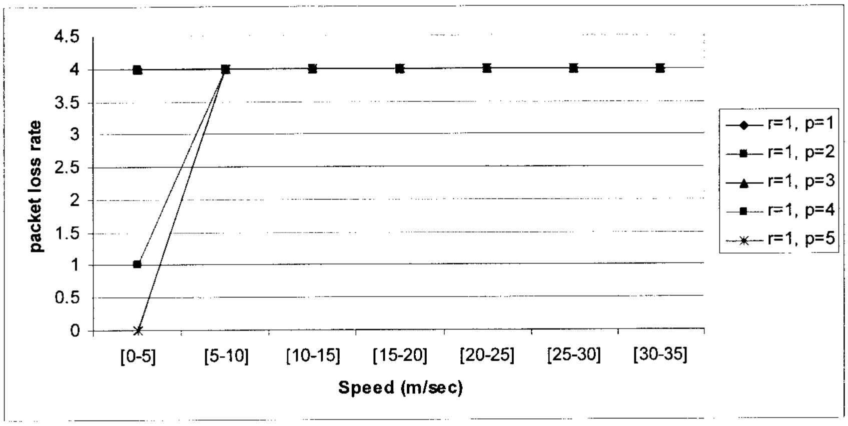

Figure 7-18: Packet loss for Random Way Point mobility Model, where r=1,buffer size=5, for CBR traffic

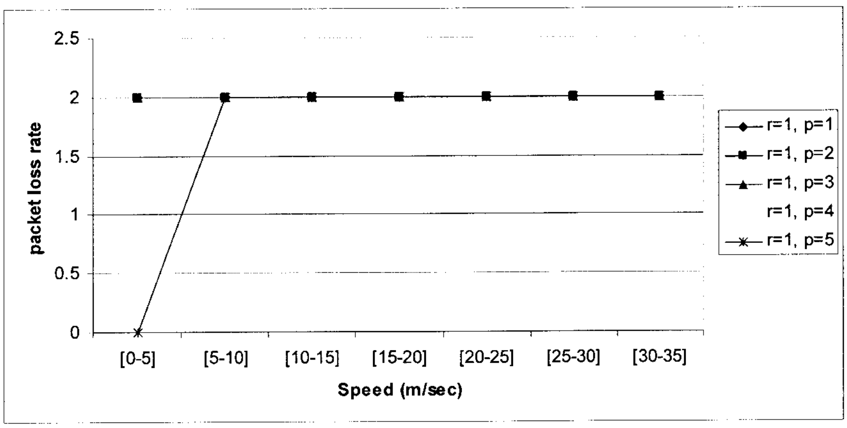

Figure 7-19: Packet loss for Random Way Point mobility Model, where r=1,buffer size=6, for CBR traffic

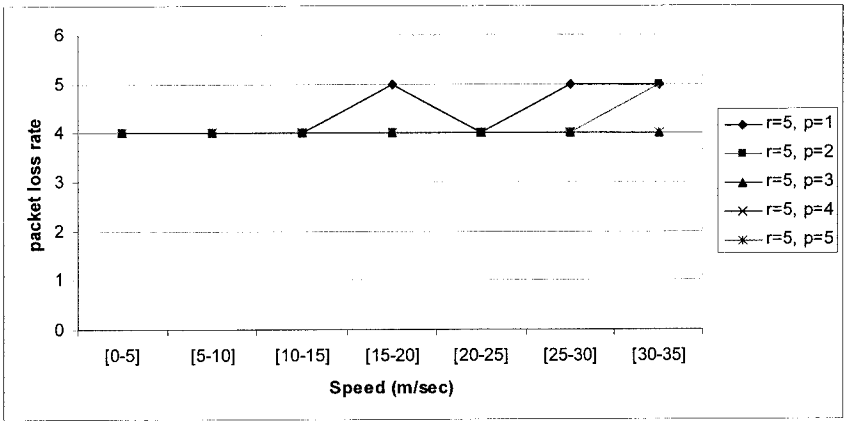

Figure 7-20: Packet loss for Random Way Point mobility Model, where r=5,buffer size=5, for CBR traffic

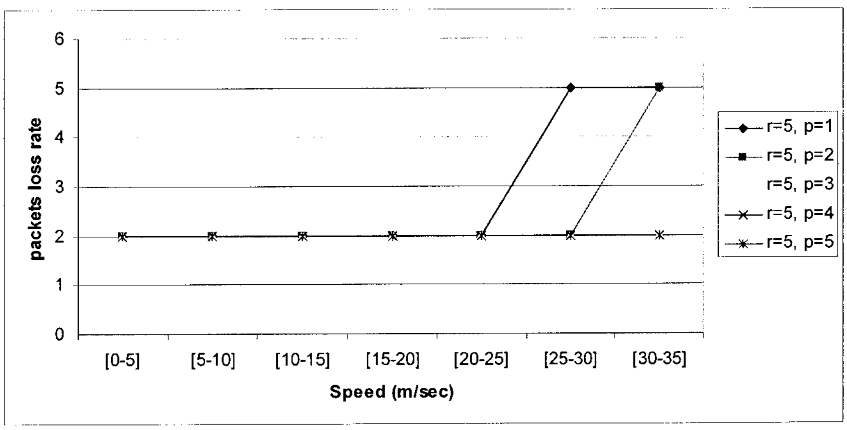

Figure 7-21: Packet loss for Random Way Point mobility Model, where r=5,buffer size=6, for CBR traffic

Figures 7-18, 7-18, 7-20 and 7-21 present the number of packets being lost where the random waypoint mobility model is used and CBR traffic is sent by the source. For Figures 7-18 and 7-19 the random factor r=1. While for figures 7-20 and 7-21, r=5.

From the Figures 7-18 and 7-19 it is evident that when the randomness of the MN move is low (in Figure 7-18 the pause time p=4 and 5, in Figure 7-19 when the pause time p=5) the number of packets being dropped decreases. From these figures we are able to conclude that as the pause time increases (resulting to a lower factor of time the MN moves) the packet loss decreases.

Figure 7-20 and 7-21 shows a slight increase in the packet loss for some higher values of the speed range.

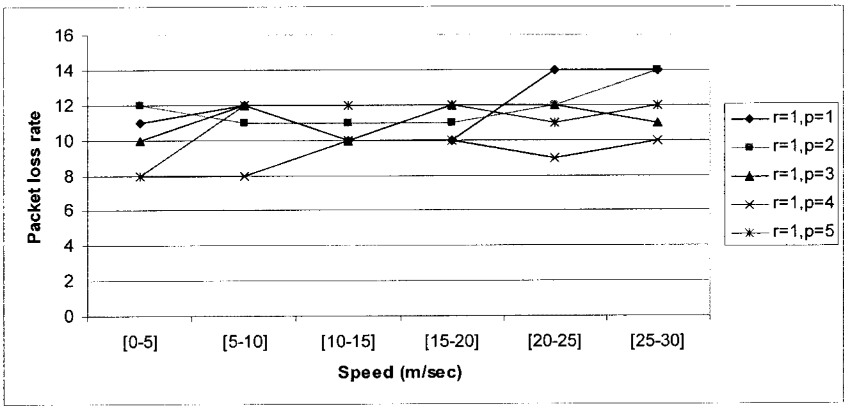

Figure 7-22: Packet loss for Random Way Point mobility Model, for r=1,buffer size=7, for Pareto traffic

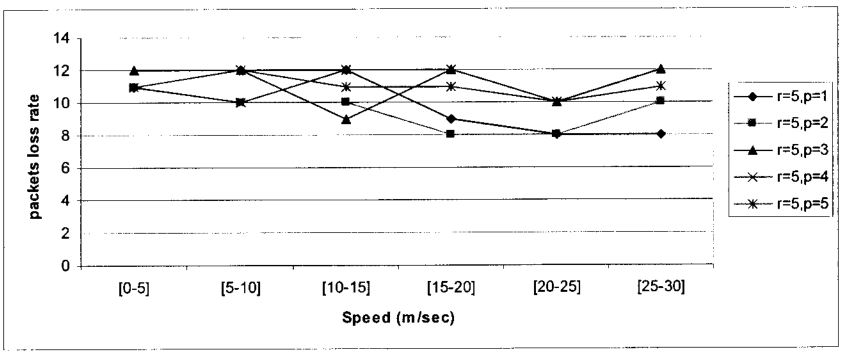

Figure 7-23: Packet loss for Random Way Point mobility Model, for r=5, buffer size=7, for Pareto traffic

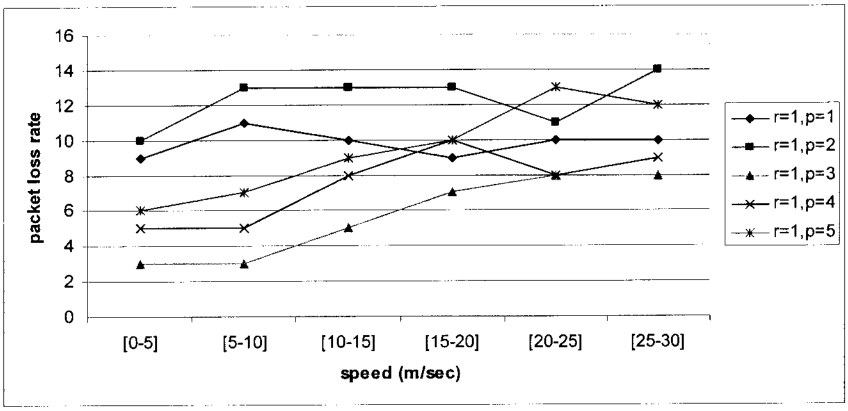

Figure 7-24 Packet loss for Random Way Point mobility Model, for r=1, buffer size=10, for Pareto traffic

