IMPROVING THROUGHPUT BY TRAFFIC AWARE ROUTING IN NON-TRANSPARENT IEEE 802.16J NETWORKS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

RIDVAN TEKDOĞAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
COMPUTER ENGINEERING

MAY 2010

Approval of the thesis:

IMPROVING THROUGHPUT BY TRAFFIC AWARE ROUTING IN NON-TRANSPARENT IEEE 802.16J NETWORKS

submitted by **RIDVAN TEKDOĞAN** in partial fulfillment of the requirements for the degree of **Master of Science in Computer Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences** ________________

Prof. Dr. Adnan Yazıcı
Head of Department, **Computer Engineering** ________________

Dr. Cevat Şener
Supervisor, **Computer Engineering Dept., METU** ________________

Assoc. Prof. Dr. İbrahim Körpeoğlu
Co-supervisor, **Computer Engineering Dept., Bilkent University** ________________

Examining Committee Members:

Prof. Dr. Müslim Bozyiğit
Computer Engineering Dept., METU ________________

Dr. Cevat Şener
Computer Engineering Dept., METU ________________

Dr. Attila Özgit
Computer Engineering Dept., METU ________________

Dr. Sinan Kalkan
Computer Engineering Dept., METU ________________

Dr. Tolga Numanoğlu
Senior Engineer, ASELSAN ________________

Date: ________________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: RIDVAN TEKDOĞAN

Signature :

ABSTRACT

IMPROVING THROUGHPUT BY TRAFFIC AWARE ROUTING IN NON-TRANSPARENT IEEE 802.16J NETWORKS

Tekdoğan, Rıdvan
M.S, Department of Computer Engineering
Supervisor : Dr. Cevat Şener
Co-Supervisor : Assoc. Prof. Dr. İbrahim Körpeoğlu

May 2010, 83 pages

Worldwide Interoperability for Microwave Access (WiMAX) is one of the rising communications technology which enables last mile broadband mobile wireless Internet connectivity. IEEE Std 802.16-2009 is the last accepted standard which targets mobile and fixed wireless broadband access. The standard defines two types of stations which are base and mobile stations. A base station has a wired connection to backhaul network and gives broadband wireless service to mobile stations. This communication architecture requires dense deployment of base stations to provide communication without interruption. IEEE 802.16j standard which is an amendment to IEEE 802.16, introduces ”Multihop Relaying” for increasing coverage and throughput. Deployment of relay stations, where the backbone network does not exist, is a cost effective solution. Two modes of operations are defined for relay station: transparent mode and non-transparent mode. Relays in transparent mode, are deployed for improving signal quality, so that mobile stations can use relay link for increasing throughput. In non-transparent mode, relays can send management packets, so that mobile stations, which are not in the direct reach of a base station, can connect to network through relay stations.

In domain specific networks, like tactical area networks used in military, main data traffic is caused by the communication between subscribers in same region. In IEEE 802.16j Networks, each network packet is sent to a base station, and that base station forwards the packet to destination in such cases. In this thesis shortcut routing scheme is proposed as sending packets to destination directly through relay station for data traffic between two subscribers with a common relay. With shortcut routing, network throughput is increased by preventing links at higher layer in topology from becoming bottleneck. Moreover, by traversing fewer hops, latency decreases. We also propose traffic aware path selection method, where a path will be chosen according to the traffic demand so that subscribers can benefit from shortcut routing more.

Keywords: WiMAX, IEEE 802.16j Non-transparent, Shortcut routing, Path Selection, Multihop Relaying

ÖZ

SAYDAM OLMAYAN IEEE 802.16J AĞLARINDA TRAFİĞE DUYARLI YÖNLENDİRME İLE VERİ KAPASİTESİNİ YÜKSELTME

Tekdoğan, Rıdvan
Yüksek Lisans, Bilgisayar Mühendisliği Bölümü
Tez Yöneticisi : Dr. Cevat Şener
Ortak Tez Yöneticisi : Yrd. Doç. Dr. İbrahim Körpeoğlu
Mayıs 2010, 83 sayfa

WiMAX son mil geniş bant kablosuz İnternet erişimi sağlayan yakın geleceğin iletişim teknolojisidir. IEEE Std 802.16-2009 kabul görmüş en güncel WiMAX standardıdır. Bu standart sabit ve mobil geniş bant ağ erişimini belirlemektedir. Standart mobil ve baz istasyonu olmak üzere iki çeşit istasyon türü tanımlamaktadır. Baz istasyonunun omurga ağına kablolu bağlantısı bulunur ve mobil istasyonlara geniş bant kablosuz ağ yayını yapar. Kesintisiz kablosuz iletişim için mevcut standart birçok baz istasyonu kurulumu gerektirmektedir. Arazi koşullarında böyle bir kurulum yapmak mümkün olmayabilir. IEEE 802.16j Sekmeli Röle iyileştirmesi, uygun maliyetli röle kullanımıyla kapsama alanını genişletmeyi ve veri kapasitesini yükseltmeyi sağlar. Rölelerin saydam ve saydam olmayan iki türlü çalışma modu vardır. Saydam modda mobil istasyonlar sadece veri aktarımını daha kaliteli röle bağlantılarıyla yaparlar. Saydam olmayan röleler, yönetim paketleri gönderme özelliğine sahip olduklarından, kapsama alanını genişletme amaçlı da kullanılırlar.

Taktik saha iletişim ağı gibi özel iletişim ağlarında, abonelerin aynı alan içerisindeki diğer abonelerle veri iletişimi İnternet trafiğinden daha yoğun olabilir. IEEE 802.16j Sekmeli Röle ağlarında, kullanıcıların ağ trafiği baz istasyonuna kadar iletilip; baz istasyonu hedef istasyona

veriyi yönlendirir. Bu çalışmada ortak röleye sahip, iletişim içerisinde olan abonelerin veri trafiğinin rölelerden yönlendirilmesi anlatılacaktır. Trafiğin akış yönüne göre mobil istasyonlar için yol seçme yöntemi önerilecektir. Bu önerme sayesinde, mobil istasyonlar alan içi görüşme yaptıkları istasyonlarla ortak rölelere bağlanmaları sağlanacaktır. Sonuç olarak ağdaki veri iletim hacmi artırılacak ve paket gecikmeleri azaltılacaktır.

Anahtar Kelimeler: WiMAX, Saydam Olmayan Sekmeli IEEE 802.16j ağları, Kestirme Yönlendirme, Yön Seçme

To My Family

ACKNOWLEDGMENTS

I am heartily thankful to my supervisors, Cevat Şener and İbrahim Körpeoğlu, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject.

It is an honor for me to thank my jury members for their valuable comments on writing this thesis.

I owe my deepest gratitude to my family; Selahattin Tekdoğan, Ayser Tekdoğan, Ömer Faruk Tekdoğan and Furkan Tekdoğan for their love and encouragement.

I am grateful to my friends and colleagues: Oğuz Özün, Füruzan Atay Onat, Tolga Numanoğlu, Eda Kılıç, Güven Yenihayat and Onur Dizdar for their help and encouragement. This thesis would not have been possible without their knowledge sharing.

I would like to thank ASELSAN and TUBİTAK for their supporting grants throughout my master program.

Lastly, I offer my regards and blessings to all of those who supported me in any respect during the completion of this thesis.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

4G 4th Generation
AES Advanced Encryption Standard
AMC Adaptive Modulation and Coding
ARQ Automatic Repeat Request
ATM Asynchronous Transfer Mode
BE Best Effort
BPSK Binary Phase-Shift Keying
BS Base Station
BWA Broadband Wireless Access
CDMA Code Division Multiple Access
CID Connection Identifier
MAC CPS MAC Common Part Sublayer
CRC Cyclic Redundancy Check
MAC CS MAC Convergence Sublayer
D/A Digital/Analog
DCD DL channel descriptor
DES Data Encryption Standard
DL Downlink
DL-MAP Downlink Map
DSL Digital Subscriber Line
FDD Frequency Division Duplexing
FTP File Transfer Protocol
GUI Graphical User Interface
IEEE Institute of Electrical and Electronics Engineers
LAN Local Area Network
LOS Line of Sight
MAC Media Access Control
MAN Metropolitan Network Area
MIMO Multiple input Multiple Output
MR-BS Multihop Relay Base Station
MS Mobile Station
NCTUns National Chiao Tung University Network Simulator
NLOS Non-Line of Sight
nrtPS Non-Real-Time Polling Service
NS2 Network Simulator 2
NT-RS Non-Transparent Relay Station
OFDM Orthogonal Frequency Division Multiplexing
OFDMA Orthogonal Frequency Division Multiple Access
PDU Protocol Data Unit
PHS Packet Header Suppression
PHY Physical Layer
PMP Point to Multipoint
QAM Quadrature Amplitude Modulation
QOS Quality of Service
QPSK Quadrature Phase-Shift Keying
RNG-REQ Ranging Request
RNG-RSP Ranging Response
RS Relay Station
RSA Rivest Shamir Adleman
rtPS Real-Time Polling Service
SDUs Service Data Units
SFID Service Flow Identifier
SNR Signal to Noise Ratio
SS Subscriber Station
MAC SS MAC Security Sublayer
Std Standard
TCP Transmission Control Protocol
TLV Type Length Value
T-RS Transparent Relay Station
UCD UL channel descriptor
UGS Unsolicited Grant Service
UL Uplink
UL-MAP Uplink Map

WIMAX Worldwide Interoperability for Microwave Access

WMAN Wireless Metropolitan Area Network

CHAPTER 1

INTRODUCTION

WiMAX is a wireless communication technology which provides broadband network access over long distances. WiMAX standardization activities are carried out by the IEEE 802.16 Working Task Group and WiMAX Forum. Current active IEEE 802.16 standard defines the air interface for fixed and mobile wireless broadband access systems. The standard defines two types of nodes which are Base Stations (BS) and Mobile Stations (MS). A base station is connected to a possible wired backhaul network and serves wirelessly connected mobile stations. The communication architecture of IEEE 802.16 is Point to Multi Point (PMP), where a BS serves multiple MSs. This communication architecture may increase the deployment cost in rural areas where obstacles make have shadow effect. Therefore, the IEEE 802.16j standard introduces a Multihop Relay specification for wireless broadband access. By using Relays Stations (RS), the coverage of Multihop Relay - Base Station (MR-BS) is extended.

The IEEE 802.16j standard defines two different relay modes, which are *Transparent Mode* and *Non-transparent Mode*. In transparent mode, relays do not forward any framing information. Transparent mode relays are used for increasing network *capacity*. Subscribers use relay links with better CINR ratio for sending data. Management messages should be directly received from MR-BS in transparent IEEE 802.16j Networks. Non-transparent mode relays can send framing information. Subscribers that are not in the range of MR-BS can connect to a relay station (RS), since relays stations send synchronization messages and they can be received but the mobile stations in the range. In this way, a mobile station can connect to an RS from which it has received a synchronization message. Consequently, by using non-transparent mode relays, network *coverage* can be extended, in addition to network capacity.

Two scheduling modes are available for IEEE 802.16j relays, which are centralized scheduling and distributed scheduling. MR-BS is responsible for scheduling uplink (UL) and downlink (DL) transmission of all MSs and RSs in centralized scheduling. Therefore, MR-BS generates ULMAPs and DLMAPs, and the relay stations forward them to mobile stations. In distributed scheduling, relay stations schedule uplink and downlink transmission of their subordinates (mobile stations and other relay stations) which are directly connected.

Considering the bandwidth requirements of today's network applications, an IEEE 802.16j network can also be used as a backhaul network for Tactical Area Communication Systems. In such a scenario there are many mobile subscribers scattered in a tactical field and communicating with each other. These communicating subscribers in a field may benefit from the same relay. However, in IEEE 802.16j, even though two communicating subscribers are connected to the same RS, their network packets are routed to the BS first and then the BS forwards them to the subscribers. Going always to the base station may consume the network resources unnecessarily in multi-hop relay networks and may increase the latency.

For achieving uninterrupted communication, deployment of RSs and MR-BS may form a dense topology. In such a topology, MSs and RSs may be in the transmission range of more than one RS and MR-BS. Therefore, each RS and MS should select an access station (relay station or MR-BS) to get service. As a result of this, an MS may have several alternative paths that could be formed to connect the MS to the MR-BS system. Hence, there is a path selection problem here, and path selection plays an important role for improving the network performance. In IEEE 802.16j standard, the MR-BS system determines the paths for mobile stations (MSs).

In literature, many path selection algorithms are proposed to construct the network topology among the mobile stations, relay stations and MR-BS, with the objective of utilizes the network resources more efficiently. SNR, hop count, load of a node and end-to-end available bandwidth are some of the metrics that are used in those proposed methods.

Forwarding data directly to the destination from a relay station is a new issue for IEEE 802.16j multihop relay networks. A local forwarding scheme is proposed which forwards data of subscribers connected to the same RS directly. Local forwarding scheme, however, does not handle routing of data from an intermediate relay station that is a common relay station on the paths from two communicating mobile stations to the same MR-BS. Additionally, there is no path selection algorithm which considers shortcut routing and traffic demand together.

In this thesis, we propose a shortcut routing technique and a traffic aware path selection method for better utilizing the resources and improving the throughput in non-transparent IEEE 802.16j networks. We propose the use of a new layer, named CROSS, on top of the 802.16 MAC layers of relay stations for performing shortcut routing and keeping IP data traffic statistics. In 802.16j, while establishing a path between an MS and MR-BS, a connection is created for the MS and indentified with a connection identifier (CID). In our CROSS layer, we maintain the association between the CID and IP address allocated to the MS and that association (mapping) is kept in a dictionary structure enabling fast lookups. Accordingly, when an UL IP packet is received at a relay station from a connection going through it, a check is made whether the destination IP address of the packet is in CID-IP mapping dictionary. If the destination address is in the dictionary, this means the destination MS is one of the subordinates of the relay station and it is possible to forward the packet in downlink (DL) direction instead of sending it towards the MR-BS (uplink direction).

Shortcut routing is meaningful when subscribers, communicating with each other, share a common RS. To benefit from shortcut routing advantages, our traffic aware path selection method selects access stations (relay stations) of subordinates according to the sent and received traffic analysis (i.e. according to the traffic demands). All subordinate stations keep a list of RSs and MR-BSs in the vicinity. When an IP packet is received, it is checked whether the IP address is in the address domain of one of the neighbor superordinate station. Hierarchical addressing is used to be able to easily and quickly determine the location of a node in the network topology.

Outline of the rest of the thesis is as follows. In Chapter 2 background information is given to facilitate the better understanding of the proposed solution. Studies on shortcut routing and path selection methods for IEEE 802.16j networks are discussed in Chapter 3. Chapter 4 presents our proposed solution that includes a shortcut routing proposal and a traffic aware path selection method. Our simulation environment, topologies and traffic scenarios used in the simulations, and our simulation results are given and discussed in Chapter 5. Finally, chapter 6 concludes the thesis together with comments about future work.

CHAPTER 2

BACKGROUND

This chapter will provide background information of the proposed solution. Brief information will be given for WiMAX and standardization activities. IEEE Std 802.16-2009 and IEEE Std 802.16j-2009 will be explained with details used in thesis. Overview of MobileIP protocol will be mentioned as the last topic of this chapter.

2.1 WiMAX Standards

”WiMAX, meaning Worldwide Interoperability for Microwave Access, is a telecommunications technology that provides wireless transmission of data using a variety of transmission modes, from point-to-multipoint links to portable and fully mobile Internet access”[4]. Nowadays WiMAX is also known as 4G of wireless communication technology. IEEE 802.16 standard is basis for the WiMAX. WiMAX Forum and IEEE 802.16 Working Group are official standardization organizations [4]. 802.16 working group composed of many task groups working on different aspects of wireless broadband standardization. List of current standards and their status are given in Table 2.1. 802.16d standard defines fixed broadband wireless access and it is first accepted standard. 802.16e introduces mobility and merged with 802.16d, both them are included in IEEE Std 802.16-2009.

2.1.1 IEEE 802.16

IEEE 802.16 standard defines fixed and mobile wireless broadband access. Last version of the standard released at 2009 which combines IEEE 802.16d and IEEE 802.16e. IEEE 802.16d

Table 2.1: IEEE 802.16 Projects and Standards

Standard	Description	Status
802.16-2001	Fixed Broadband Wireless Access (10-63 GHz)	Superseded
802.16.2-2001	Recommended practice for coexistence	Superseded
802.16c-2002	System profiles for 10-63 GHz	Superseded
802.16a-2003	Physical layer and MAC definitions for 2-11 GHz	Superseded
P802.16b	License-exempt frequencies (Project withdrawn)	Withdrawn
P802.16d	Maintenance and System profiles for 2-11 GHz (Project merged into 802.16-2004)	Merged
802.16-2004	Air Interface for Fixed Broadband Wireless Access System (rollup of 802.16-2001, 802.16a, 802.16c and P802.16d)	Superseded
P802.16.2a	Coexistence with 2-11 GHz and 23.5-43.5 GHz (Project merged into 802.16.2-2004)	Merged
802.16.2-2004	Recommended practice for coexistence (Maintenance and rollup of 802.16.2-2001 and P802.16.2a)	Current
802.16f-2005	Management Information Base (MIB) for 802.16-2004	Superseded
802.16-2004/Cor 1-2005	Corrections for fixed operations (co-published with 802.16e-2005)	Superseded
802.16e-2005	Mobile Broadband Wireless Access System	Superseded
802.16k-2007	Bridging of 802.16 (an amendment to IEEE 802.1D)	Current
802.16g-2007	Management Plane Procedures and Services	Superseded
P802.16i	Mobile Management Information Base (Project merged into 802.16-2009)	Merged
802.16-2009	Air Interface for Fixed and Mobile Broadband Wireless Access System (rollup of 802.16-2004, 802.16-2004/Cor 1, 802.16e, 802.16f, 802.16g and P802.16i)	Current
802.16j-2009	Multihop relay	Current
P802.16h	Improved Coexistence Mechanisms for License-Exempt Operation	in progress
P802.16m	Advanced Air Interface with data rates of 100 Mbit/s mobile & 1 Gbit/s fixed	in progress

Standard specifies the fixed wireless broadband access. Only PMP LOS communication was supported in this standard. 802.16e introduces the mobility and compared to its legacy IEEE 802.16d, has lower data rate. IEEE 802.16 standard specifies the MAC and PHY layers. MAC layer supports point-to-multipoint (PMP) architecture. Design of MAC supports different types of PHY layer specification [7]. OFDMA PHY is suitable physical layer for mobile WiMAX systems.

2.1.1.1 IEEE 802.16 OFDMA PHY

IEEE 802.16 OFDMA (Orthogonal Frequency Division Multiple Access) PHY operates in frequencies less than 11 GHz. OFDM frequency division multiplexing scheme divides frequency band into many subcarriers. Data in each subcarrier is modulated traditionally as if sending data on parallel streams. OFDM supports single user at a time. OFDMA groups subset of subcarriers to form subchannels. By assigning these subchannels different subscribers, multiple users can communicate concurrently on same channel. Example OFMDA symbol shown in Figure 2.1. As shown in the Figure 2.1, subchannels may be constructed from inconsecutive subcarriers. In addition to OFDMA, IEEE 802.16 PHY support Adaptive Antenna Systems (AAS) and Multiple Input Multiple Output (MIMO) technology which enhances the capacity and coverage of network [8].

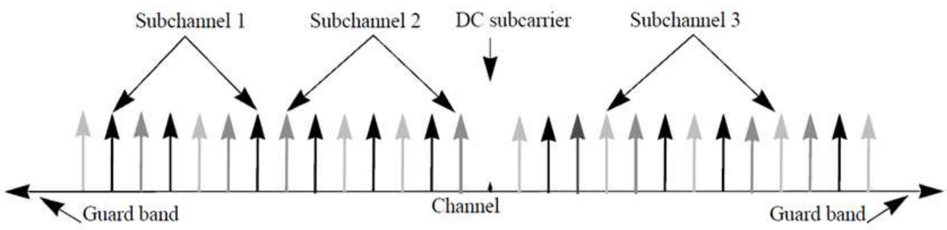

Figure 2.1: OFDMA Frequency description [7]

Minimum data allocation unit of OFDMA PHY is a slot. Slot is two dimensional data structure with frequency and time domain. Time domain is composed of OFDMA symbols and frequency domain consists of subchannels. A Burst is a user data to be sent in one OFDMA frame. Since multiple users can communicate at the same time OFDMA frame consists of many bursts. When transmitting data, bursts should be mapped on slots in a frame. Example allocation of slots in OFDMA frame is shown in Figure 2.2. ODMA frame is divided into two main parts for Uplink (UL) and Downlink (DL) data. Each frame starts with preamble and ends with a guard. UL and DL parts are also separated with guard.

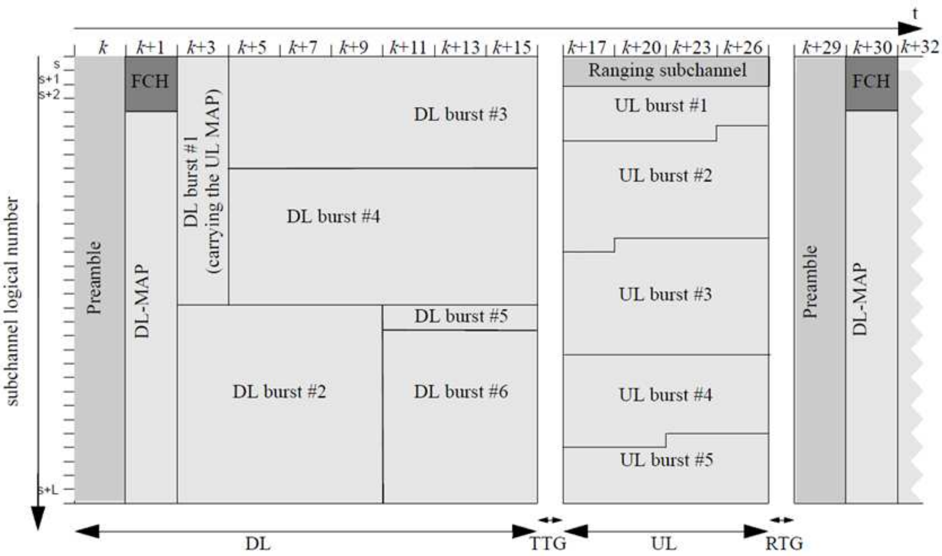

Figure 2.2: Example OFDMA Frame [7]

2.1.1.2 IEEE 802.16 MAC

IEEE 802.16 MAC layer composed of three sublayers which are Service Specific Convergence Sublayer (CS), MAC Common Part Sublayer (MAC CPS) and Security Sublayer. Reference model is given on Figure 2.3. CS is adaptation layer for different upper layer protocols like ATM or TCP/IP. CPS is responsible for system access, connection management and bandwidth allocation.

MAC layer supports point-to-multipoint (PMP) architecture which means one or more mobile subscribers (MS) connect to one base station (BS). In downlink (DL) direction, BS transmits data without coordinating with subscribers. DL transmission is broadcast, every subscribers receives the transmitted frame. In DLMAP, sent in DL frame, it may not be explicitly stated which portion of subframe belongs to a specific subscriber. In such cases subscribers inspect the connection id (CID) of incoming packets and determine whether it belongs to itself. In uplink (UL) direction, users share uplink frame by adhering on a transport protocol. UL subframe allocation is determined by BS according to subscribers' demands. MAC layer allocates bandwidth according the transport connections' QoS class. UL frame sharing is handled by scheduling services implemented using unsolicited grants, contention and polling mechanisms [7].

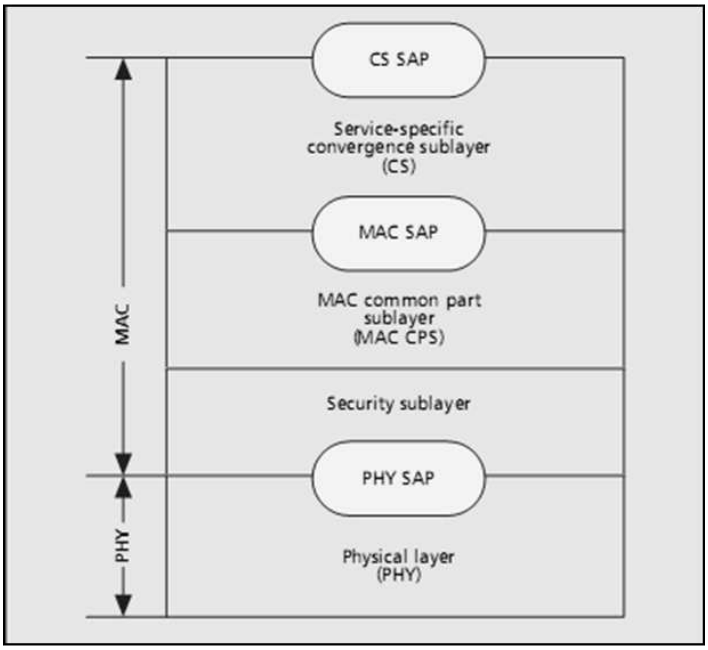

Figure 2.3: IEEE 802.16 Reference Model [8]

Addressing and Connections

Every 802.16 device has 48 bit MAC address which uniquely identifies the air interface. This address is used to establish the appropriate connections for SS. Connections in IEEE 802.16 are identified with 16 bit CIDs. During SS initialization three pairs of management connections are established which are basic connections (UL and DL), primary connections (UL and DL) and secondary connections (UL and DL). These connections are used for different level QoS management messages. Basic connections are used for urgent, short management messages. For longer and delay tolerant messages, primary connections are utilized. Secondary connections are dedicated to delay tolerant standard based messages like TFPT and DHCP.

CIDs of connections should be assigned in the RNGRSP, REGRSP or MOB-BSHO-REQ/RSP messages. UL and DL connections of same type are assigned to same CID. For each QoS class bandwidth demand, a new transport connection is established.

MAC PDU Formats and Management Messages

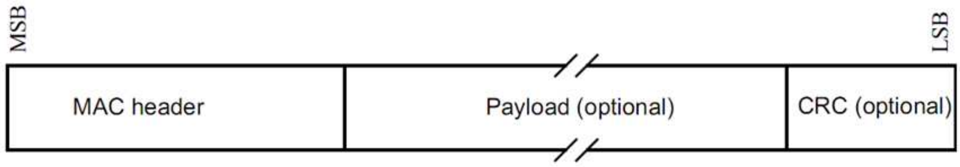

Figure 2.4: MAC PDU Format

IEEE 802.16 MAC PDU consists of three parts as shown in Figure 2.4. MAC header and CRC parts are fixed length and Payload is variable size. Some of the UL management messages may consist of just MAC header. OFDMA PHY layer requires CRC part included in MAC PDUs. Payload may contain zero or more subheaders.

MAC management messages starts with Type field and message payload follows it. Management messages carried on basic, primary, initial ranging and broadcast connections. Messages carried on basic and primary connection are not fragmented. Some of the MAC management messages listed in Table 2.2 with their descriptions and connections.

Table 2.2: MAC Management Messages [7]

Type	Message Name	Message Description	Connection
0	UCD	UL Channel descriptor	Fragmentable broadcast
1	DCD	DL Channel descriptor	Fragmentable broadcast
2	DL-MAP	DL Access definition	Broadcast
3	UL-MAP	UL Access definition	Broadcast
4	RNG-REQ	Ranging request	Initial and basic
5	RNG-RSP	Ranging response	Initial and basic
6	REG-REQ	Registration request	Primary management
7	REG-RSP	Registration response	Primary management
11	DSA-REQ	Dynamic service addition request	Primary management
12	DSA-RSP	Dynamic service addition response	Primary management
13	DSA-ACK	Dynamic service addition acknowledge	Primary management
14	DSC-REQ	Dynamic service change request	Primary management
15	DSC-RSP	Dynamic service change response	Primary management
16	DSC-ACK	Dynamic service change acknowledge	Primary management
17	DSD-REQ	Dynamic service deletion request	Primary management
18	DSD-RSP	Dynamic service deletion response	Primary management
26	SBC-REQ	SS basic capability request	Basic
27	SBC-RSP	SS basic capability response	Basic
56	MOB_BSHO-REQ	BS HO request message	Basic
57	MOB_MSHO-REQ	MS HO request message	Basic
58	MOB_BSHO-RSP	BS HO response message	Basic
59	MOB_HO-IND	HO indication message	Basic
60	MOB_SCN-REP	Scanning result report message	Psrimary management

Scheduling

For each transport connection, there is an associated scheduling service to handle the data. Scheduling mechanism differs according to service flow over transport connections. There are four scheduling types defined in MAC layer to meet bandwidth and delay requirements of different user demands. QoS classes and descriptions explained in Table 2.3. Requirements are parameterized in order to be used in scheduling algorithm. Algorithms employ polling, unsolicited grants and contention methods according to QoS parameters. For example for VoIP traffic, UGS QoS is selected and unsolicited grants given to subscriber to guarantee bandwidth for constant bitrate traffic. Unsolicited grants method reserves bandwidth upon a demand from subscriber if enough resource is available. Polling methods are used for services that requires different amount of bandwidth during transmission. BS polls subscriber whether it should increase or decrease the bandwidth. Contention is used for services which are delay tolerant like FTP. Every subscriber tries to get as much as in contention period and all of them have equal chance.

Table 2.3: QoS Classes [8]

QoS category	Applications	QoS specifications
UGS Unsolicited grant service	VoIP	Maximum sustained rate Maximum latency tolerance Jitter Tolerance
rtPS Real-time polling service	Streaming audio or video	Minimum reserved rate Maximum sustained rate Maximum latency tolerance Traffic priority
nrtPS Non-real-time polling service	File transfer Protocol (FTP)	Minimum reserved rate Maximum sustained rate Traffic priority
BE Best effort service	Data transfer, Web browsing, etc.	Maximum sustained rate Traffic priority

Network Entry and Initialization

Network entry and initialization scenario for SS is described in activity diagram given in Figure 2.5.