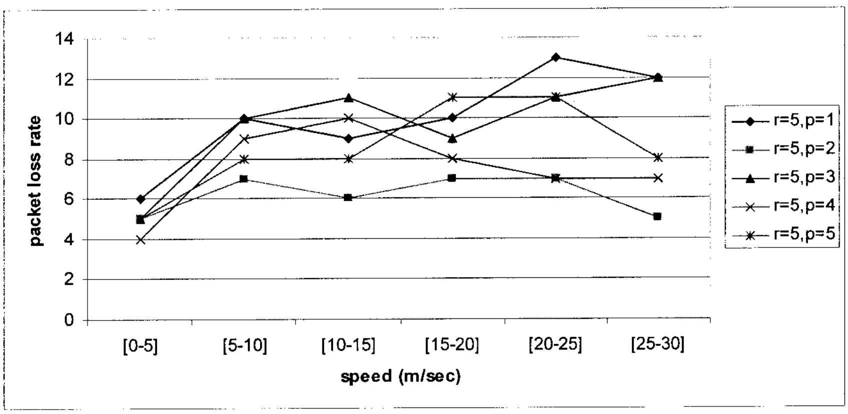

Figure 7-25: Packet loss for Random Way Point mobility Model, for r=5, buffer size=10, for Pareto traffic

Figure 7-22, 7-23, 7-24, 7-25 show the packet loss for Pareto distribution data traffic when the MN follows the random waypoint mobility model. Figure's 7-22, 7-23 show the packet loss for buffer size = 7, with random factor r=1, r=5 respectively, and Figure 7-24, 7-25 shows the packet loss for buffer size =10 for random factor r=1, r=5 respectively.

When we compare the CBR and Pareto traffic distribution's, it is noticeable that Unlike CBR Pareto traffic distribution it does not show a very linear behavior for the speed range we have performed our simulation (packet loss increases with the increase in the speed of the MN). This no-linear nature of graph is due to the ON-OFF nature of the Pareto distribution. Packets and send only during the ON period. If the handoff occurs during an

OFF period the packet loss will be zero or atleast lower compared to the case where the source sending packets during the entire duration of handoff. In case of CBR traffic packets are always send through out the simulation period, thus the probability of packets being dropped remains constant through out.

Our model supports the mobility and keeps the average packets loss for internet traffic (Pareto traffic) below 15 for the speed range for both random walk and random way point.

WWW Web Traffic

In this simulation Environment we use a Single User Equivalents, with traffic rate of 32 Kbps

Number of User Equivalents	1
Traffic rate	32 Kbps
Number of Streams	1

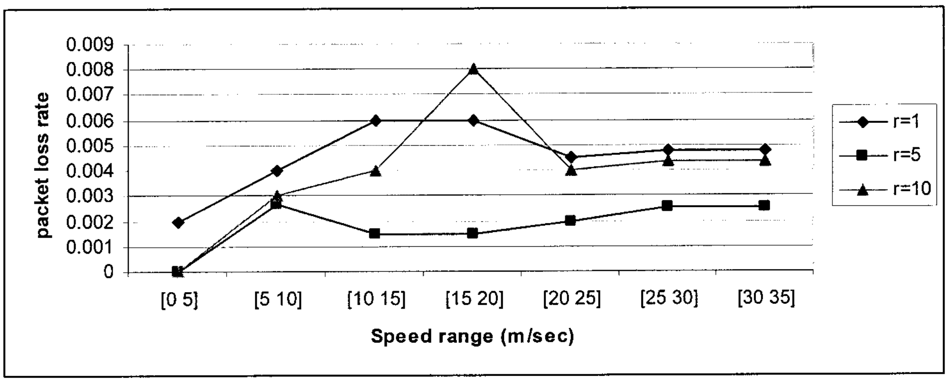

Figure 7-26: Packet loss for Random Walk mobility Model, buffer size=5, for WWW traffic

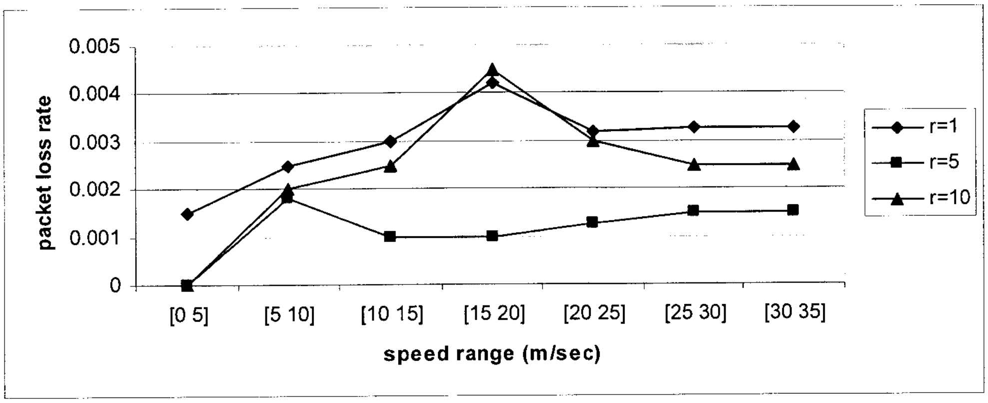

Figure 7-27: Packet loss for Random Walk mobility Model, with buffer size=7, for WWW traffic

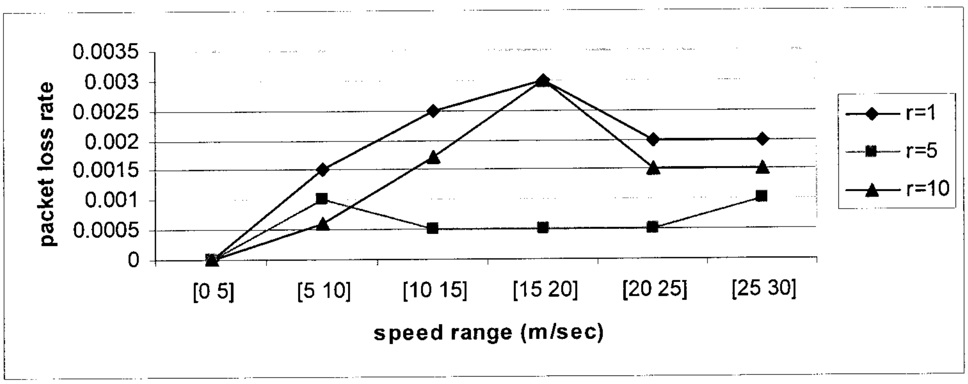

Figure 7-28: Packet loss for Random Walk mobility Model, with buffer size=10, for WWW traffic

Figure 7-26, 7-27, 7-28 present the rate of packet loss for WWW traffic, single UE, random walk mobility model(r=1, 5, 10) and buffer sizes of 5,7 & 10 respectively. From the graphs we can draw the following observations: 1) the maximum packet loss rate for random factor r=1 is obviously higher when the buffer size is 5, which is 0.8% , and reduces to 0.45 % when the buffer size is increases to by 7 and to 0.3 % for buffer size 10. 3) The packet loss is considerably users with lower mobility . This can be concluded by comparing the results corresponding to the r=7 with those of r=10 and r=10. 3) As we explained before, in the case of Pareto traffic, www traffic does not show linear behavior in terms of packet loss, i.e. the speed range does not have any direct proportionality to the packet loss. This non-linear nature of the graph is due to the bursty nature of the web traffic. 4) From these graphs we could see that our model shows pretty good behavior for speeds as high as 30-35m/sec and of course the maximum packet loss, it being 0.8%, occurs when the buffer size is 5.