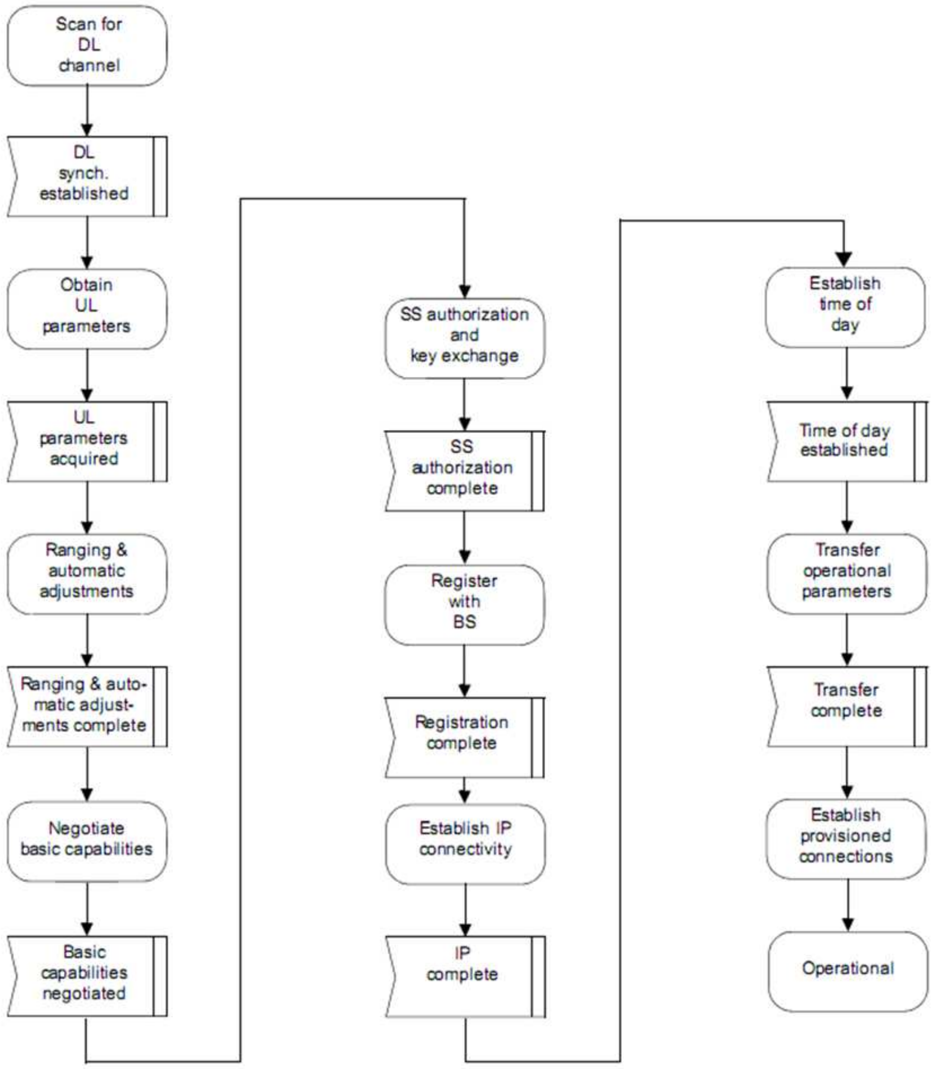

Figure 2.5: Initialization of SS [7]

When a new SS enters to network, it starts to scan the wireless media and obtains DL-MAP and UL-MAP of neighbor BSs. Receiving a DL-MAP, SS is synchronized with BS. DL-MAP and messages describes the burst allocations of a frame for downlink and uplink respectively. SS prepares RNG-REQ for initialing ranging process after receiving DCD and UCD which give information about downlink and uplink channel respectively. Receiving RNG-RSP message ranging process is ends. SS sends SBC-REQ which contains information about SS capabilities and BS respond back with SBC-RSP which contains information about the BS's capabilities. By this way SS and BS negotiate with their common capabilities. SS sends REG-REQ the register BS after authorization phase passed. If BS accepts SS registration request, it responds with positive REG-RSP message. After this point if SS is accepted, then management and transport connections established and SS become operational.

Handover

Handover is process of a SS moves from radio interface of serving BS to radio interface of neighboring BS. Main reasons for handover are better signal quality and better QoS. When a SS moves, signal received from serving BS may attenuate. Interference and fading also affects the signal quality received by SS. In such cases taking service from other neighboring BS which has better signal, is more appropriate. Sometimes SS could not get enough service quality from the serving BS. Connecting to BS which has more available bandwidth is preferable.

BSs know the topology of network by communicating with each other over backbone network. BS broadcast topology information periodically with MOB_NBR-ADV message. MSs learn the neighboring BSs from this message. Also SS could requests scanning interval from BS to obtain neighboring BS information. MS obtains scanning interval by sending MOB_SCN-REQ to serving BS. Handover decision could be taken by either MS or serving BS. If MS decides, it sends MOB_MSHO-REQ to serving BS. If vice versa, BS sends MOB_BSHO-REQ to MS. After this point MS starts synchronize with target BS as in initialization phase. The rest of the process is similar to MS network entry and initialization as depicted in Figure 2.6.

2.1.2 IEEE 802.16j Multihop Relay Specification

IEEE 802.16j standard is developed as an amendment to IEEE Std 802.16 by IEEE Multihop Relay (MR) task group. The standard updates and expands IEEE 802.16 for covering multihop relay operations. IEEE 802.16j describes data structures and control mechanisms for MAC common sublayer, security sublayer and physical layers of the MRs [1]. IEEE 802.16j standard designed to operate with legacy IEEE 802.16 mobile subscribers (MS).

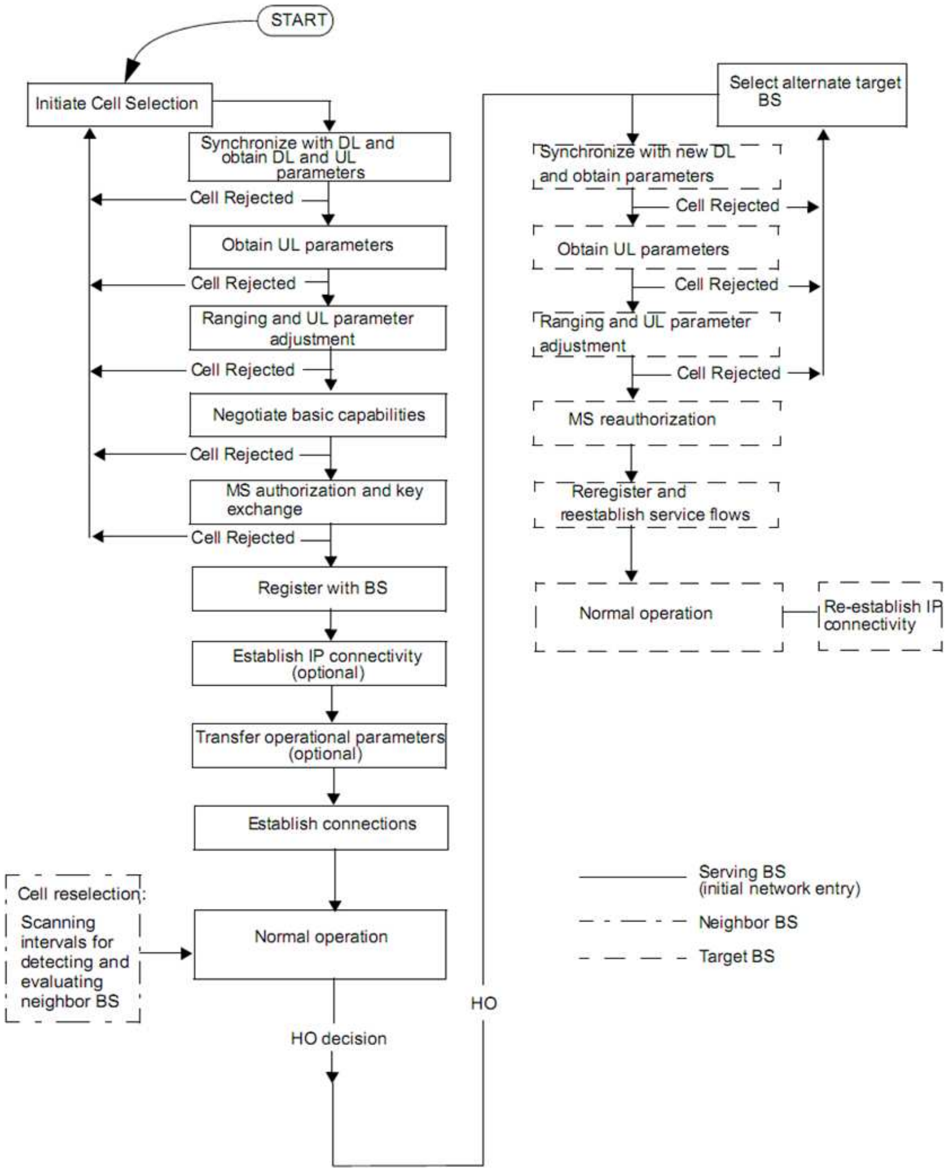

Figure 2.6: Network Initialization and Handover [7]

Introducing multihop relaying concept brings cost effective solutions for extending coverage and increasing the capacity. Specification supports mobility for relay stations (RS). Three types of relays exist according to deployment scenario, which are mobile, fixed and nomadic. Main use cases of RSs are shown in Figure 2.7.

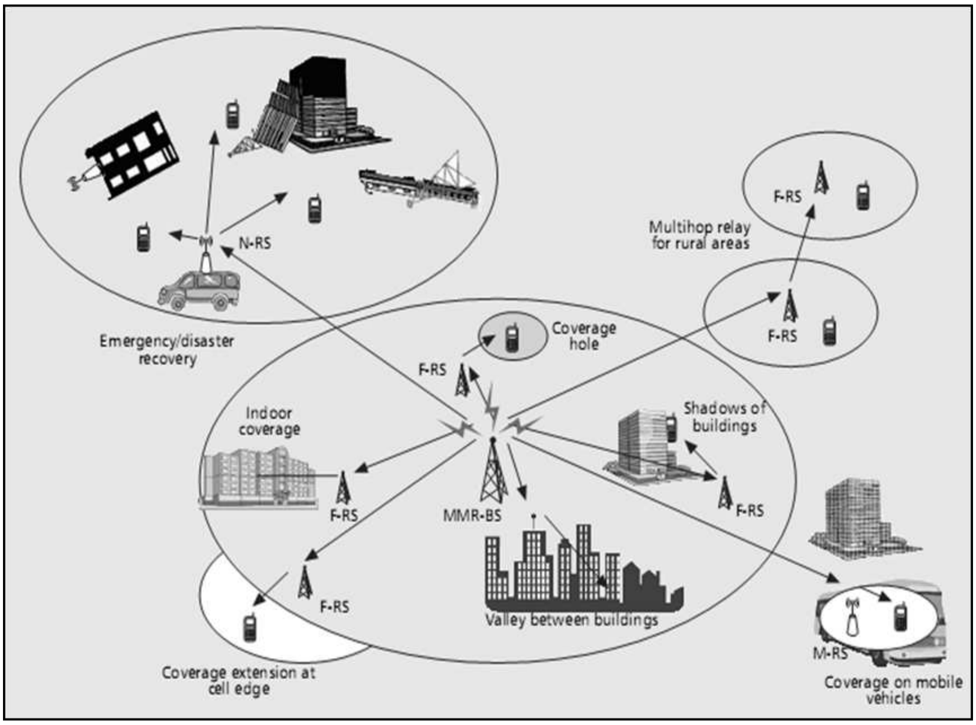

Figure 2.7: Examples of usage scenarios for fixed (F-Rs), nomadic (N-RS), and mobile stations (M-RS) [3]

IEEE 802.16j defines two types of operation modes for relays which are transparent and non-transparent mode. In transparent mode RSs do not send framing information which contains scheduling information for subscribers. Management messages should be received directly from the BS by subscribers in this mode. Non-transparent relays can send framing information, therefore management messages sent by RS to its subordinate subscribers.

There are two types of scheduling related to these operation modes: centralized scheduling and distributed scheduling. In centralized scheduling, BS is responsible for assigning UL and DL slots to each subscriber even if connected over RS. In distributed scheduling every RS generates DLMAP and ULMAP which contains the scheduling information [2].

Transparent mode:

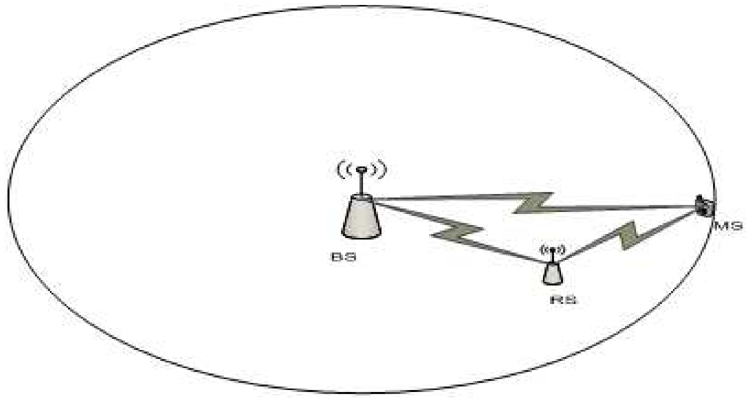

Figure 2.8: Usage scenario for transparent relays

Relays operating in transparent mode do not have transmits framing information. Therefore centralized scheduling mode is used. Subscriber could connect to BS at most over two hops. Consequently relays operating in transparent modes do not extend the coverage of BS. These types of relays used for increasing the network capacity. Example usage is shown in Figure 2.8.

Non-transparent mode:

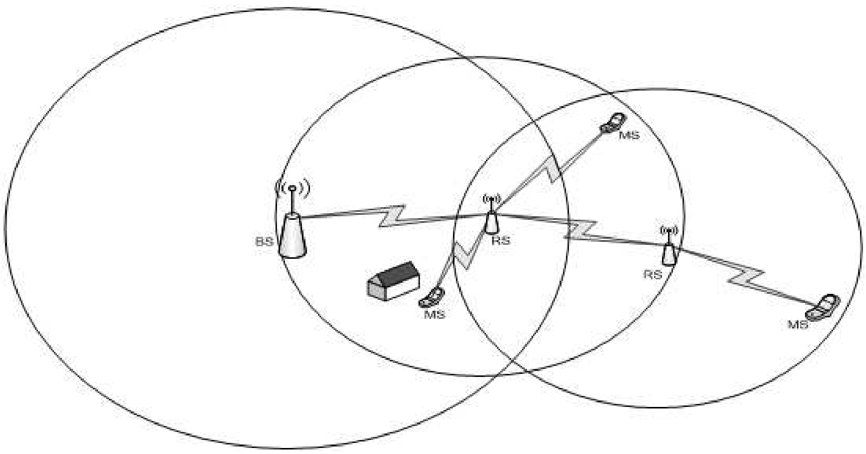

Figure 2.9: Usage scenarios for non-transparent relays

Relays in this mode enables extending coverage by sending framing information. Framing information can be generated by RS or can be directly forwarded according to selected scheduling mechanism [2]. Multihop topologies supported in both distributed and centralized scheduling. Due to interference caused by sending management messages by relays, capacity enhancement is not well as in transparent mode. Example usage of non-transparent relays is shown in Figure 2.9. Comparison of two modes is given in Table 2.4.

Table 2.4: Comparison between transparent and non-transparent modes of operation [2]

	Transparent RS	Non-transparent RS
Coverage Extension	No	Yes
Number of Hops Inter RS cell interference	2 None	2 or more High
Performance	In BS coverage: high Outer BS coverage: -	In BS coverage: same as 16e Outer BS coverage: medium
RS Cost	Low	High
Scheduling	Centralized scheduling only	Centralized/distributed scheduling

2.1.2.1 PHY Layer Specification

IEEE 802.16j provides some modifications to frame structure to capture the multihop relaying needs. IEEE 802.16 frame consist of UL and DL subframes. UL and DL subframes divided into two zones, one is for ordinary access zone which contains MS bursts and the other is relay zone for relay burst. UL and DL access zone structure is similar to legacy IEEE 802.16 UL and DL access zone. DL Relay zone starts R-MAP which contains framing information for UL and DL transfers, and relay fast recovery channel (R-FCH). The rest of DL zone is allocated for relay DL bursts. UL Relay zone contains relay UL bursts. UL and DL bursts contain data generated from or destined to relay nodes. The structure of an IEEE 802.16 PHY frame is shown in Figure 2.10.

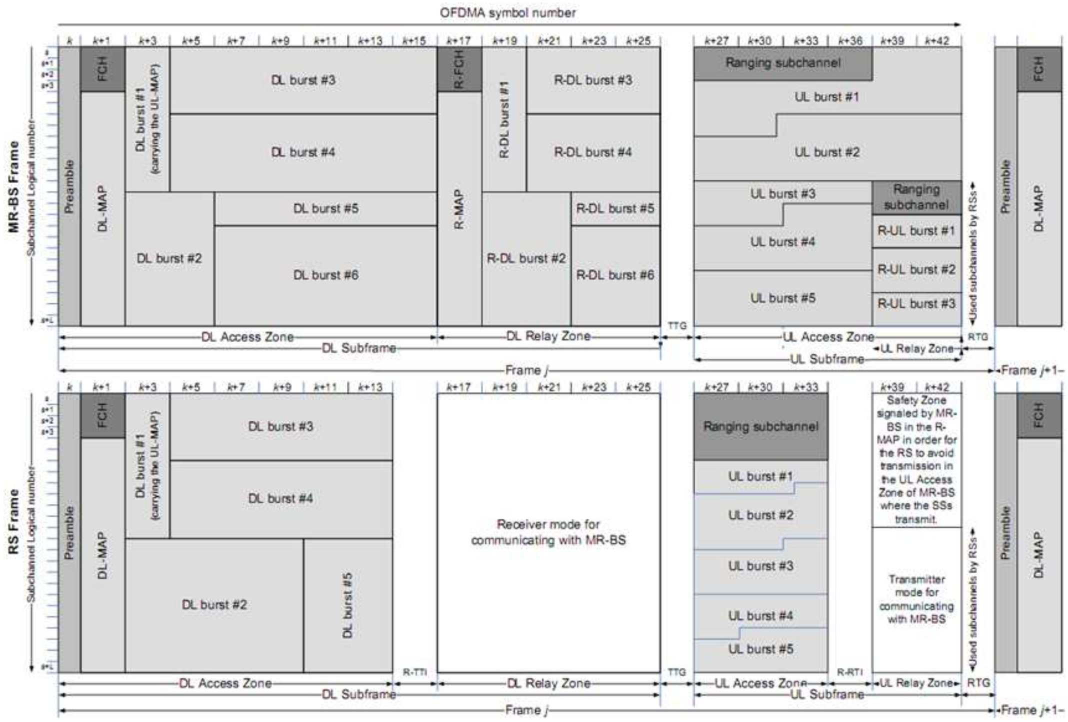

Figure 2.10: IEEE 802.16j Relay Frame Structure [1]

2.1.2.2 MAC Layer Specification

IEEE 802.16j MAC layer specification includes data structures and control mechanisms for multihop relaying. New management messages are introduced and some of the former messages modified for relay stations. Concerned management messages listed in Table 2.5.

Addressing and Connections

Connections defined in IEEE 802.16 standard are applicable for RS and MR-BS. Connections between MR-BS and SS may span over one or more RSs. Each RS must be associated with a BS in order to serve SSs and subordinate RSs. Connections are identified with CID as in former standard and CIDs must be unique in each cell. Multihop relay specification introduces tunnel connection concept. Tunnel connections are established between RS and MR-BS or superordinate RS. MAC PDUs from different connections may use same tunnel connection. Using tunnel connection, end-to-end transport connections passing through the intermediate links decreased to one. Usage of tunnel connections is optional.

Table 2.5: Relay MAC Management Messages

Type	Message Name	Message Descriptions	Connection
0	UCD	UL Channel Descriptor	Fragmentable Broadcast RS Primary Management Multicast Management
1	DCD	DL Channel Descriptor	Fragmentable Broadcast RS Primary Management Multicast Management
70	RCD	R-link channel descriptor	RS Primary Management RS Multicast Management
71	MR_NBR-INFO	Multihop relay neighbor information	RS Primary Management
72	MR_RNG-REPORT	Multihop relay ranging report	RS Basic
87	RS_AccessRS-REQ	RS access station selection request	RS Basic
92	RS_Access-MAP	MAP information in centralizes scheduling mode	RS Basic RS Multicast Management

Paths between MR-BS and SS can be established either by using tunnel connections or CID based connections. In CID based connection, R-links with same CID are established between RS and its superordinate station. When a PDU arrived to RS, packet forwarded to connection with CID stated in the MAC header. Figure 2.11 illustrates the CID based connections.

Relay Path Management and Routing

Based on topology, MR-BS determines the path between itself and access station of MS. Path selection decision is made according to tree topology constraints and available resources. In tree topology a RS should connect to one superordinate station. Path selection algorithm is not defined in IEEE 802.16j standard, it is left to vendors.

MR-BS selects access station of RS during network entry. RS performs measurement report to inform MR-BS about neighbors as shown in Figure 2.12 and wait for RS_AccessRS-REQ message. When MR-BS receives measurement report from RS via RS_NBR_MEAS-REP

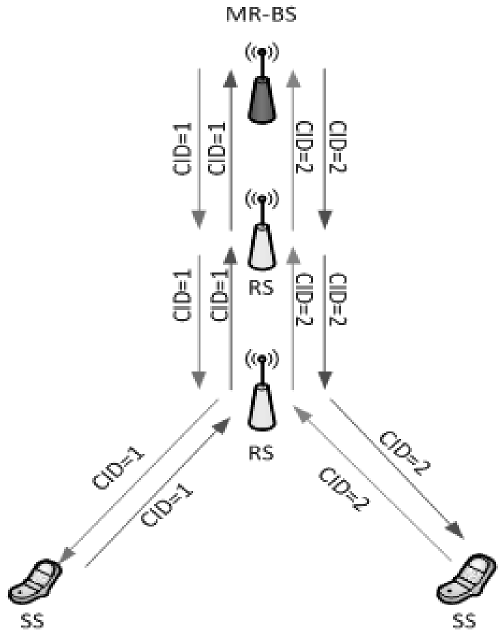

Figure 2.11: CID based connection management

message, it performs path selection for RS as shown in Figure 2.13. If the access station for RS is changed MR-BS sends RS_AccessRS-REQ message. Upon receiving the message RS sends MR_Generic-ACK to MR-BS and initiate network reentry procedure as shown in Figure 2.14.

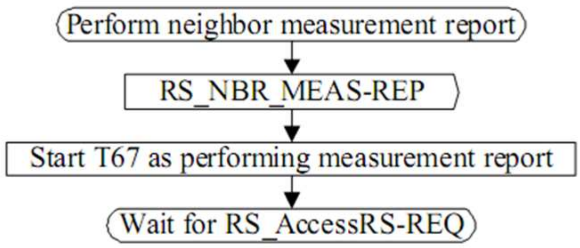

Figure 2.12: Perform neighbor measurement report at a RS

There are two types of path management. One of them is embedded path management where

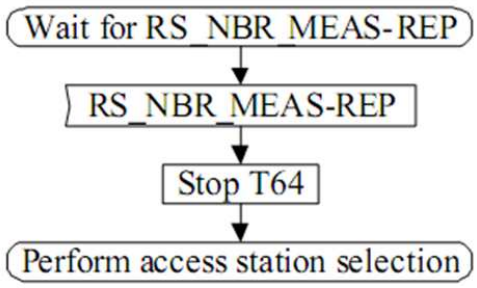

Figure 2.13: Handling RS Measurement Report

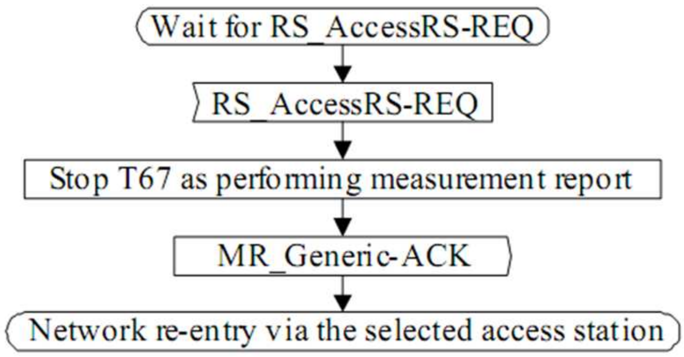

Figure 2.14: Network reentry after path selection

there is no need to keep routing table. Second is explicit path management which requires routing table to keep the association between path and CIDs. In embedded path management CIDs assigned systematically so that arrived MAC PDU's next R-link is determined according to CID. MR-BS assigns CID blocks to its subordinate RSs and each RS also assigns consecutive block of CIDs to its subordinates. Example assignment is shown in Figure 2.15. According to the assignment in Figure 15, when it is needed to establish a transport connection between MR-BS and a MS connected to RS tagged with 'H', CID of the connection is picked between 1 and 100. If a MAC PDU with CID=9, arrives to intermediate RS B then RS will forward the PDU to R-link to RS D. RS D then forwards it to RS H.

MR-BS detects the topology changes when a new MS or RS connect or disconnects from

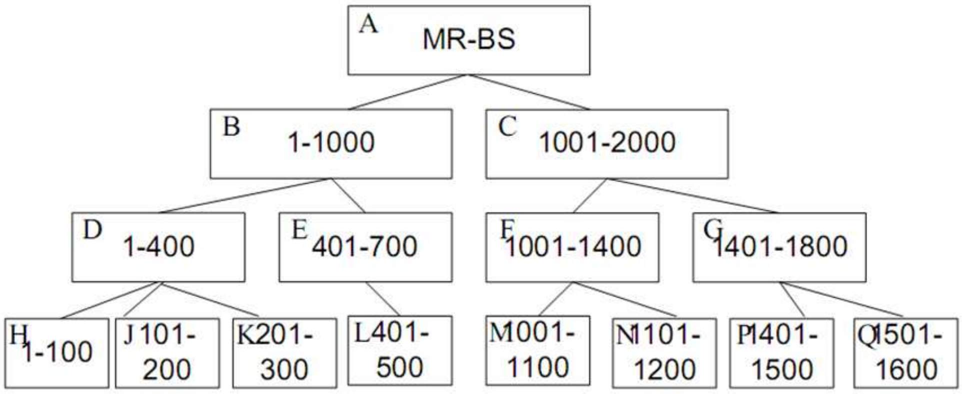

Figure 2.15: CID assignment in embedded path management

the network. Upon such topology changes, MR-BS rearranges the CID and PATH mapping. When a new MS or RS is attached to MR-BS, path for that subordinate station is determined by MR-BS and all the intermediate RSs are informed about the PATH-CID mapping with DSA-REQ message. Furthermore, when MR-BS decides to remove a path, it sends a DSD-REQ message to all intermediate RSs.

When a new connection is established, PATH-CID mapping in routing table should be updated. After MR-BS decides the path for routing the new connection, it sends path information which contains PATH ID and CID, to all intermediate RSs on the path. Upon receiving DSA-REQ message, RS insert entry to routing table and forwards message to subordinate RS which is the next hop on the path. A DSA-REQ message propagates until it reaches the access station. Access station responds with DSA-RSP message to inform all superordinate RSs and MR-BS about success of path creation. For removal of a path, MR-BS sends DSD-REQ messages to RSs on the path, and all the intermediate RS updates the routing table by removing associated the path entry.

2.2 Mobile IP

MobileIP is extension to IP protocol developed by IETF (Internet Engineering Task Force) to handle mobility issues. MobileIP enables mobile devices stay connected without changing home address while roaming. Mobile nodes which support Mobile IP have two IP addresses which are Home Address and Care Of Address. Home Address is assigned by Home Agent which runs at mobile nodes home network. Care Of Address is assigned when mobile node moves to another network. Care Of Address is temporary whereas home address is permanent. Care of address is not transparent to upper layer protocols. All upper layer protocols and applications use home address for addressing source and destination nodes.

2.2.1 Overview of Protocol

There are two types of agent defined in protocol. Home Agent is running on home network of the mobile node. Home agent maintains the mobility binding table which contains the home address and care of address mapping of mobile node. Foreign agent is running on the foreign network where the node is visiting temporarily, and gives service. Both services periodically broadcast the agent information in order to be discovered by mobile nodes.

When mobile node leaves the home network and enters to foreign network, it sends a registration request with lifetime information to foreign agent. Foreign agent resolves the home network of node by inspecting its home address, and sends a request with care of address to be assigned to node. Upon receiving the request from foreign agent, home agent keeps a record of mobile nodes current location and its care of address and then responds to request. After receiving response from home agent, foreign agent accepts connection request of visiting mobile node and updates the visiting lists. Registration process is summarized in Figure 2.16.

After registration a tunnel between home agent and foreign is established to forward packets destined to home address of mobile node its current address. Since upper layers transparent to location change, IP packets are constructed with source address is home address. Packets are routed to destination as usual. Destination node of the traffic never realizes the location of the mobile node. If the correspondent node intends to send packet to mobile node, it uses home address as destination address. Routers will forward the packet to home network of the mobile node. Receiving a packet destined to mobile node, home agent checks the mobility binding table and forwards the packet to care of address of the node by encapsulating packet with a new IP header. Encapsulated packet is sent to foreign agent over tunnel established in registration phase. Upon receiving encapsulated packet, foreign agent extracts original IP packet and sends it to mobile node. Tunneling operation is shown in Figure 2.17.

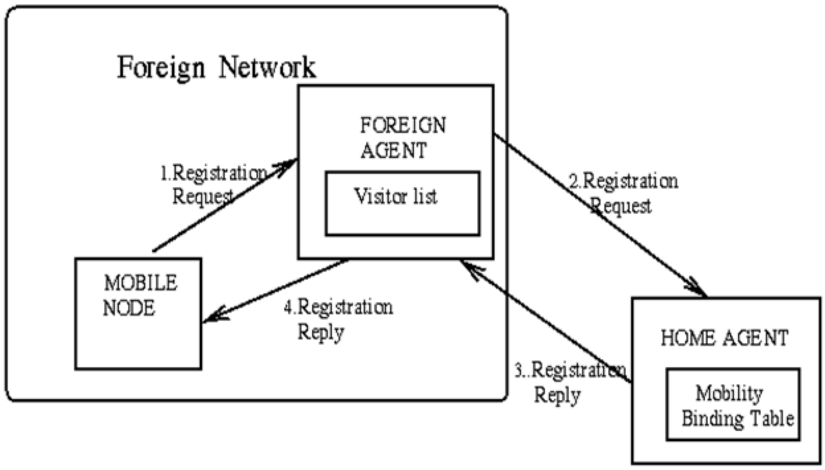

Figure 2.16: CID based connection management

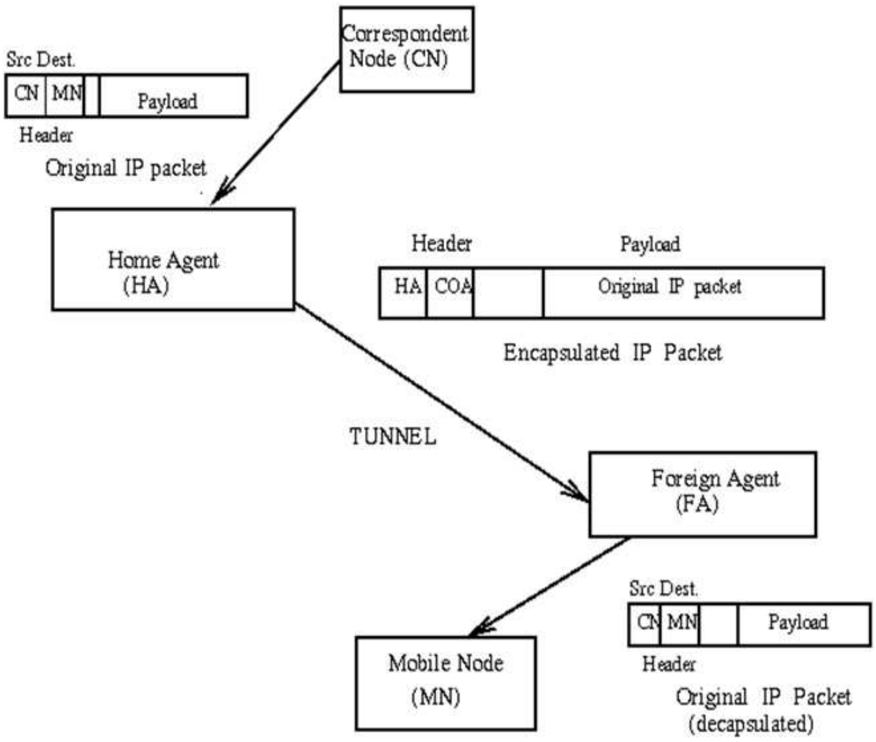

Figure 2.17: CID assignment in embedded path management

CHAPTER 3

LITERATURE SURVEY

IEEE 802.16j Standard specifies path establishment and management procedures. Path selection method is intentionally left to vendors. Relays route data according to path defined by MR-BS. According to the standard all UL data should be routed to MR-BS. As a result user data between two subscribers connected to same access RS station will traverse over MR-BS. For the subscribers connected to same access RS station, a routing path like SS1-RS-SS2 would be more efficient. Routing data to subordinates instead of routing to superordinate station in relay based systems is called shortcut routing, and introduced in [10]. For the path selection algorithm, there are many studies proposes efficient path selection algorithms for MSs and RSs. There are various metrics that can be used for path selection like hop count, SNR value of link and available bandwidth of access station. Related works on shortcut routing and path selections will be discussed in separate sections. There is not any work on path selection in IEEE 802.16j MMR networks with shortcut routing enabled.