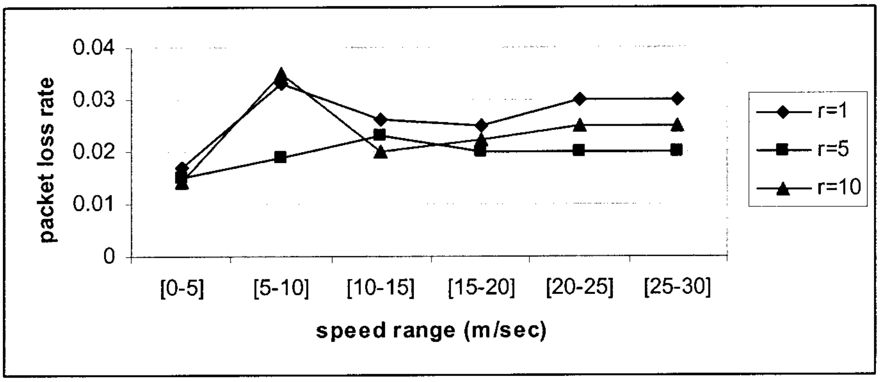

Figure 7-29: Packet loss for Random Walk mobility Model, with buffer size=10, for WWW traffic

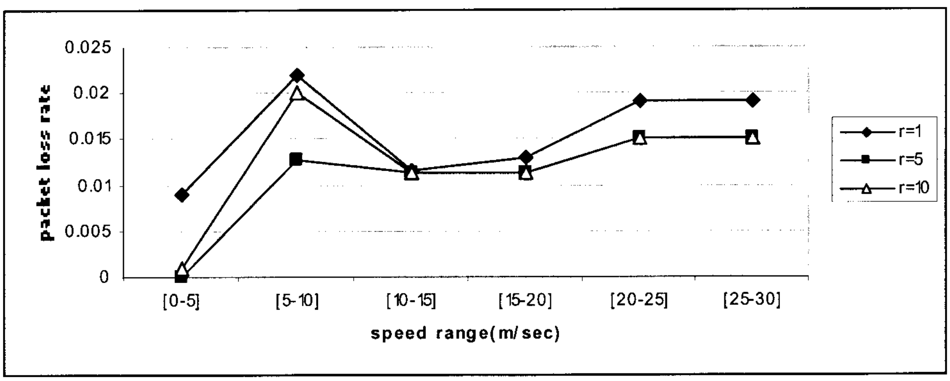

Figure 7-30: Packet loss for Random Walk mobility Model, with buffer size=15, for WWW traffic

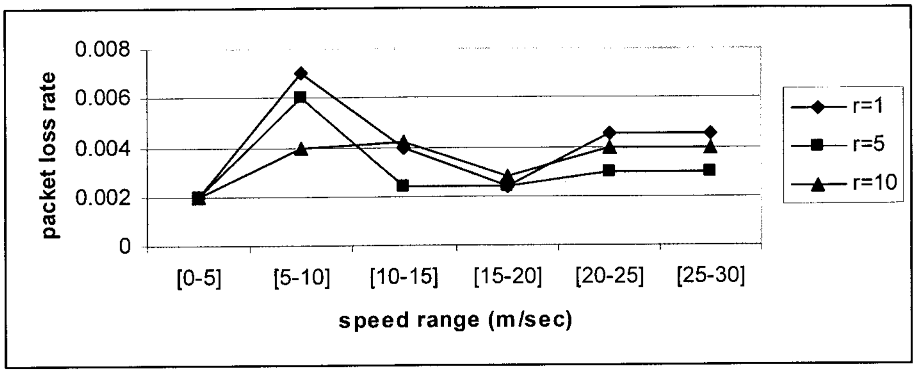

Figure 7-31: Packet loss for Random Walk mobility Model, for with buffer size=20, for WWW traffic

Figure 7-29, 7-30, 7-31 present the rate of packet loss for a WWW traffic random walk mobility model (r=1, 5, 10) and buffer size being set to 10,15 & 20 respectively. In this case the total traffic rate is increased to 64 Kbps, and the number of traffic sources is increased to 3, with 2 of the sources sending traffic at the rate of 16kbps, and the actual correspondent node sending traffic at the rate of 32 Kbps. This scenario was simulated to study the performance of the model, when other competing traffic or noise gets introduced along with the actual traffic into the network. It might also be noted the buffer size had to be increased in order to generate realistic output.

Comments that can be made by observbing the produced graphs are the following: 1) The nature of the graphs remains the same with those generated for a single source, with traffic rate of 32 Kbps. 2) There is a slight increase in the packet loss, with the packet loss curve getting elevated to a maximum of 3.1% for r=1 and speed range of 5-10 m/sec

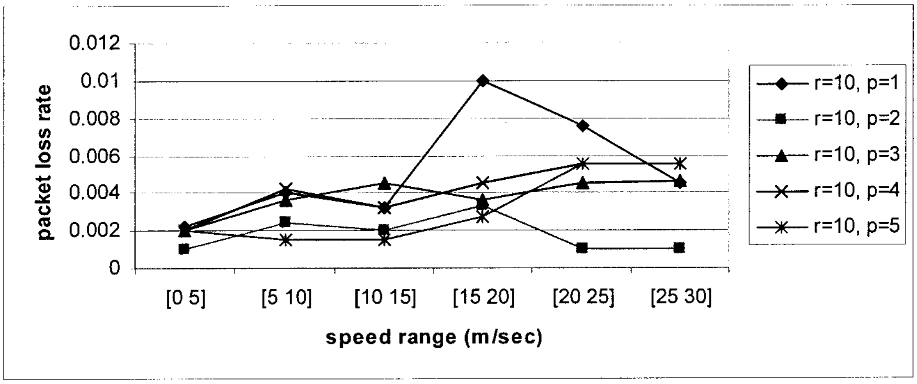

Figure 7-32: Packet loss for Random Way Point mobility Model, for r=10, with buffer size=5, for WWW traffic

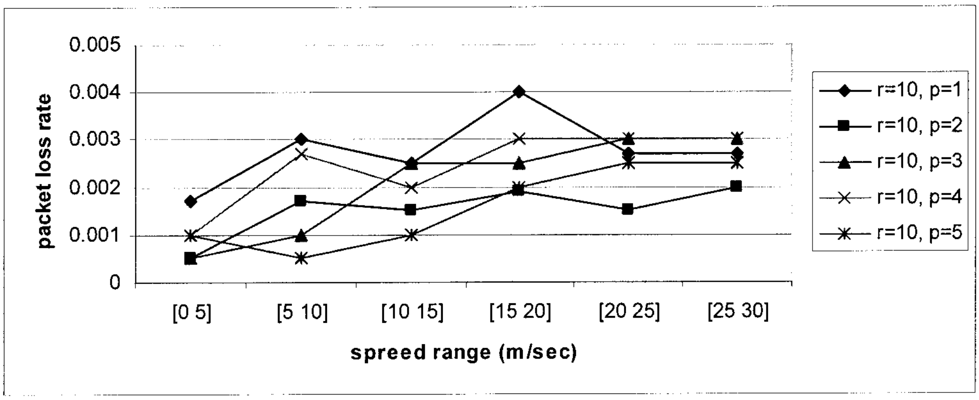

Figure 7-33: Packet loss for Random Way Point mobility Model, for r=10, with buffer size=7, for WWW traffic

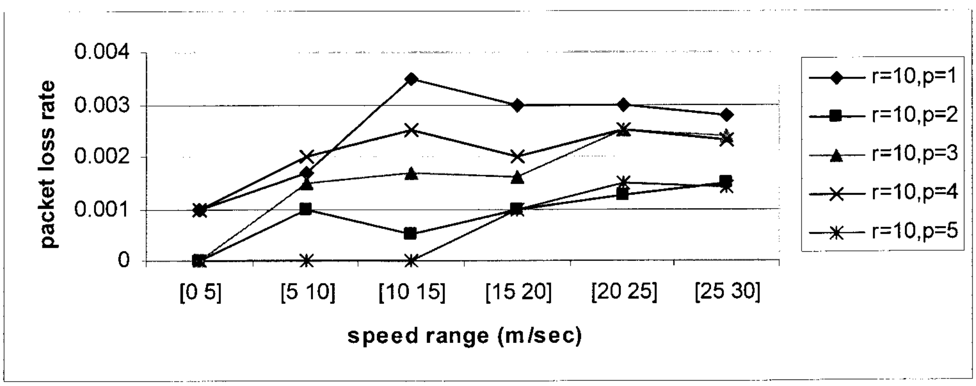

Figure 7-34: Packet loss for Random Way Point mobility Model, for r=10, with buffer size=10, for WWW traffic

Figures 7-32, 7-33, 7-34 display the results for the packet loss rate for WWW traffic and single UE, when the random way point mobility model with for random factor 10 is applied. A set of different pause times (p=1, 2, 3, 4, 5) and buffer size (5, 7, and 10) are used as parameters. The packet loss rate goes up to a maximum of 1 % when the buffer size is 5, and is reduced to 0.4 % when buffer size is increased to 7, as well as drops down to 0.35 % for buffer size 10. From the above results we could see that the1) packet loss rate WWW traffic is not linear as the nature of the traffic is not constant. 2) All the 3 graphs show higher packet loss for highly random moving users i.e. r=10, p=1 3). The traffic is to be best handled by a less randomly moving user, eg, a user moving with r=10 p=5 mobility factors.