3.1 Shortcut Routing

Hu et. al in [11] proposes local forwarding scheme for OFDMA-based multihop cellular networks. [11] categorizes the communication paths into three as in Figure 3.1. The first one is traffic destined to the node which is not connected to the MR-BS of source node. In this path, all data generated by subscriber will be routed to MR-BS. The second is the path where source and destination node connected to same MR-BS with different access RS stations. In this path data will be first routed to MR-BS and MR-BS forwards it to destination station. The last one is communication path where both source and destination stations are connected to same access RS. Here [11] suggests a forwarding scheme so that data need not to traverse over

MR-BS. Figure 3.1(c) shows paths established from MR-BS to SS1 and SS2. Data generated by SS1 will be first routed to MR-BS on dotted path and then forwarded to path established for SS2. Proposed solution enables RS to forward MAC PDUs directly to SS2 instead of sending to MR-BS.

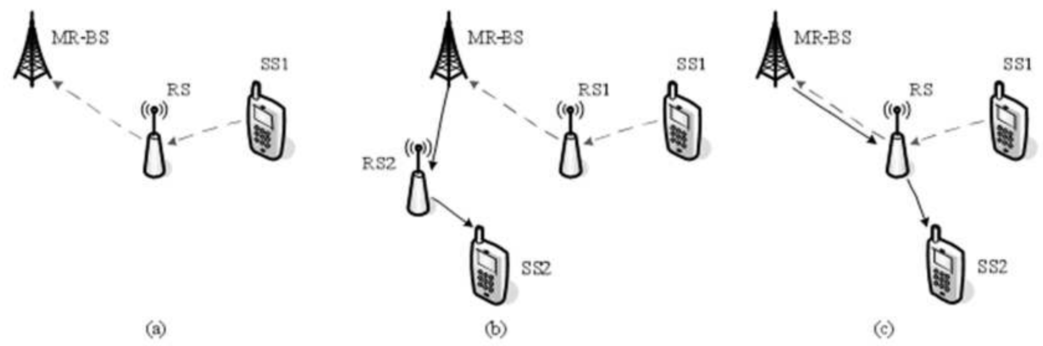

Figure 3.1: Communication path of the MR networks [11]

Local forwarding mechanism decreases the hop counts which decreases latency and increases throughput of network. For local forwarding scheme, [11] modifies the path creation procedure defined in the IEEE 802.16j standard. SS initiates service flow creation procedure by sending DSA-REQ to MR-BS. Upon receiving DSA-REQ MR-BS realizes the destination is connected to same access RS. MR-BS sends DSA-REQ message only address to access RS instead of addressing all the intermediate RSs on the path. When access RS receives DSA-REQ with path information from MR-BS, it responds with DSA-REP. After affirmation, MR-BS responds to MS with DSA-RSP message. Access RS forwards user data on established service flow directly to destination MS. Comparison of regular DSA message flow and flow in local forwarding scheme showed in Figure 3.2.

According to simulation results, throughput increases 12.38% where the probability of %50 of data communication is type of (c) in Figure 3.1.

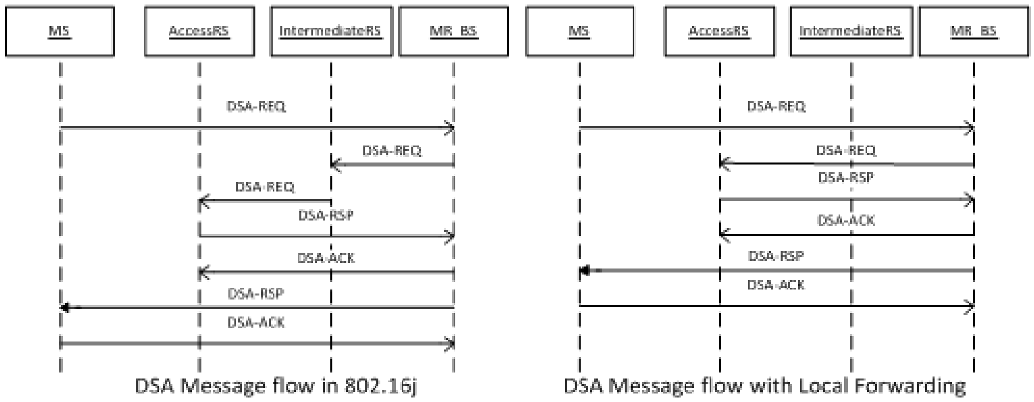

Figure 3.2: DSA message flow in 802.16j and local forwarding

3.2 Path Selection

Path selection in MMR networks, affects latency and throughput of overall system. Various metrics have been proposed in literature for efficient path selection. IEEE 802.16j multihop relay specification does not define path selection method but suggests CINR as a metric. In addition to CINR, Hop count and available bandwidth of access station are common metrics used in literature. Studies focused on path selection for MS and RS separately. Also path selection methods differ according to requirements of subscribers. Ann et. al. proposes path selection method for RSs in IEEE 802.16j MMR Networks in [13]. They use hop count, link available bandwidth and SNR as path metrics. Distributed path selection approach is applied so every RS obtains path metrics from its neighbors. [14] employs similar metrics in different names for selecting paths for MS and RS in centralized manner. Link spectral efficiency and link load are the metric names used in [14]. [15] proposes traffic aware routing algorithm by inspecting the load of intermediate RSs. Since none of the studies concerned about the traffic between the subscribers with common superordinate station, there is no path selection metric according to traffic demands.

When a RS is in the transmit range of more than one RS, it should choose one of them as an access station. In order to give best service to its subscribers, RS should select the suitable RS which has more efficient path to MR-BS. Proposed scheme in [13] employs UCD messages

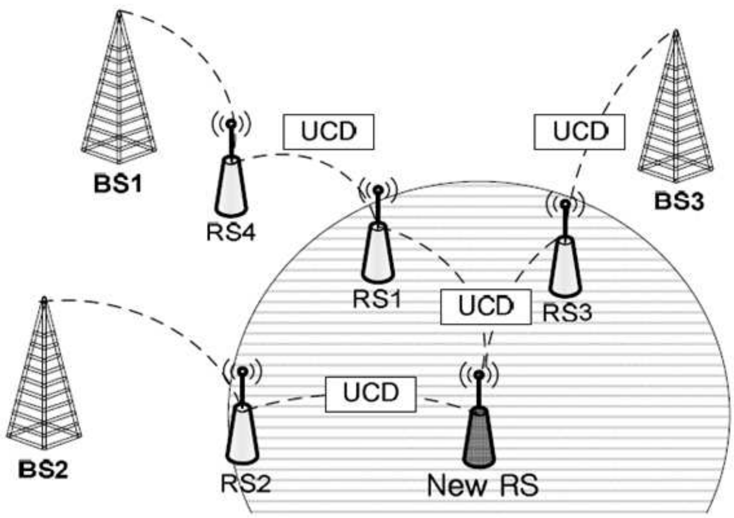

Figure 3.3: Path selection of RS

to publish the path selection parameters. Parameters are added as TLV encoded information into UCD message. In the topology given in Figure 3.3, RS initiates network entry procedure and starts to scan the wireless media. RS obtains neighbor information with received UCD messages. After RS receives the first UCD message, it waits for while to make path selection. The waiting period is duration between two UCD messages originated by same RS. Since UCD message is sent periodically, receiving UCD from sender of first UCD message means all UCD messages have been received from neighbors. RS decides access station with path selection method and starts registration process with sending RNG-REQ to MR-BS. Network entry procedure is described in Figure 3.4.

[13] uses link available bandwidth, channel MCS and hop count as path selection metrics. Aim is to select the best path for increasing the network throughput while decreasing the latency. According to SNR of channel, different MCS level is used. SNR thresholds for different MCS levels are shown in Table 3.1. MCS with link available bandwidth determines the link data rate. Less hop count is preferable while selecting the path because less network resources used and noise ratio is lessen. For the path selection, minimum available link data rate along the path and number of hops the path are used as metrics. Link data rate calculated by multiplying link available bandwidth with MCS level. Past cost formula is

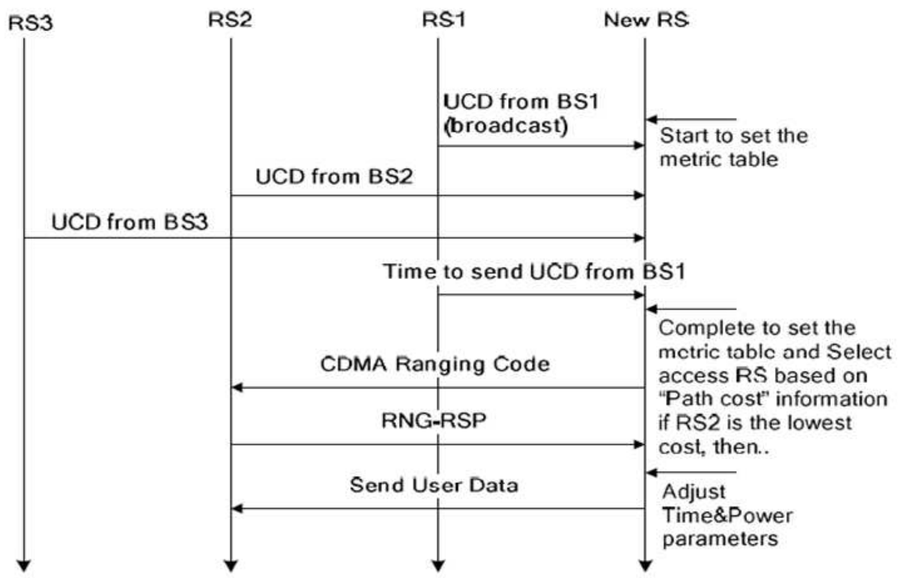

Figure 3.4: Network Entry with path selection

$$P_{r-BS} = L_{r-BS} / H_{r-BS} \quad (1)$$

Where L_{r-BS} is the minimum link data rate for path between relay r and BS and H_{r-BS} is number of hops in the path. New RS selects the RS with maximum path cost.

Table 3.1: SNR thresholds for MCS levels

Assigned Level	Modulation	Coding Rate	Receive SNR Treshold (db)
1	BPSK	1/2	6.4
2	QPSK	1/2	9.4
3	QPSK	3/4	11.2
4	QAM16	1/2	16.4
5	QAM16	3/4	18.2
6	QAM64	1/2	22.7
7	QAM64	3/4	22.4

Link spectral efficiency, proposed metric in [14], means net bitrate divided by the bandwidth in hertz of link. According to the definition, link spectral efficiency (LSE) is amount of data that can be transmitted over given bandwidth of a link. LSE is dependent to modulation and coding scheme used in OFDMA symbol. Modulation and coding scheme is also used in [13] as a metric. Link load is the other metric for path selection in [13]. Every RS is aware of link load between itself and subordinate station. When selecting the path, some of the links may be overloaded due to being a common link of multiple paths. So selecting the RS in accordance to this constraint will improve network throughput. Path selection decision is taken by MR-BS in [14] as defined in the IEEE 802.16j standard. Each RS informs superordinate station about the UL traffic load. MR-BS selects the most appropriate path for new comer MS or RS according to metrics gathered from the subordinate stations.

Chen et.al offer path selection methods which prevent intermediate relay to become bottleneck with heavy load [15]. Each MR-BS or RS has cumulative load of its subordinates. MR-BS obtains the load information with bandwidth request of SSs. First path selection algorithm is minimum average load first (MALF). MR-BS chooses the path with minimum accumulated traffic per hop count. An example of MALF path selection is depicted in Figure 3.5. There are four paths available for new comer subscriber SS2-3. Loads of the RSs and SSs are shown under the node names. According to MALF path costs are 9, 6, 5, and 4 respectively for P1, P2, P3 and P4. MR-BS selects P4 for the SS2-3 as an access station.

Second proposed scheme is traffic aware routing algorithm (TARA). MALF achieves partial load balance. RSs at higher level may still become bottleneck with MALF. MALF does not consider new comer subscribers' load and its effects on nodes at higher level in topology. In TARA, two metrics are used for path selection. One is total traffic load of paths and second is load balance factor. Total traffic load of paths is sum of loads of each intermediate RS on the paths for new subscriber. For each path, possible effect of joining new subscriber to total traffic load is calculated. The path with minimum total traffic load is first criteria for path selection. Load balance factor is average load per link, which is total traffic load divided by number of links in all possible paths. When BS selects the access station of new MS, it checks a path with minimum total traffic load. If there is more than one path satisfying the condition, then the path with minimum load balance factor is chosen. Since load of subordinate RS affects its superordinate RS's load, BS would select the RS with minimum hop count for MS.

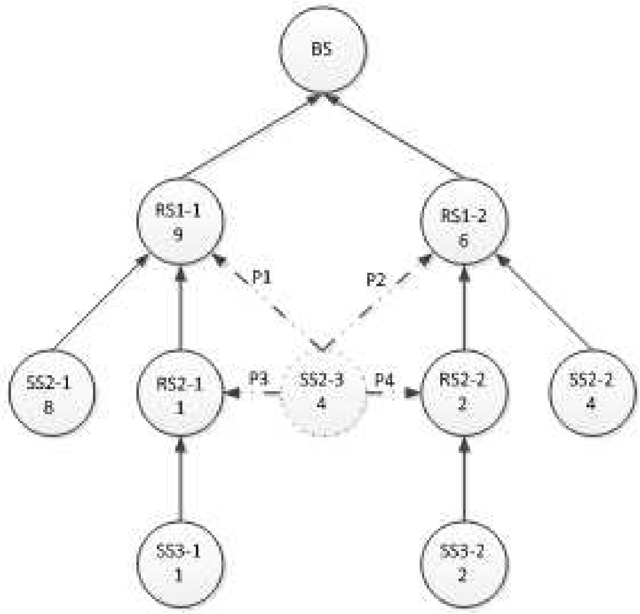

Figure 3.5: An Example of MALF

Therefore latency of the data packets will also decrease.