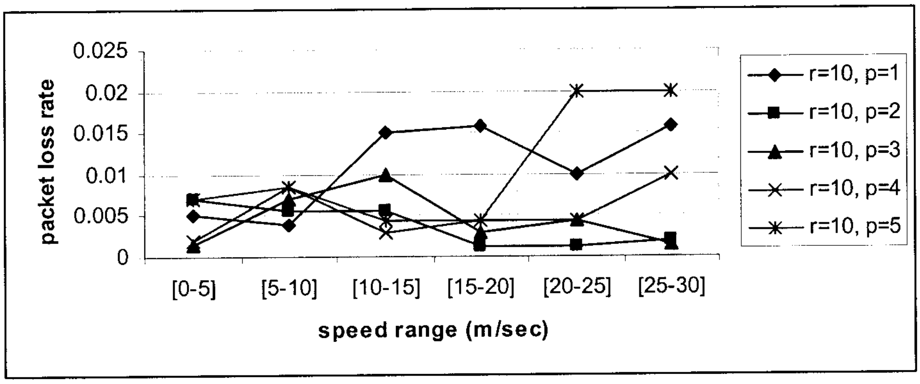

Figure 7-35: Packet loss for Random Way Point mobility Model, for r=10, with buffer size=15, for WWW traffic

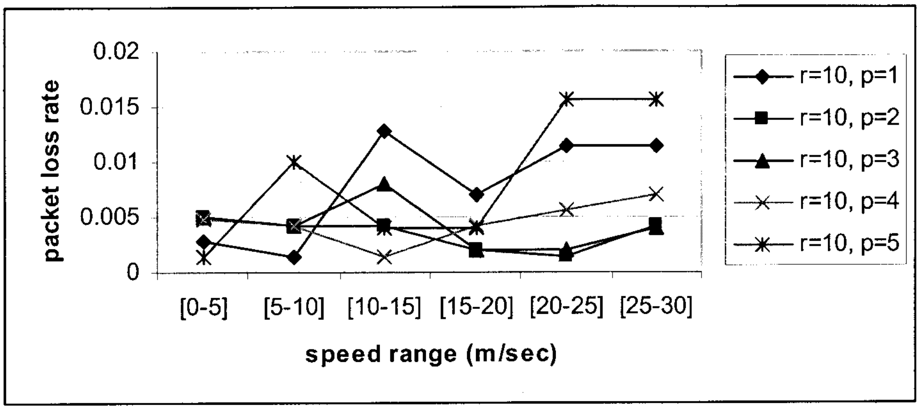

Figure 7-36: Packet loss for Random Way Point mobility Model, for r=10, with buffer size=20, for WWW traffic

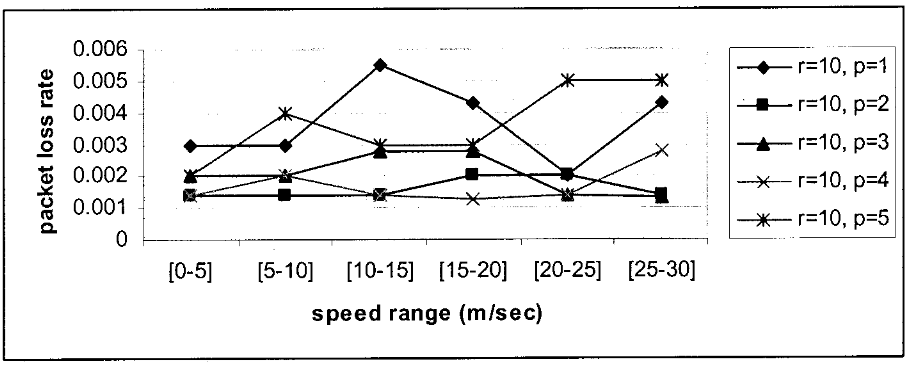

Figure 7-37: Packet loss for Random Way Point mobility Model, for r=10, with buffer size=25, for WWW traffic

Figures 7-35, 7-36, 7-37 present the rate of packet loss for a WWW traffic random waypoint mobility model(r=1, 5, 10) and buffer size being set to 10,15 & 20 respectively. The simulation scenario is similar to the one explained for figures 7-29, 30, 31, with number of traffic source increased to a total of 3, 2 sources sending packets at 16 Kbps and the correspondent node sending packets at 32Kbps making a total of 64 Kbps traffic in the network. the simulation parameter, buffer size had to be increased to 15, 20, and 25 respectively. It should be noted that as the amount of sent traffic increases, the simulation parameter, buffer size had to be increased in order to generate realistic output. The nature of the graph remains the same with the one corresponding to a single source and traffic rate of 32 Kbps, but could observe a

Voice Traffic

Table 6: Voice Traffic Results

Source Rate	64 kb/sec
Packet size	80bits
Principal Talkspurt	1.000 secs
Principal Gap	1.350 secs
Packet Inter-arrival time	0.00125

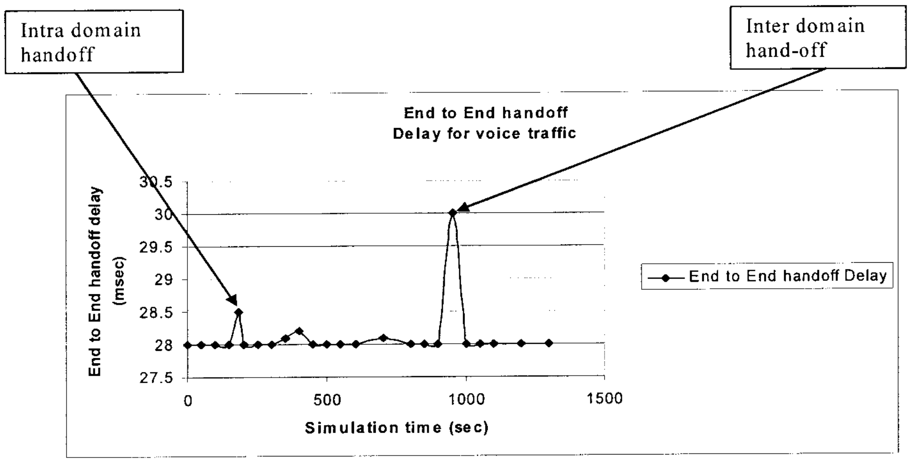

Figure 7-38: End to End Handoff Delay for Voice Traffic

Figure 7-38 shows the end to end intra domain and inter domain hand off delay for the voice traffic when the MN moves at a speed of 1m/sec. The intra-domain handoff delay comes out to be 0.5msec and the inter-domain handoff for voice packet to be 2msec.

In should be noted that for voice traffic the acceptable packet delay is a maximum of 32msec. as voice packet with higher delay is considered to be obsolete. And the acceptable % of packet loss is 1%. For voice traffic anything higher that 1% of packet loss is considered to be a poor performance.

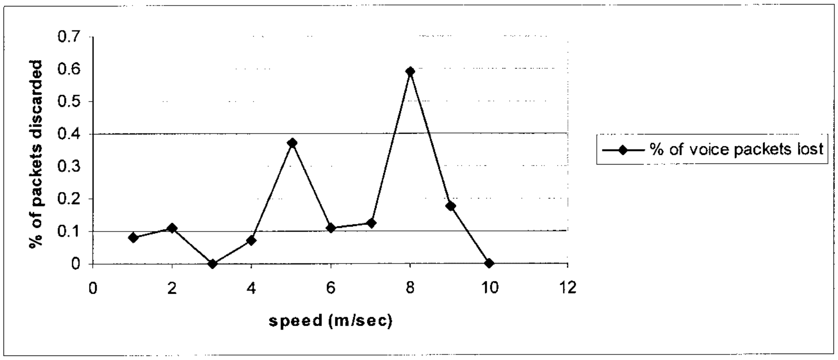

Figure 7-39: % of Voice Packets Dropped

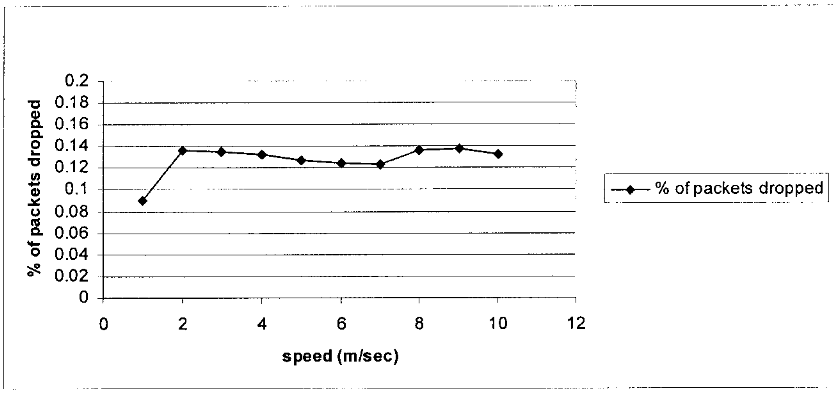

Figure 7-40: % of Voice Packets Dropped (for multiple simulations with different seed value)

Figure 7-39 and Figure 7-40 shows the percentage of packet loss experienced by voice traffic for various speed ranges, when the MN moves in a random walk fashion.

Figure 7-39 shows the result of a single simulation run (with a single seed value). The result demonstrates that the % of packet loss, does not increase at a constant rate as the speed increases, but it tends to show a bit irregular behavior unlike the CBR traffic. This zig-zag nature of the graphs shows that there is no linear nature to the packet loss for voice traffic while the speed of the MN increases (for the speed limit with which the simulations are performed), which would have been what ideally was expected.

This behavior is due to the nature of voice traffic, where the packets are not sent at a constant rate during voice activity (ON period) but instead has a talk-silent period (ON-

OFF) model. In our simulation we are assuming that during talk time the packets are sent at the rate of 64 Kbps.

As it can be seen the packet loss rate of voice packets (due to experimentation) appears to be independent or at the most having very small dependency with the speed. To our understanding this is due to the fact that the distance between consecutive base stations and the speeds considered produce a long duration between 2 consecutive handoffs. For example for 600 meter distance between 2 base stations and under the extremely unlikely case that the mobile will always remain on the straight line linking the BS, it will take atleast 60 seconds for a 10m/sec MN. This duration makes almost certain that the MN will be passing through 2 consecutive BS, while not making transition from one state to the other. Thus the state the MN is, when entering the handoff process of a BS is statistically independent from what was the state it was at the previous or will be at the next state.

In Figure 7-40 the packet loss slightly reduces for medium range speed (5 m/sec, 6 m/sec, 7m/sec), and again slightly increases for higher speeds like 8, 9, 10 *M/sec)

Figure 7-40 depicts that the SIP over MM-MPLS model shows very good results for voice traffic, by keeping the packet loss rate lower than 0.14%.

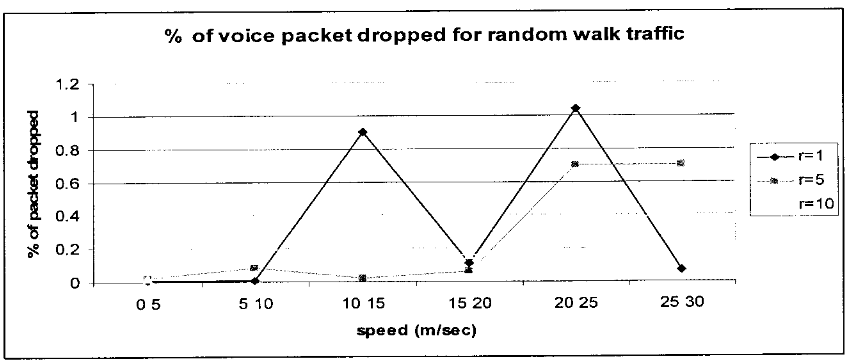

Figure 7-41: % of Voice Packets Dropped when Random Walk Mobility is used (single simulation run)

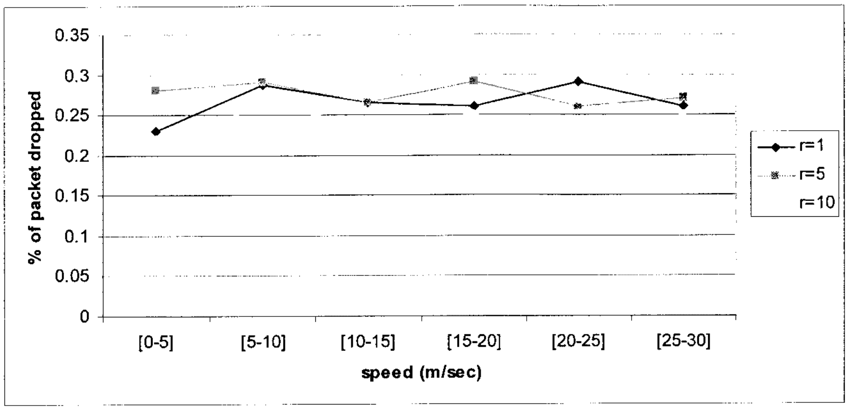

Figure 7-42: % of Voice Packets Dropped when Random Walk Mobility is used (average of multiple simulations run)

Figure 7-41 and Figure 7-42 shows the percentage of packet dropped for random walk mobility model for voice traffic for random factor r= 1, 5, 10.

Figure 7-41 shows the result of a single simulation run (with a single seed value). Figure 7-42 shows the same results for multiple simulation runs with different seed values.

The results demonstrate that the % of packet loss, does not increase at a constant rate as the speed increases, but it tends to show a bit non-linear nature as the speed increases. This zig-zag nature of the graphs shows that there is no linear nature to the packet loss for voice traffic while the speed of the MN increases. This nature could be explained due to the ON-OFF nature of the voice traffic.

We could also notice that, for smaller values of r ie; when r =1, r=5 it shows slightly higher percentage of packet loss when compared to lower value of random factor r. This is due to the fact that smaller value of 'r' means the random nature of the MN is high.

The model presented in this dissertation shows fairly good results for percentage of packet loss for voice traffic for MN node moving in random walk fashion. The percentage of packet loss remains below 0.3% for speeds upto speed of 25-30m/sec.

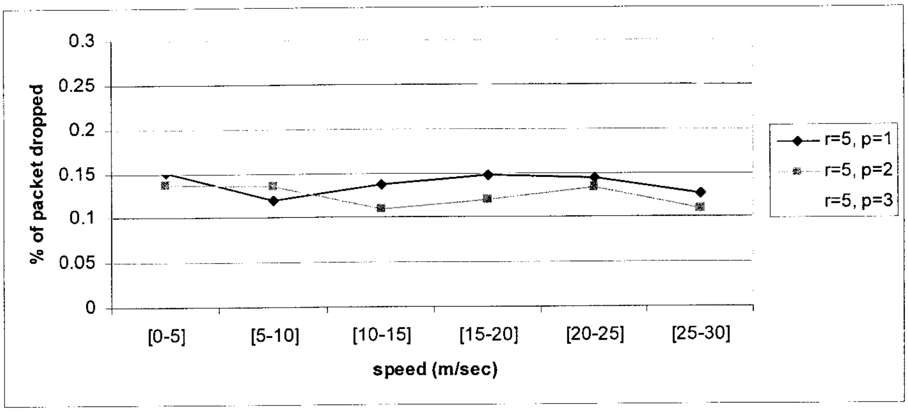

Figure 7-43: % of Voice Packets Dropped when the Random Way Point Mobility model is used

Figure 7-43 shows the % of packet loss for voice traffic when the MN follows a random way point mobility model. The random factor (time after which the speed of the MN changes) is 5 seconds and the displayed curves correspond to pause times (p=1, p=2, p=3).

The rate of packet loss remains steady with slight variation at different speed for pause time 1 sec and 2 sec. The packet loss shows a slightly higher value when the pause time increases, because the longer pause time will keep the MN in out of coverage are for longer this in turn increases % of packet loss.

Comparing the results for random walk and random waypoint mobility model (Figure 7-42 and 7-43), SIP over MM-MPLS model shows a slight better behavior with the random waypoint mobility model.

Video Traffic

Table 7: Video Traffic

Source Rate	64 kb/sec
Packet size	VBR (H263 codec)
Number of frames	15 frames/sec

Note: Itshould be noted that for voice traffic the acceptable packet delay is a maximum of 66.5 msec. as voice packet with higher delay is considered to be obsolete.

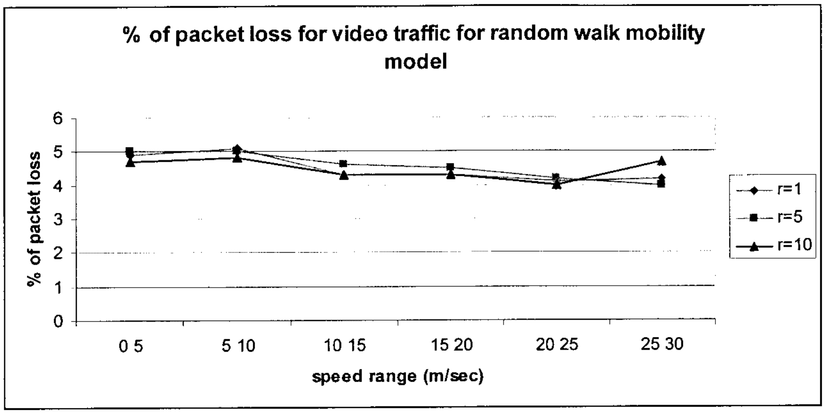

Figure 7-44: % of Video Packets Dropped for Random Walk Mobility

Figure 7-44 shows the % of video packets dropped during the mobility of the MN in random walk fashion. Unlike the voice traffic the video traffic shows a considerably steady behavior. There is not much difference in the % of lost packets for different speed ranges or different random factor. This is due to the nature of the video traffic.

With our model, the maximum packet loss remains under 5% for up to a speed of 25-30 m/sec.

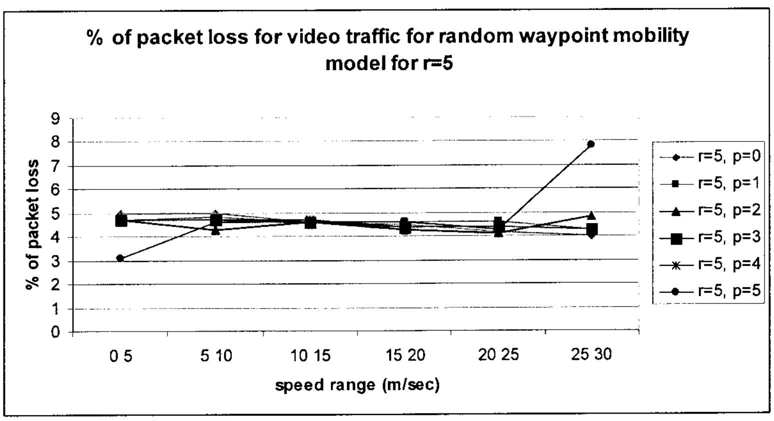

Figure 7-45: % of Video Packets Dropped for Random way point Mobility

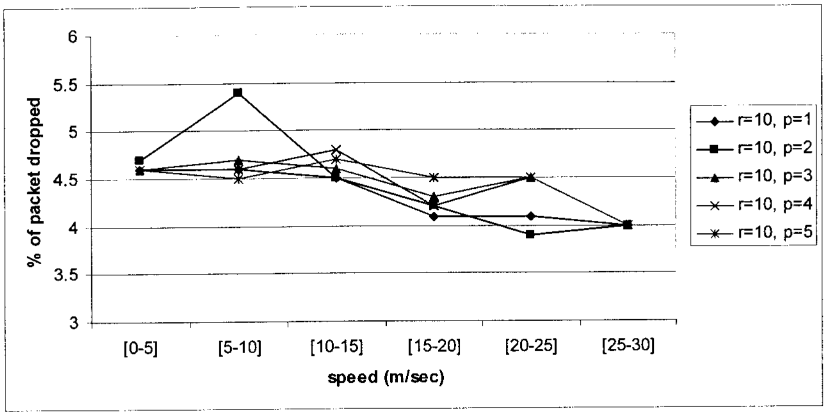

Figure 7-46: % of Video Packets Dropped for Random Way Point Mobility model for video traffic

Figure 7-45 and 7-46 shows the simulation results for the % of packets dropped for video traffic when the MN is moving in random waypoint mobility model, for random factor r=5 and r=10 respectively. Comparing the figures, it could be noted that the % of packet loss is more or less steady. But in both the figures it could be noted that there is a slight increase in the packet loss when the randomness with which the MN moves is high (ie for lower values of pause time eg: p=2.)