CHAPTER 4

PROPOSED METHODS FOR TRAFIC AWARE ROUTING IN IEEE 802.16J NON-TRANSPARENT NETWORKS

IEEE 802.16j Multihop Relaying extends the network coverage and increases the network throughput. One of the usage areas of 802.16j is the rural areas where shadowing affect is more likely to occur because of geographic conditions. Considering a tactical area network established in a rural area and using the 802.16j technology, the majority of the whole network traffic is due to the communication among the subscribers in the same region. In such cases, routing all data over the MR-BS for the communicating pairs in the same region is a waste of resources. Moreover, since all the communication paths are established between the MSs and MR-BS, relay links at higher levels of the tree topology may become congested and bottleneck links. For data traffic between subscribers with a common ancestor RS in the tree topology of the network, forwarding packets directly to downlink connection of the destination subscriber (shortcut routing) may be an effective solution that will decrease the congestion on higher levels of the topology and can increase the aggregate network throughput. Selecting a good path between the communicating subscribers may be another effective solution for increasing the overall network throughput.

In IEEE 802.16j, maximum resource utilization is obtained in non-transparent mode with a distributed scheduling. Consider the IEEE 802.16j non-transparent network topology given in Figure 4.1. Imagine that there are communications between MS1-MS2 and MS2-MS3. At the same time, assume all the subscribers are downloading data from the Internet. In this scenario, without shortcut routing, communication path of MS1-MS2 is $MS1 \rightarrow RS2 \rightarrow RS1 \rightarrow MR-BS1 \rightarrow RS1 \rightarrow RS2 \rightarrow MS2$. When shortcut routing is enabled, RS2 would forward the data destined to MS2 directly to MS2 instead of forwarding to RS1. In this case, communication path between MS1 and MS2 will be $MS1 \rightarrow RS2 \rightarrow MS2$. Hence when shortcut routing is used, hop-count of the path decreases from six to two. Moreover, uplink resources of RS1 and MR-BS1 will not be wasted.

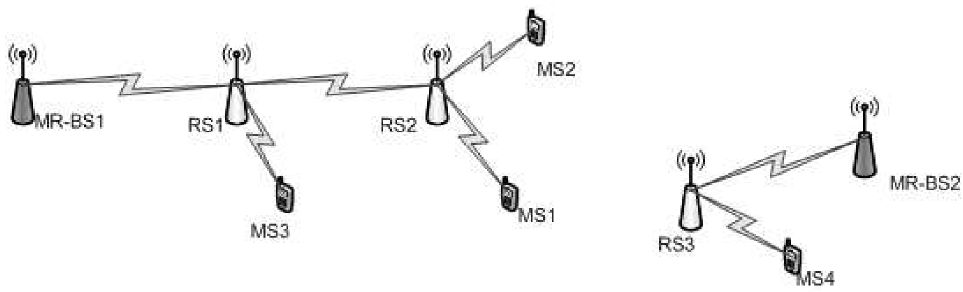

Figure 4.1: IEEE 802.16j Non-transparent network

In Figure 4.1, if the subscriber MS4 initiates a network flow destined to subscriber MS1, data will be sent over the backbone network. However, since MS1 is also inside the coverage of relay station RS3, it would be more appropriate to use RS3 as an access relay station of MS1, instead of RS2. MS1 could make a path selection (hence relay selection) based on its communication demand (which is to MS4) and decide to use RS2, instead of RS2. We propose the use of such a path selection mechanism based on the traffic demands of the mobile stations. Hence, the decision to which RS to connect to will be made by a path selection algorithm that will consider the traffic demands. Taking traffic demand into consideration while selecting the paths and access relays stations may prevent unnecessary traversals of extra relay stations.

4.1 Shortcut Routing

Routing is the responsibility of the Network Layer in IP based networks. In IEEE 802.16j MR Networks, Relay Stations implement the Data Link and Physical layers. Therefore IP based routing is not available in RSs. RSs forward the incoming data directly to either downlink or uplink according to CID. For traffic between two subscribers having a common ancestor RS, forwarding data directly to destination subscriber from the RS may increase network utilization. By keeping the <destination IP address, transport CID> pair in a routing cache, RS can forward data packets by inspecting the destination IP address in the received packets.

Since IP address is IP layer information and CID is MAC layer information, to keep both information together we need an additional layer (module), which we call as CROSS layer, on top of the MAC layer. Hence, we use a cross-layer mechanism to implement short-cut routing at the intermediate relays stations.

In this section, integration of the cross layer on top of the MAC layer of IEEE 802.16j NT RS will be introduced and discussed. This mechanism is our first contribution in this thesis.

4.1.1 Assumptions

Assumptions for our shortcut routing are listed below:

- Only one link layer (MAC) transport connection is established for each MS. The connection is between the MR-BS and MS.
- All service flows are of type UGS (unsolicited grant service).
- RSs run in non-transparent mode with distributed scheduling applied.

4.1.2 Cross Layer for Shortcut Routing

Link layer transport connections are created upon service flow creation request initiated either by MR-BS or MS. Service flow creation procedure is started with DSA-REQ message. Service flows must be associated with a transport connection. This connection may be an existing one or it can be created upon request. MR-BS decides the path for transport connection and informs all the intermediate relay stations about the path. Relay stations keep <CID, link> pairs in a MAC layer routing table. When a data packet is received at a relay station, the relay station checks the CID of the MAC packet and forwards it to corresponding link. An example path creation scenario is shown in Figure 4.2.

In Figure 4.2, the mobile station is the originator of the DSA-REQ message. This message is sent to MR-BS on the management connection. After MR-BS determines the path for the data transport connection, it sends DSA-REQ or DSC-REQ message to the relay stations that are the intermediate stations on the path. Each relay station checks the available resources whether it can satisfy the service demand. If an intermediate relay station accepts the service

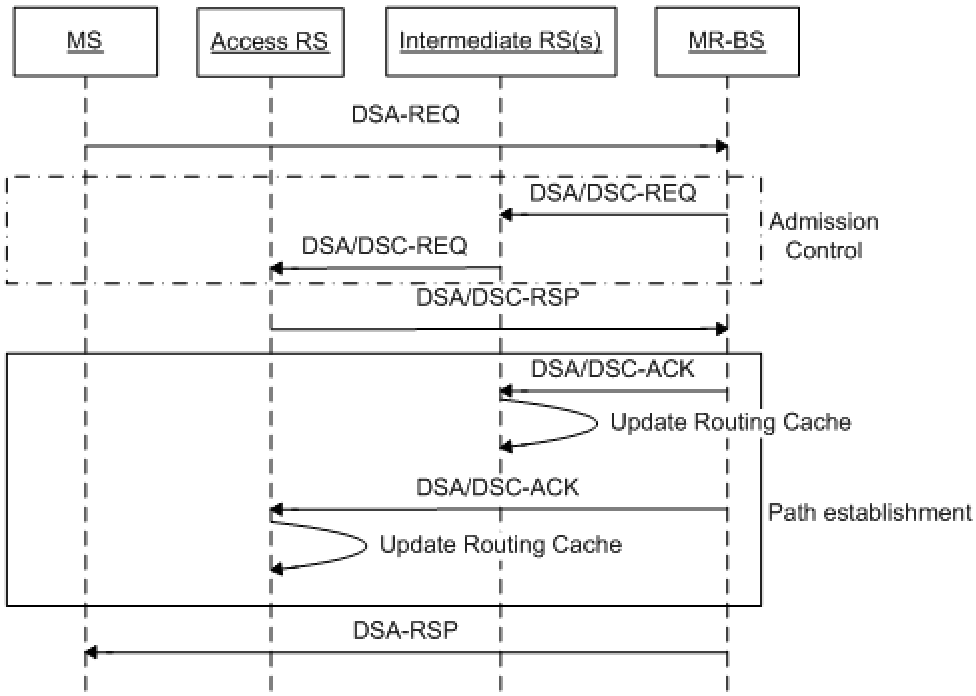

Figure 4.2: Multihop path creation procedure

request, it forwards the request message to the next relay station on the path towards the MS until the message reaches to access relay station that is directly connected to the MS. If one the intermediate relay stations rejects the request, the message will not reach the access relay station and a transport connection will not be created. Upon access relay station receives the DSA-REQ message, it accepts the service request if it has available resources. MR-BS informs the intermediate relay stations about the path created via a DSA/DSC-ACK message. MR-BS also sends a DSA-RSP message to the MS to inform about acceptance of service request by all the relay stations on the path.

Relay stations keep CID and path information in the MAC layer. When a data packet is received from an uplink transport connection, the relay station looks to the CID in the packet and forwards the packet to next hop accordingly. For our shortcut routing, target station's IP address is the key value for making a direct forwarding decision. Hence, not only the CID, but also the IP address inside a received MAC packet is examined by the relay station to make a short-cut routing decision. CROSS layer is inserted on the protocol stack of a relay station as in Figure 4.3 to handle IP based routing operations.

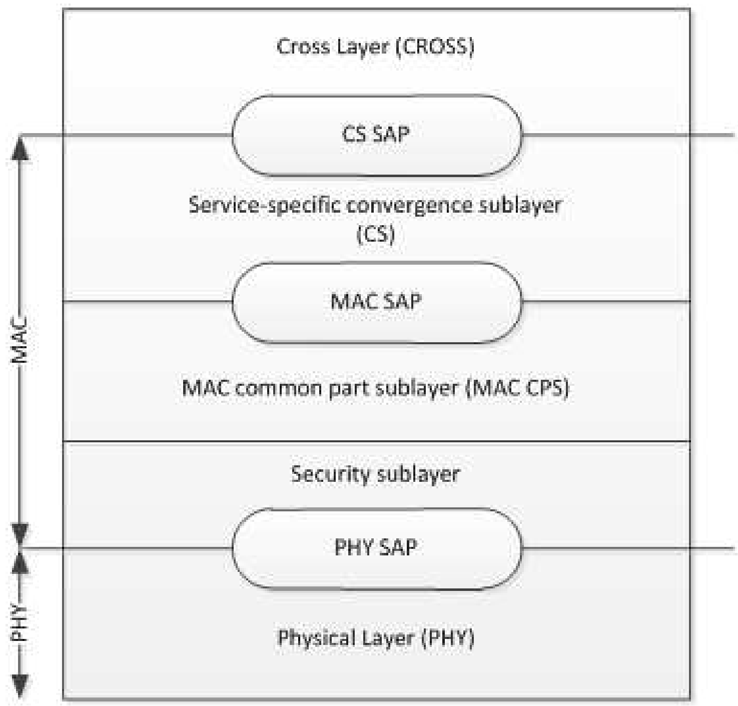

Figure 4.3: Adding Cross Layer to RS protocol stack

The Cross Layer in a relay station keeps <CID, IP> pairs in a routing cache to be used IP based short-cut routing. In our proposal and implementation, the routing cache is a dictionary structure where the key is an IP address and the value is the corresponding CID. Since dictionary structure does not let inserting an item with existing key, only one connection could be associated with given IP address. Routing cache is constructed and updated upon connection creations. As depicted in Figure 4.2, after Access RS accepts a service acquisition request, MR-BS informs intermediate RSs on the path with DSA-ACK which includes path and CID information in TLV encoded message section. Upon intermediate RSs receive this message, the CID and IP address of the MS is sent to the CROSS layer to update the routing cache. For disconnection or handover scenarios of MS, MR-BS sends MS_INFO-DEL message to Access Station to delete MS related information and to release resources allocated to it. This message is sent on the management connection going over intermediate RSs. Intermediate RSs would inspect this message to update their routing caches.

Figure 4.4 shows how shortcut routing decision is taken in our Cross Layer. Shortcut routing mechanism is used only for uplink data (i.e. data traveling from an MS towards the BS). Data packets received from upper layer relay stations (superordinate stations) (i.e. data packets data are originated from BS and traveling to an MS) are forwarded as normal. When a MAC data packet is received from a subordinate station, it is sent to the cross layer instead of finding the destination link from the MAC layer routing table.

For uplink packets, the cross layer first extracts the data packet from the MAC packet. If the packet is an IP packet, then the routing cache is checked whether there is an existing record for the destination IP address. If a record is found, then a new MAC packet constructed with the CID obtained from the cache and pushed into the downlink queue corresponding to that destination connection (to that CID obtained from the cache). If there is no record found for the destination address, no action is taken and the MAC packet is sent back to the MAC layer in its original form. Then the MAC layer forwards the packet to its superordinate station (to a neighboring relay station on the path towards the BS).

4.2 Hierarchical Addressing and Routing for IEEE 802.16j Multihop Relay Networks

Determining the destination of an incoming data packet at a relay station requires checking the routing cache at the cross layer. Handovers and access station changes require reconstructing the routing cache. Using a hierarchical addressing scheme, it can be easier to understand whether a destination node with a hierarchical IP address is a subordinate station of a relay station or not. By this way when a packet is received, it is first checked if the destination IP in the packet is in the address block of the current station. If it is, then the destination CID can be obtained from the routing cache.

For traffic aware routing that we propose for 802.16j network, shortcut routing mechanism was our first step. The main factor that determines the routing path in IEEE 802.16j networks is the path selection method. Location information of a destination node is necessary to make path selection according to traffic demand. By using hierarchical addressing, location of a destination node inside the addressing hierarchy (tree) can be determined by inspecting the destination IP address of a packet.

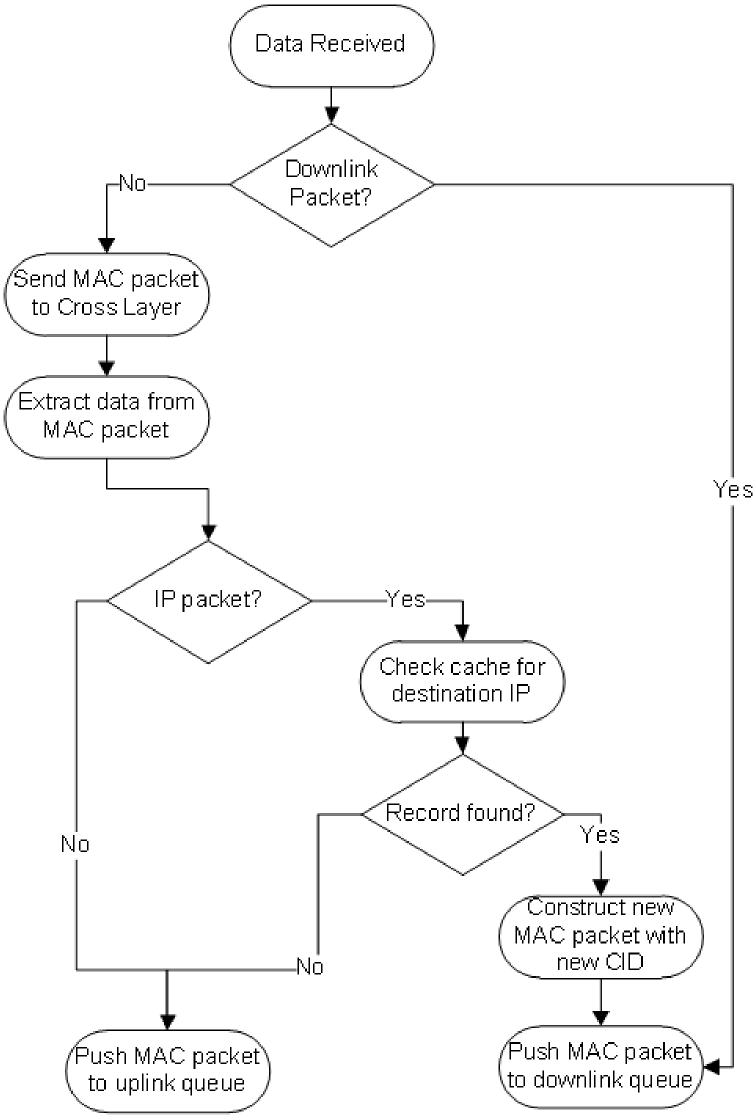

Figure 4.4: Shortcut Routing with Cross Layer

In an IEEE 802.16j MR Network, every intermediate and leaf station has one access station and they are interconnected together to form a tree topology. Position of a node in the tree can be easily determined when hierarchical addressing used. Address assignment is done in a systematic manner so that each subordinate station connected to the same access station will obtain addresses from the same CIDR address block.