CHAPTER 8
CONCLUSIONS AND FUTURE WORK

This dissertation has introduced a method to facilitate transparent roaming between independent and heterogeneous wireless network access technologies. Simulation evaluations demonstrate that the proposed technique enables roaming in integrated networks, reduces handoff delay, and retains transparency by hiding complexities from users. This chapter of the thesis summarizes the contributions of the research and describes potential avenues for further research.

8.1 Summary & Conclusion

The main objective of this research was to develop a model to handle seamless mobility across converged networks. After the research work as outlined in Chapter 2, it was observed that supporting mobility in the converged networks poses challenging issues viz. how

- To minimize mobility disruptions while roaming, by handling both intra-domain as well as inter-domain mobility and
- To reduce the packet loss.

To address these issues, a suite of integrated Mobile IP and SIP model is proposed. An extension of Lower layer MIP protocol (MM-MPLS) is proposed to handle the intra domain mobility and the Application Layer SIP protocol is proposed to handle the inter-domain mobility. Research has proven that intra-domain mobility management is best handled by lower layer protocols like HMIP and MM-MPLS. The benefit of using SIP for (application layer) mobility management is that it allows applications to adapt their service behavior based on the mobility management strategy selected, in order to provide the best possible end user experience.

The hypothesis was thoroughly evaluated by simulating the proposed model with OPNET. Results show that it is possible to minimize the handoff delay and hence minimize packet loss for real time applications during heterogeneous network handoffs, enabling transparent mobility. The simulation results also show that the proposed approach outperforms the existing approach in most cases. It is believed that the work presented here is an important step towards supporting VoIP service over wireless mobile environment.

8.2 Future work

This work identifies the following problem areas for further study and research.

- QoS, is an open issue. It is useful to undertake a research on how to add QoS support to mobility management protocols, rather than keeping these issues (QoS and mobility) dissasoociated. To enable seamless mobility, it is important to secure QoS when the MN roams to a new access router. However, it minimizes the mobility impact by assuring a particular level of service across heterogeneous networks. Some work has already been performed in this area [72], but it is still not completely explored.

- The model introduced in this dissertation was tested using the simulation tool OPNET. More research can be done by implementing it in a test bed and comparing the test results.

- Security issues for the MN, neighboring users and proxy servers need to be studied further.

GLOSSARY

A.1 Definition of terms and concepts

Access router: An access network router residing on the edge of an access network and connected to one or more access points. An access router offers IP connectivity to mobile hosts. The access router may include intelligence beyond a simple forwarding service offered by ordinary IP routers.

Base station: Also called access point, it is the point of attachment of a MN to the Internet. Binding The association of the home address of a MN with the care-of address of that MN, along with the remaining lifetime of that association.

Care-of address: A unicast routable address associated with a MN's visiting a foreign link; the subnet of this IP address is a foreign subnet prefix.

Context-aware handover: A handover that is governed by a certain specific requirement to be fulfilled while handing the connection between two access routers.

Correspondent node: A peer node with which a MN is communicating.

Eager cell switching Node should switch to the new access router as early as possible, or as soon as the MN receives a router advertisement from the new access router.

Fast handover: A handover that aims primarily to minimize delay, with no explicit interest in packet loss.

Handover: The act of changing the attachment point of a MN, switching the communications from one access point to another access point, also know as handoff.

Handover latency: Handover latency is the time difference between when a mobile host is last able to send and/or receive an IP packet by way of the old access router, until when the mobile host is able to send and/or receive an IP packet through the new accessrouter.

Home address: A unicast routable address assigned to a MN, used as the permanent address of the terminal.

Home Agent: A router on the MN's home link with which the MN has registered its current care-of address.

Horizontal handover: Also know as intra-technology handover, a handover between two cells (or access points) employing the same air interface technology.

MN: A node that can change its point of attachment from one link to another, while still being reachable via its home address.

Movement: A change in a MN's point of attachment to the Internet.

Network domain: A grouping of network objects, such as computers, that simplifies the naming of network services. Within a domain, all the names must be unique.

Registration: The process during which a MN sends a binding update to its home agent or a correspondent node, causing a binding for the MN's to be registered.

Seamless handover: A handover that is both smooth and fast, thus provides fast lossless handover between two access routers.

Vertical handover: Also called inter-technology handover, a handover between two cells employing different air interface technologies.

REFERENCES

[1]. Goodman, David. “A Dummy’s Introduction to GPRS”, WINLAB, February 22, 1999. [Online], 2006 Available at: http://win.rutgers.edu/pub/symposiums/GPRS/David%20Goodman.pdf

[2]. Langton, Charon. “The Concept of signal spreading and its uses in communications.” CDMA, tutorial, 2002, [Online], 2009 Available at: http://ece.ut.ac.ir/Classpages/F86/ECE637/FILES/material/CDMA.pdf

[3]. David P. Hole, Fouad A. Tobagi, “Capacity of an IEEE 802.11b Wireless LAN supporting VoIP”, *in Proc. IEEE Int. Conference on Communications (ICC)*, Vol. 1, pp.196-201, 2004.

[4]. K. W. Richardson. “UMTS Overview”, *Electronics & Communication Engineering Journal* Vol.12 (3) pp. 93-100, June 2000.

[5]. A.Dutta, P.Agarwal, S.Das, M. Elaoud, D. Famolari, S. Madhani, A. McAuley, B. Kim, P. Li, M. Tauil, S. Baba, Y. Ohba, T. Kodama, N. Nakajima, Jyh-Cheng Chen, Henning Schulzrinne “Realizing mobile wireless Internet telephony and streaming multimedia testbed”, *Computer Communicaton* Vol. 27, Issue 8, pp. 725-738, 2004.

[6]. J.Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M.Handley, and E. Schooler. “SIP: Session Initiation Protocol”, RFC 3261, June 2002.

[7]. Tingzhou Yang, Yixin Dong, Yihan Zhang, Dimitrios Makrakis, “Practical Approaches for Supporting Micro Mobility with MPLS”, *International Conference on Telecommunications (ICT 02)*, Beijing, China, Vol. 2, June 2002.

[8]. “Emerging Wireless Technologies: A Look into the Future of Wireless Communications beyond 3G.”, SAFECOM. [Online], 2006 Available at: http://www.safecomprogram.gov/NR/rdonlyres/5C74C631-ACF6-433F-B313-C04D041A5489/0/Look_Future_Wireless_Communications_Beyond3G.pdf

[9]. R. Gellen, “Mapping Between the Multimedia Messaging Service (MMS) and Internet Mail”, RFC 4356, Jan 2006.

[10]. Mitch Tulloch, Ingrid Tulloch, "Encyclopedia of Networking", 2nd edition, Microsoft Press, ISBN: 0735605734, June 2000.

[11]. Richard. C. Braley, Ian C. Gifford, Robert F. Heile. "Wireless Personal Area Networks an Overview of the IEEE P802.15 Working Group", *ACM SIGMOBILE Mobile Computing and Communications*, Vol. 4 , Issue 1, pp. 26-33, Jan 2000.

[12]. Brian P Crow, Indra Widjaja, Jeong Geun Kim, Prescott T. Sakai. "IEEE 802.11 Wireless Local Area Networks", *IEEE Communications Magazine,* September 1997.

[13]. Abd-Elhamid Taha, Abd-Elhamid Mohamed., "A Framework for Radio Resource Management in Heterogeneous Wireless Network", A Thesis submitted to the Department of Electrical and Computer Engineering, Queen University, September 2007, [Online], 2006 Available at: http://qspace.library.queensu.ca/bitstream/1974/745/1/taha_abd-elhamid_m_200709_PhD.pdf

[14]. Subir Das, Anthony McAuley, Ashutosh Dutta, Archan Misra, Kaushik Chakraborty, Sajal K. Das. "IDMP: Intra-Domain Mobility Management Protocol for Next Generation Wireless Networks", *IEEE Wireless Communications.* Vol. 9, pp. 38-45, 2002.

[15]. K. Lai, M. Roussopoulos, D. Tang, X. Zhao, M. Baker. "Experiences with a Mobile Testbed", *Proceedings of the Second International Conference on Worldwide Computing and its Applications (WWCA 1998),* Volume 1368/1998, pp. 222-237, 1998.

[16]. A. .T. Campbell, J. Gomez-Castellanos, C. Wan, S. Kim, Z. Turanyi, A. Valko. "Cellular IP. Internet Draft", IETF, January 2000. [Online], (2009), Available at: draft-ietf-mobileip-cellularip-00.txt.

[17]. R. Ramjee, T. La Porta, S. Thuel, K. Varadhan, S. Y. Wang. "HAWAII: A Domain-Based Approach for Supporting Mobility in Wide-Area Wireless Networks", *Proceedings of the Seventh International Conference on Network Protocols*, Vol. 10, Issue:3, pp.183-292, 1999.

[18]. A. T. Campbell, J. Gomez-Castellanos, S. Kim, A. Valko, C. Wan, Z. Turanyi., "Design, Implementation, and Evaluation of Cellular IP", *IEEE Personal Communications Magazine*, Vol. 7, Issue. 4, pp. 42-49, 2000.

[19]. S. Das, A. Misra, P. Agrawal, S. K. Das. "TeleMIP: Telecommunication- Enhanced Mobile IP Architecture for Fast Intra-Domain Mobility", *IEEE Personal Communications Magazine* Vol.7, Issue. 4, pp. 50-58, 2000.

[20]. Petros Zerfos, Xiaoqiao Meng, H. Starsky, Y. Wong, Vidyut Samanta, Songwu Lu. "A Study of the Short Message Service of a Nationwide Cellular Network", *Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement*, Vol. 6, pp. 263-268, 2006.

[21]. R. Hsieh, A. Seneviratne. "A Comparison of Mechanisms for Improving Mobile IP Handoff Latency for End-to-End TCP", *Proceedings of the Nineth ACM International Conference on Mobile Computing and Networking (MobiCom 03)*, pp. 29-41, Sep 2003.

[22]. Y. Mun, "Fast Macro Mobility Handovers in HMIPv6", 2007, [Online], (2009), Available at: draft-mun-mipshop-fhmacro-2.txt.

[23]. Robert Hsieh, Zhe Guang Zhou, Aruna Seneviratne., "S-MIP: A Seamless Handoff Architecture for Mobile IP", *IEEE, INFOCOM.* Vol. 3, pp.1774-1784, Apr 2003.

[24]. Pablo Vidales, "Seamless mobility in 4G systems.", *Technical report based on a dissertation submitted May 2005, University of Cambridge*, ISSN 1476-2986, 2005, [Online], (2009), Available at: http://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-656.pdf

[25]. A.C.Snoeren, H. Balakrishnan. "An End-to-End Approach to Host Mobility", *Proceedings of the Sixth ACM International Conference on Mobile Computing and Networking (MobiCom 2000)*, pp.155-166, 2000.

[26]. A.C.Snoeren, H. Balakrishnan, M. F. Kaashoek, "Reconsidering Internet Mobility", *In Proceedings of the Eighth Workshop on Hot Topics in Operating Systems (HotOS-VIII)*, pp.41–46, May 2001.

[27]. "Introduction to Mobile IP." Cisco White paper. "Mobility and Mobile IP, Introduction", White paper- ipUnplugged, 2003. [Online], (2009), Available at: http://www.cisco.com/en/US/docs/ios/solutions_docs/mobile_ip/mobil_ip.pdf

[28]. "MPLS Overview". JUNOS 7.2 MPLS Applications Configuration Guide, [Online], (2009), Available at: http://www.juniper.net/techpubs/software/junos/junos72/swconfig72-mpls-apps/toc/collapsedTOC.htm

[29]. "Introduction to MPLS", Cisco White paper Release 9.3.10, Part Number 78-11658-01 Rev.B0, May 2001, [Online], (2009), Available at: http://www.cisco.com/univercd/cc/td/doc/product/wanbu/bpx8600/mpls/9_3_1/mpls01.pdf

[30]. "Virtual Private Networking: An Overview", Microsoft TechNet. , [Online], (2009), Available at: http://technet.microsoft.com/en-us/library/bb742566.aspx

[31]. "MPLS Label Distribution Protocol (LDP)", Corporate Headquarters: Cisco Systems, Inc-2005, [Online], (2009), Available at: http://www.cisco.com/univercd/cc/td/doc/product/software/ios124/124newft/124_x/ftldp41.pdf

[32]. D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, Dec 2001.

[33]. B. Jamoussi, R. Callon, R. Dantu, L. Wu, P. Doolan, T. Worster, N. Feldman, A. Fredette, M. Girish, E. Gary-Sandburst, J. Heinanen, T. Kilty, and A. Malice. "CR-LDP: Constraint-based Label Distribution Protocol Overview", RFC 3212, 2002.

[34]. E.Rosen, A. Vishwanathan, and R.Callon., "Multiprotocol Label Switching Architecture", RFC 3031, Jan 2001.

[35]. Chen. Khong, Tham Chun, Choong Foo, and Chi-Chung Ko. "Integration of Mobile IP and Multi-Protocol Label Switching", *Communications, 2001. ICC IEEE International Conference,* Vol. 7, pp.2123-2127, 2002.

[36]. Abd-Elhamid M Taha, Hossam Hassanein, and Hussein T. Mouftah. Integrated "Solutions for Wireless MPLS and Mobile IP: Current Status of Future Directions", *Proceedings of the Electrical and Computer Engineering 2004 Canadian Conference.* Vol. 3, pp.1463-1466, May 2004.

[37]. Tingzhou Yang, Dimitrios Makrakis. "Hierarchical Mobile MPLS: Supporting Delay Sensitive Applications Over by Broadband Wireless and Internetworking", *Proceedings. ICII 2001. Beijing 2001 International Conferences*, Vol. 2, pp.453-458.

[38]. "SIP Protocol Overview" by RADVISION VOIP Experts. [Online], (2009), Available at: http://www.radvision.com/NR/rdonlyres/51855E82-BD7C-4D9D-AA8A-E822E3F4A81F/0/RADVISIONSIPProtocolOverview.pdf

[39]. K. Zeilenga, ed. "Open LDAP Foundation, Lightweight Directory Access Protocol (LDAP)", RFC4510, June 2006.

[40]. M. Handley, V.Jacobson, "Session Description Protocol Overview", RFC 2327, April 1998.

[41]. Cisco Systems, Inc. "DHCP Overview", [Online], (2009), Available at: http://cco.cisco.com/en/US/docs/ios/ios_xe/ipaddr/configuration/guide/iad_dhcp_ovrvw_xe.pdf

[42]. Cisco Systems, Inc. "Introduction to DNS", [Online], (2009), Available at: http://www.cisco.com/en/US/docs/net_mgmt/network_registrar/6.2/user/guide/UG15_DNS .pdf

[43]. M. O'Doherty, "Pico SIP Optimized Link State Routing Protocol (OLSR)", RFC 3626, Oct 2003.

[44]. K.D. Dutta, J. Wong, R. Burns, A. Jain, K. McAuley, Young, H. Schulzrinne, "Realization of integrated mobility management protocol for ad hoc networks", *MILCOM 2002*, Vol. 1, pp. 448-454, 2002.

[45]. R. Matei, "JAIN SIP Approach in Ad-Hoc Networks Mobility Management", *Ad Hoc Mobile Wireless Networks Research Seminar on Telecommunications Software*, 2002, [Online], (2009), Available at: http://www.tml.tkk.fi/Studies/T110.557/2002/papers/razvan_matei.pdf

[46]. I. Chakeres, J. Macker, and T. Clausen. "Mobile Ad-hoc Networks Architecture (MANET)". Internet Engineering Task Force, 2007, [Online], (2009), Available at: http://tools.ietf.org/pdf/draft-chakeres-manet-arch-01.pdf.

[47]. Henning Schulzrinne, ElinWedlund, "Application-Layer Mobility Using SIP", *ACM SIGMOBILE Mobile Computing and Communications Review*. Vol. 4, Issue. 3, pp. 47-57, July 2000.

[48]. Elin Wedlund, "Mobility Support using SIP", *Proceedings of the 2nd ACM International Workshop on Wireless Mobile Multimedia*, pp. 76-82, Aug 1999.

[49]. Perkins, Charles, and David B. Johnson. "Route Optimization in Mobile IPv6", Working Group Internet Draft, RFC 4866, Feb 2000.

[50]. H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", RFC 1889, Jan1996.

[51]. Cisco Networks. "Overview Application Layer Gateway and Translation Type Support." [Online], (2009), Available at: http://www.cisco.com/en/US/technologies/tk648/tk361/tk438/technologies_white_paper09186a00801af2b9.pdf

[52]. Nakamura, Motonori, and Jun-ichiro itojun Hagino, "SMTP Operational Experience in Mixed IPv4/IPv6 Environments", RFC 3974, 2002.