In IEEE 802.16j, IP address assignment is handled by MR-BS during RS and MS network entry process. There is not any regulation for IP address assignment in the standard. With hierarchical address assignment, every superordinate station (RSs and BSs) will be considered as a subnetwork. An IP address block will be assigned for each subnetwork. Since both RS and BS are superordinate stations, BS would become a network of subnetworks. The same situation is also valid for RSs which have subordinate RSs. For superordinate stations which have subordinate RSs, IP address block assignment should also contain address block for subordinate RSs. Example tree topology and IP address assignment with CIDR notation is shown in Figure 4.5.

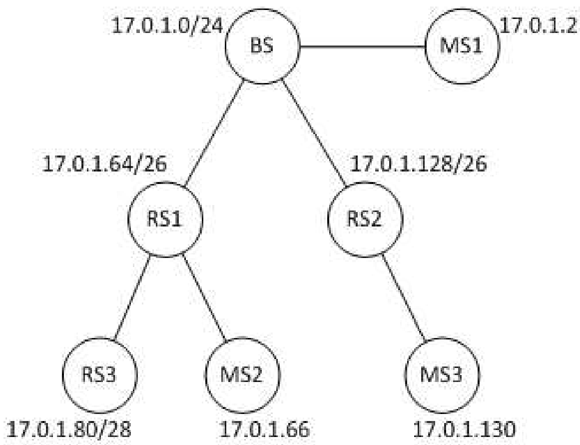

Figure 4.5: Hierarchical IP Address Assignment for 802.16j Network

In the topology shown in Figure 4.5, CIDR block of BS is 24 which means $2^{(32-24)} - 1$ nodes can be addressed with this address block. Since RSs are also superordinate stations, they are assigned an address block in network entry process. Address block assignment is done in a systematic manner. If the subnet mask of an address block is less than 29, address block is divided into four parts. Three parts are reserved for subordinate RSs and remaining part is used for MSs directly connected to the station. In the example topology, BS assigns MS1's IP address from the first part of address block which is 17.2.1.0/26. BS's IP address is the first IP of first part of the IP address block which is 17.0.1.1. Second and third parts of the address block are assigned to RS1 and RS2 respectively.

Systematic assignment of address blocks limits breadth and depth of the topology. Breadth limitation comes from division of each subnet block into four parts, which means a superordinate station can have at most three subordinate RS. CIDR of the assigned address block of the BS limits the depth of the topology. Since smallest CIDR address block that can be assigned to RS is in the form of a.b.c.d/30, depth of the topology will be at most (32 - (subnet length of MR-BS))/2.

Main motivation behind using hierarchical addressing is to be able to gather information about the incoming and outgoing traffic of a station. Examine the mobile station MS2 in Figure 4.1. MS2 is in the coverage of both RS1 and RS2. BS selected RS1 as the access station of MS2 by using a path selection method. However, when MS2 and MS3 start communicating with each other, since MS2 is in the range of RS2, MS2 can understand that the incoming traffic from MS3 will follow a shorter path if the packets from source IP 17.0.1.130 (i.e. from MS3) are passed through RS2. Eventually this gives clue about selecting RS2 as an access station of MS2 and this may increase the overall throughput in the network, and decrease the latency and congestion.

4.3 Enhanced Mobile IP for Shortcut Routing Enabled IEEE 802.16j Networks

Handling mobility is a challenging issue in mobile networks. Seamless handover and addressing mobile nodes are key issues for uninterrupted communication. IEEE 802.16 specification provides a solution for handover and describes the procedure in details. For IP based networks, addressing of mobile nodes are handled with Mobile IP protocol which is also applicable in WiMAX networks.

As stated in the background section, if Mobile IP is used, mobile nodes have two addresses which are ”home address” and ”care-of address”. Foreign and home agents run on MR-BSs in traditional WiMAX networks. When a packet is received at an RS, if the destination address belongs to one of the subordinates of RS, RS forwards the packet directly to the destination

MS. If MS performs handover operation, then the packets forwarded from the RS will be lost. To prevent this situation, every RS should have home and foreign agent functionality. When a subordinate station initiates handover or a path change process, the requests pass through the intermediate stations. These request packets are processed at intermediate RSs to update MobileIP tables, and then forwarded to MR-BS. This solution prevents packet losses caused by migrated nodes. Also, when a subscriber initiates registration process with new BS, requests and responses pass through the intermediate stations, so all intermediate stations will know about the address of new comer.

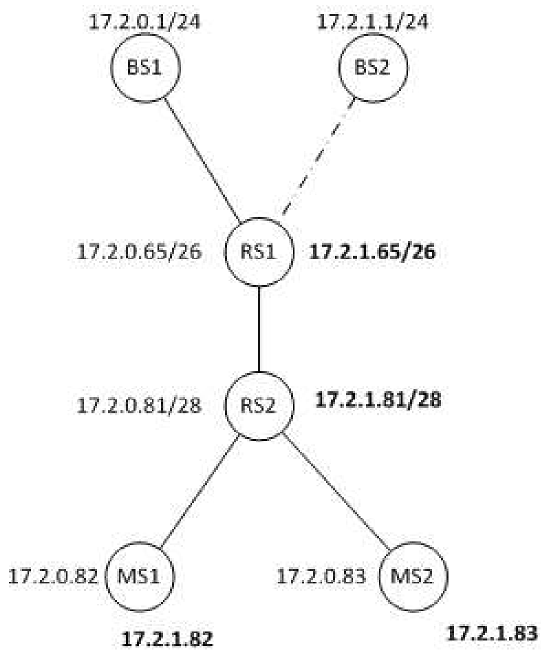

Figure 4.6: Care of Address Assignment after Access Station Change

Subordinates may change the access station because of mobility or to get better service from another superordinate station. These two access station change scenarios are valid for both RS and MS. For RSs the access station change task is more complicated since they might have subordinate stations. When an RS with subordinate stations changes the access station, all subordinates connected to the RS should renew their IP addresses. Renewing the subordinate stations' IP addresses is the responsibility of the superordinate RS. In Figure 4.6, RS1 changes its access station from BS1 to BS2. While connecting to BS1, MobileIP procedure works and BS2 sends request to BS1 for assigning new care of address for RS1. BS1 responds, and BS2 accepts the registration of RS1 and assigns its new care of address. For the completeness of hierarchy in IP addresses, RS2 and its subordinates should have changed their IP addresses. RS1 assigns new addresses to its first level subordinates. Then each RS is responsible for assigning IP addresses to its first level subordinates. In this case, RS1 assigns the IP address of RS2 and RS2 assigns the IP addresses of MS1 and MS2. Since the proposed IP address assignment method limits the depth and breadth of topology, some nodes may be disconnected. Disconnected MSs should connect to another superordinate station in the vicinity if available. Path selection method can prevent having disconnected nodes after access station change.

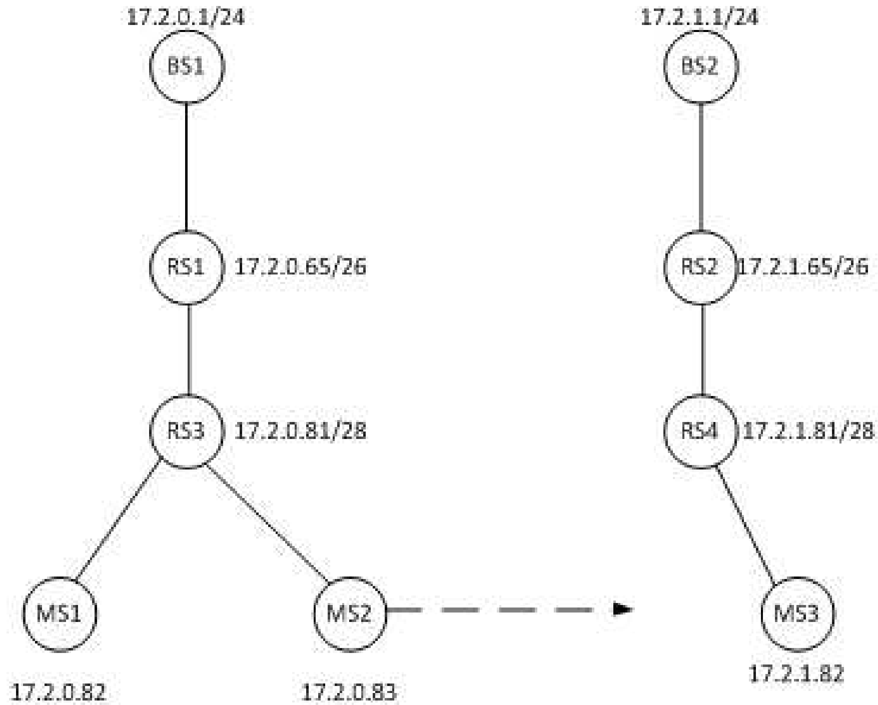

Figure 4.7: IP Address Assignment before Handover

Figure 4.7 and Figure 4.8 illustrate how agents work in the proposed solution. MS2 leaves the network where BS1 is the gateway station to the backbone network, and MS2 moves to the network of BS2. Every node has a home address which is taken from the first network joined. In the figures, current care of addresses of the stations are written. Before leaving the BS1's network, MS2 initiates handover request with candidate neighbor station. If the handover request is accepted by BS1, MS2 initiates the registration request to BS2. Upon receiving registration request of MS2, BS2 inserts MS2 into visitors' list record and requests BS1 to join MS2 its network with MS2's new care of address. If BS1 accepts BS2's request, it keeps the mobility information of MS2 and respond back to BS2. Each intermediate RSs which receives MS2 handover request and its response, will know that MS is leaving the network. If home agent address of MS2 is declared to be the IP address of RS3, then MobileIP request will be destined to RS3. So RS1 and BS1 indirectly learn about the care of address of MS2. Each intermediate RS receiving MS2 registration request and its response will know about MS2's IP address and keep its care of address.

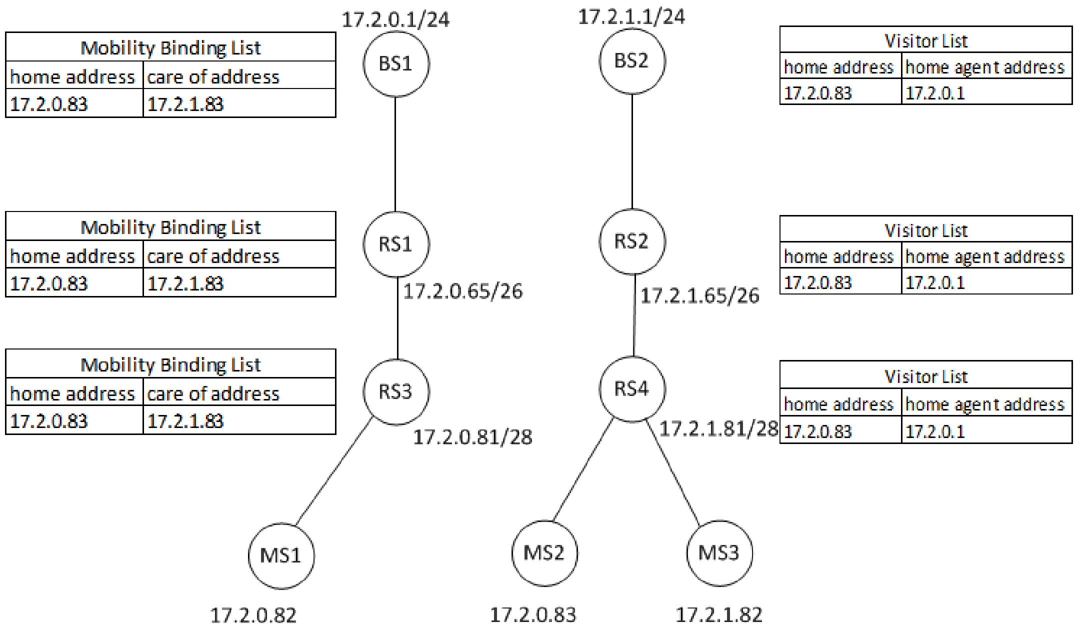

Figure 4.8: Mobility Binding List and Visitor List after handover

In the scenario examined in Figure 4.7 and Figure 4.8, if MS2 has an ongoing communication with many other nodes, communication should not be disturbed because of handover. For outgoing traffic there will not be any problem since the source addresses of the packets originating from MS2 have the home address of the MS2. For incoming data to MS2, there are two cases to consider in an intermediate node: 1) a node transferring data with shortcut routing, 2) a node transferring data over MR-BS. For the first case, since MS2 sends handover request to BS1 before leaving the network, all the intermediate stations along the path learn about the handover. Between handover request and registration to BS2, shortcut routing is valid for MS2. Since during registration, BS2 sends request to RS3 about MobileIP assignment of MS2, RS3 and its superordinates learn the care of address of MS2. Every RS and BS updates the mobility binding list when request from BS2 is received. After RS3 accepts the request of BS2, MS2 will be registered. After this point, when a packet is destined to MS2's home address, if the sender is outside the network of BS1, then the packet will be routed to the care of address of MS2 from BS1. If the source of the packet is inside the BS1's network, this means source node and MS2 had at least one common superordinate station. Shortcut routing should forward the packet to previous access station of MS2 upon receiving such a packet. However, since there is a mobility record for MS2, receiving RS directly forwards the incoming packet to its access station as if in a traditional IEEE 802.16j network. At the end, BS1 sends the packet to MS2 over the backbone network.

4.4 Traffic Aware Path Selection for Shortcut Routing Enabled IEEE 802.16j Multihop Relay Networks

Path selection has an important role for utilizing network resources efficiently. Path selection is the process of selecting a routing path between two mobile stations or between a mobile station and the base station. There are various metrics that can be used in a path selection method. According to QoS requirements of network traffic, path selection metrics and methodologies may change. For shortcut routing enabled IEEE 802.16j Networks, we consider traffic demand as an important criterion for path selection to find the shorter routes for carrying the traffic. Without knowing the traffic demand and destinations, it is hard to select an appropriate access station that will provide shorter path for a communication. We propose the monitoring and use of the traffic demand (amount of traffic flowing) between any two mobile stations to select the best access relay stations and routing path for those mobile stations. In our proposal, the current traffic demand between any two stations is predicted according to the traffic statistics gathering from the previous and recent data transmissions between the two stations. Hence, received and sent data statistics between any two stations are used as the path selection metrics in our proposed method to select a good path between the stations.