[53]. Clifton, Lin. "AODV Routing Implementation for Scalable Wireless Ad-Hoc Network Simulation (SWAN)", [Online], (2009), Available at: http://jist.ece.cornell.edu/docs/040421-swans-aodv.pdf

[54]. Sze-Yao Ni, Yu-Chee Tseng, Yuh-Shyan Chen, and Jang-Ping Sheu. "The Broadcast Storm Problem in a Mobile Ad Hoc Network", *International Conference on Mobile Computing and Networking, Proceedings of the 5th annual ACM/IEEE International Conference on Mobile Computing and Networking,* Vol. 8, Issue. 2/3, pp 151–162, Aug 1999.

[55]. Ochi Yoshikage, Takahisa Okazaki, Kazuhiko Kinoshita, Hideki Tode, and Koso Murakami. "An Extended AODV Routing for Reduction of Control Messages in Ad hoc Networks", *Consumer Communications and Networking Conference, CCNC,* Vol. 4, pp.74-78, May 2007.

[56]. Choon Hoong Ding, Sarana Nutanong, and Rajkumar Buyya. "Peer-to-Peer Networks for Content Sharing.", Grid Computing and Distributed Systems Laboratory, Department of Computer Science and Software Engineering, University of Melbourne, Australia, [Online], (2009), Available at: http://www.buyya.com/papers/P2PbasedContentSharing.pdf

[57]. Kai Shen, and Yuan Sun. "Distributed Hash Table on Pre-structured Overlay Networks", *"WCW 2004: international workshop on web content caching and distribution",* Vol. 3293, pp. 63-81, Oct 2004.

[58]. Nilanjan Banerjee, Arup Acharya, and Sajal K. Das. "Enabling SIP-based Sessions in ad hoc Networks.", *Journal of Wireless Networks,* Vol. 13, Issue 4, August 2007, 461-479. ISSN: 1022-0038.

[59]. N. Freed, and J. Klensin., "Multipurpose Internet Mail Extensions (MIME) Part Four: Registration Procedures", RFC 4289, December 2005.

[60]. David Goodman, "A Dummy's Introduction to GPRS." WINLAB, February 22, 1999, [Online], (2006), Available at: http://winwww.rutgers.edu/pub/symposiums/GPRS/David%20 Goodman.pdf

[61]. "What is a Wireless LAN?" White Paper. Proxim, [Online], (2006), Available at: http://www.sss-mag.com/pdf/proximwhatwlan.pdf

[62]. K. W. Richardson, "UMTS Overview.", [Online], (2009), Available at: http://www.roke.co.uk/download/articles/ UMTS_Overview.pdf

[63]. Hewlett Packard White Paper. "CDPD Conceptual Overview.", [Online], (2009), Available at: http://cp.literature.agilent.com/litweb/pdf/ 5965-6326E.pdf

[64]. J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks,M. Handley, and E. Schooler. "SIP: Session Initiation Protocol RFC 3261", June 2002.

[65]. SuKyoung Lee, "A MPLS-Based Micro-mobility Supporting Scheme in Wireless Internet", *Journal of Lecture notes in computer science*, vol. 3314, pp. 160-165, Dec 2004.

[66]. Tang Tang, MI Zhengkun, and PENG Rongqun. "Adaptive Service Provisioning through Context-aware SIP Proxy", *Fourth International Conference on Networking and Services (icns 2008)*, pp.277-281, 2008.

[67]. Ashutosh Dutta, Subir Das, David Famolari, Yoshihiro Ohba, Kenichi Taniuchi, Victor Fajardo, Rafa Marin Lopez, Toshikazu Kodama, and Henning Schulzrinne, "Seamless Proactive Handover across Heterogeneous Access Networks", *Wireless Personal Communications: An International Journal.* Volume 43, Issue 3, 837- 855, Nov 2007.

[68]. B. Aboba, D. Simon. "PPP EAP TLS Authentication Protocol", IETF RFC 2716. Oct 1999.

[69]. Rigney Willens, S., A. Livingston, Merit Rubens, and Daydreamer W. Simpson. "Remote Authentication Dial in User Service (RADIUS)", RFC 2865, June 2000.

[70]. Bryan J. Byerly, and David Williams, "SIP Authentication Using CHAP Password", October 2000, [Online], (2009), Available at: http://tools.ietf.org/pdf/draft-byerly-sip-radius-00.pdf

[71]. Nilanjan Banerjee, Sajal K. Das, and Arup Acharya. "SIP-based Mobility Architecture for Next Generation Wireless Networks.", *Proceedings of the 3rd IEEE International Conference on Pervasive Computing and Communications (PerCom),* pp. 181-190, March 2005.

[72]. Dimitra Vali, Sarantis Paskalis, Alexandros Kaloxylos, and Lazaros Merakos, "An Efficient Micro-Mobility Solution for SIP Networks", *Global Telecommunications Conference, (GLOBECOM '03). IEEE.* Vol. 6, pp. 3088- 3092 Dec 2003.

[73]. Jin-Woo Jung, Hyun-Kook Kahng, Ranganathan Mudumbai, and Doug Montgomery, "Performance Evaluation of two Layered Mobility Management using Mobile IP and Session Initiation Protocol.", *Global Telecommunications Conference, (GLOBECOM '03). IEEE.* Vol. 3. pp. 1190- 119, Dec 2003.

[74]. J. D.Goodman, and S. X. Wei. "Efficiency of Packet Reservation Multiple Access", *Vehicular Technology, IEEE Transactions on.* Vol. 40, No. 1, pp 170-176, Feb 1991.

[75]. Gunasundari, R., S. Shanmugavel and K. Bakyalakshmi. "Fault Tolerant HAWAII Protocol for IP Services in Wireless Networks", *ainaw. 22nd International Conference on Advanced Information Networking and Applications Workshops*, pp. 457-462, 2008.

[76]. H. C. Soliman, Castelluccia, K. El Malki, and L. Bellier. "Hierarchical Mobile IPv6 Mobility Management (HMIPv6)", RFC4140, Aug 2005.

[77]. M. S. Bargh, H. Zandbelt, and A. Peddemors. "Managing Mobility in 4G Environments with Federating Service Platforms (an Overview)", *Proceedings of EVOLUTE Workshop, Evolute '03, (Guildford, UK)*, pp.1-6, Nov 2003.

[78]. M. S. Bargh, D. Bijwaard, H. Zandbelt, E. Meeuwissen, A. Peddemors. "Mobility Management in Beyond 3G Environments", *High speed networks and multimedia communications. IEEE international conference*, vol. 3079, pp. 545-555, July 2004.

[79]. B. Roach, "Session Initiation Protocol (SIP): Specific Event Notification" RFC 3265, Network Working Group, June 2002.

[80]. Christos Politis, Kar Ann Chew, Rahim Tafazolli. "Multilayer Mobility Management for All-IP Networks: Pure SIP vs Hybrid SIP/Mobile IP", *Vehicular Technology Conference, 2003. VTC 2003-Spring. The 57th IEEE Semiannual*, Vol. 4, pp. 2500- 2504, Apr 2003.

[81]. Karl Andersson, Muslim Elkotob, and Christer Ahlund, "A New MIP-SIP Interworking Scheme." *In ACM International Conference Proceeding Series, Proceedings of the 7th International Conference on Mobile and Ubiquitous Multimedia (MUM2008),* pp. 117-120, Dec 2008.

[82]. Chung-Ming Huang, Chang-Zhou Tsai, "*Proceedings of the 5th international conference on Ubiquitous Intelligence and Computing*", Vol. 5061, pp. 636 – 646, 2008.

[83]. Willem A. Romijn, Dirk-jaap Plas, Dennis Bijwaard, Erik Meeuwissen, Gijs van Ooijen "Mobility Management for SIP session in a Heterogeneous Network Environment", Bell Labs Technical Journal, Vol. 9, Issue 3, pp. 237-253, 2004.

[84]. Huei-Wen Ferng, Nigel Hsiung, David Shiung, Jeng-Ji Huang, "A SIP-Based Mobility Management Architecture Supporting TCP with Handoff Optimization", *Vehicular Technology Conference, VTC2007-Spring. IEEE 65th*, pp. 1224-1228, Nov 2007.

[85]. Wang Miao, Zhang Yu-Jun, Li Jun, "A Fast Handover Solution for SIP-based Mobility", *Wireless and Mobile Computing, Networking and Communications,. (WiMob'2006). IEEE International Conference,* pp. 315-320, June 2006.

[86]. Stefano Salsano, Andrea Polidoro, "SIP-Based Mobility Management in next Generation Networks", *IEEE Wireless Communications*, Vol.15, issue 2, pp. 92-99, Apr 2008.

[87]. Paul Barford, Mark Crovella, "Generating Representative Workloads for Network and Server Performance Evaluation", *In Proceedings of the 1998 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems,* pp.151-160, July 1998.

[88]. "Code Division Multiple Access (CDMA).The Concept of Signal Spreading and its Uses" Intuitive Guide to Principles of Communications, [Online], (2009), Available at: http://www.complextoreal.com/CDMA.pdf.

[89]. M.E. Crovella and A. Bestavros. "Self-similarity in worldwide web traffic: Evidence and possible causes". *In Proceedings of the 1996 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems,* pp. 160-169, May 1996.

[90]. M.F. Arlitt and C.L. Williamson. "Web server workload characterization: The search for invariants", *In Proceeding of the ACM SIGMETRICS '96 Conference*, pp. 126-137, April 1996.