Without shortcut routing, all MS traffic would pass through MR-BS. In such a case, path selection aims to find the best routing path between MS and MR-BS. There are various metrics proposed for this problem which increases throughput and decreases congestion and latency. These metrics would be used in shortcut routing enabled networks as well, since traffic destined to outside of the network will follow a routing path towards the MR-BS. Our focus in this study, however, is topologies and traffic scenarios where the main data traffic is among the subscribers in a limited region, like a tactical area. In this case, with short-cut routing, not all traffic from a mobile stations have to go up to the base station, but can be routed to the destination mobile stations using shorter paths by just using intermediate relay stations and not going up to the base station.

4.4.1 Path Selection Metrics

Metrics used for path selection aim to utilize the network resources efficiently. Shortcut routing avoids unnecessary packet traversals for the traffic to flow inside the network. So the problem is finding an optimum path which does not degrade the end to end throughput of MS and BS, at the same time increase the chance of shortcut routing.

Modulation and Coding Scheme (MCS)

WiMAX supports adaptive modulation and coding which means according to link quality, modulation and coding technique used can be changed dynamically. MCS is chosen according to SNR of a link. If channel is noisy then choosing modulation technique with lower bitrate and coding scheme with higher redundancy is more appropriate. There are seven MCS levels used as in [13]. Table 3.1 shows SNR and MCS levels.

Given a bandwidth for channel, selecting MCS level as high as possible increases the throughput. If the bandwidth allocated to a user is 100 MHz and the selected MCS is BPSK 1/2, then the data rate of the user will be 100 * 1/2 = 50 Kbps. Whereas if QAM64 3/4 is chosen, then it will be 100 * 6 * 3/4 = 450 Kbps. If a routing path consists of more than one link, the link with the smallest MCS level will become a bottleneck. So, while selecting a routing path, link qualities for all links on the path should be taken into consideration.

For the shortcut routing case, even if the selected path between an MS and BS has poor quality links, communication between MSs with a common RS may not be affected. Consider the

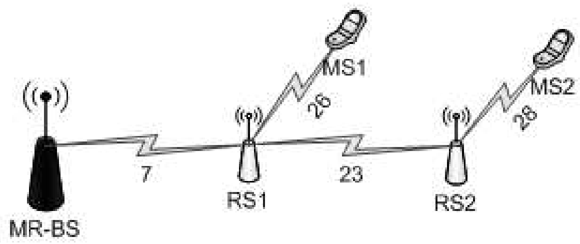

Figure 4.9: MCS metric

SNR values of the links in the topology given in Figure 4.9. MS1 and MS2 are connected to access relay stations with high quality links and QAM64 3/4 is selected as MCS. The link between RS1 and RS2 has QAM64 1/2 as MCS. The bottleneck link in this topology is the link between RS1 and MR-BS, which is the most necessary link for traffic going to or coming from the outside of the network (i.e., external traffic). MCS level for the bottleneck link is BPSK 1/2. Since, when shortcut routing is applied, the traffic flowing between MS1 and MS2 does not use the link between RS1 and MR-BS, the minimum MCS for routing path between MS1 and MS2 is QAM64 1/2.

Hop Count

Hop count metric is the number of links between MR-BS and MS. Hop count affects the latency and load of intermediate RSs. For traffic scenarios which use shortcut routing, hop count between MS and MR-BS is not meaningful.

Load

Load of the intermediate RSs gives clue about congestion and resource availability. In uplink direction, there is bandwidth reservation, so there is no problem for subscribers. However, in downlink direction, there is no bandwidth reservation, so heavily loaded RSs become congested. To avoid congestions and packet drops; selecting RS that are congested less is preferable.

Received and Sent Traffic Statistics

Previous three metrics are used for selecting a good path between an MS and the BS. By using these metrics, however, it is not possible to handle shortcut routing scenarios. In some cases subscribers may connect to an access station which has more hop count to MR-BS, or may connect to heavily loaded superordinate stations, or may use a relay link with poor quality. Even in this case, the total network throughput may be optimal. Communicating with neighboring MSs rather than with the nodes outside the network may generate such scenarios.

4.4.2 Integration with IEEE 802.16j Standard

In IEEE 802.16j networks, MR-BS is responsible for selecting the access stations of RSs and MSs. The standard describes the access station selection method for RSs during network entry phase. For MSs, changing access station is considered as handover. Therefore even if MS does not change its MR-BS, handover process is applied for access station change. Either MR-BS or MS itself can take decision of handover. Consequently, proposed path selection method will run in MSs and MR-BSs.

Metrics that are used for path selection should be collected from network. In literature, available management messages of 802.16j are used for this purpose. As in [13], we use UCD messages to gather the metrics from network for MS. MR-BS obtains neighbor information from RS_NBR_MEAS-REP sent by RSs. We propose that our traffic demand metrics (sent and received traffic) to be carried in these messages as well. There are two different path selection scenarios which are path selection in network entry and path selection in operational mode. In the first scenario, MS/RS determines the access station and initiates a ranging process. In the second scenario, MR-BS or MS initiates a handover process. For an RS, MR-BS requests the RS to change its access station.

RS and MS path selection during network entry

MS starts scanning the air interface during network entry phase to obtain neighbor stations' channel descriptors. MS path selection procedure is shown in Figure 4.10. MS interprets the information obtained from UCD messages with channel measurements to calculate the metrics. MS waits for a second UCD message received from same MR-BS to finish scanning. This duration is enough for receiving all UCDs sent from neighbor stations. MS starts selecting its access station to connect based on the metrics it obtained in the scanning phase. MS starts a ranging process with the selected RS as described in the standard. Registration of MS to MR-BS may affect the metrics for intermediate RSs. MR-BS may change the access stations of the intermediate RS during a path creation phase as described in previous sections.

The same procedure is valid for an RS for its path selection during its network entry phase.

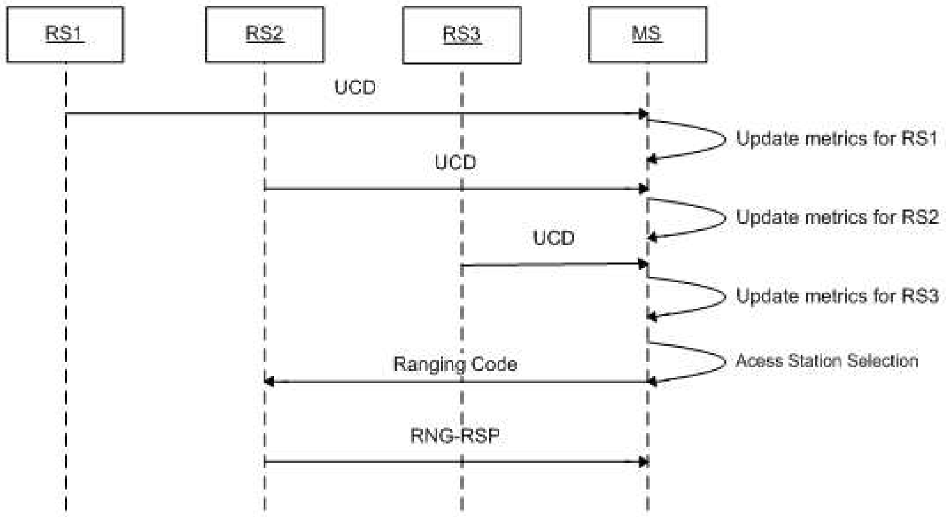

Figure 4.10: Path Selection of MS at Network Entry

MS access station change in operational state

Due to the mobility or service requirements, MS may change its access station while in operational state. MR-BS has knowledge about the network with the information obtained from the backbone network and the RSs. An MS or MR-BS may initiate a handover process for providing better service for the MS. Path selection method is used to take the handover decision. Example path selection of an operational MS is shown in Figure 4.11. The MS requests scanning interval from MR-BS with MOB_SCN-REQ. If MR-BS accepts the requests, MS starts scanning the air interface in the interval specified by MR-BS. After obtaining information from UCD messages received from neighbor stations, MS starts a path selection procedure and selects its candidate access station. MS sends MOB_MSHO-REQ to initiate handover with the candidate access station. If MR-BS accepts, MS starts a registration procedure as described in the previous chapters. After MS disconnects from the previous access station, MR-BS informs the RS with MS_INFO-DEL message.

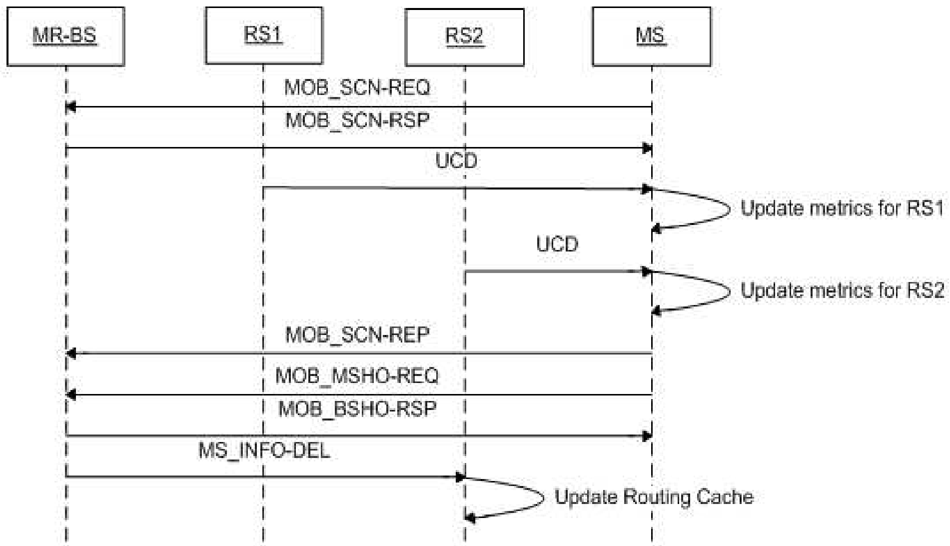

Figure 4.11: MS access station change in operational state

RS access station change in operational state

RSs periodically collect neighbor information to send reports to MR-BS. MR-BS determines access stations of RSs in operational state. Figure 4.12 depicts access station change for RS3. RS2 is the initial access station of RS3. Due the change in the metrics, MR-BS selects RS2 as a new access station of RS3. MR-BS informs RS3 about new access station with RS-AccessRS_REQ message. Upon receiving AccessRS_REQ message, RS3 sends generic ACK to MR-BS and starts network entry procedure by sending RNG-REQ message to RS1. The rest of the procedure works as defined in the standard.

4.4.3 Path Selection Method

Each RS and MS keeps statistics about the source of its incoming traffic and destination of its outgoing traffic. Statistics are kept in a dictionary where the key is a neighbor superordinate station and value is the amount of data transferred (i.e., total incoming and outgoing traffic of the MS or RS). For send and receive cases, two dictionaries are kept. When some amount of data is sent to a destination node, it is checked whether the destination IP address is inside

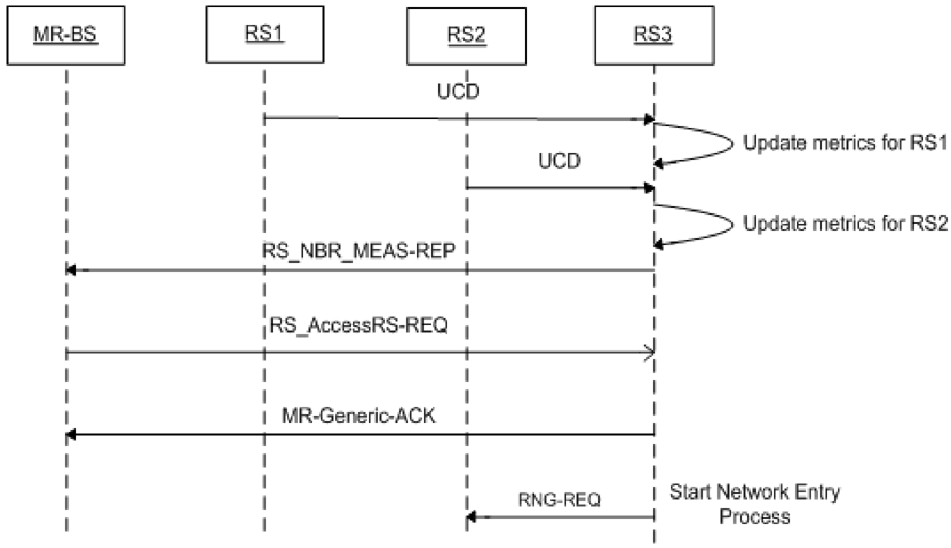

Figure 4.12: Access station change for operational RS

one of the address blocks of the superordinate stations. Similarly, when some amount of data (for example, one data packet) is received from the access station, the source IP address of the data is checked whether it is a subordinate of the neighbor stations.

Figure 4.13 explains the usage scenario of ”sent & received traffic” metric. In the given topology, MSs are connected to RSs and communicating with each other. MS7 enters the region which is the intersection of RS2’s and RS4’s transmit range. Since MS7 is new in the region, there should not be any traffic statistics in its dictionary for superordinate stations in the range. MS7 creates dictionary entries for the new neighboring RSs and selects of them as the access station according the previous three metrics. According to the previous three metrics, hop count and MCS metrics are the same for both of the paths going through RS2 and RS4. However, load of RS4 is more than RS2 because of the number of subordinates connected to them. Therefore, MS7 selects RS2 as its access station at the beginning. After a while later, MS7’s major traffic becomes the traffic to/from MS5, MS4 and MS6. When traffic load prevails the other metrics, then MS7 selects RS4 its an access station to benefit from shortcut routing.

Each subordinate station calculates alternative path costs periodically. Alternative paths may

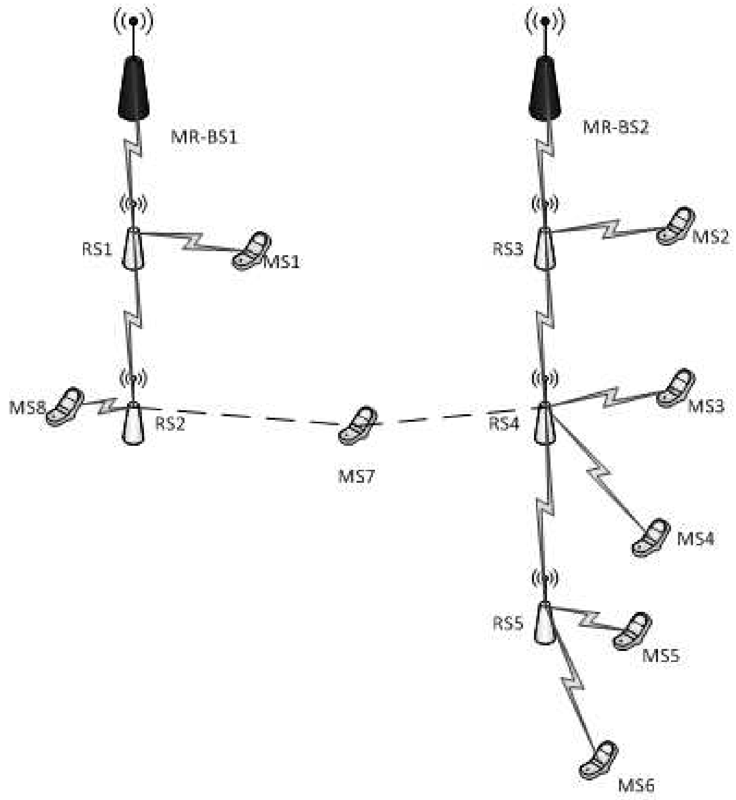

Figure 4.13: Sent and Receive Data Metric Usage Scenario

end up with same MR-BS or not. If the rating alternate paths prevails that the current and the alternate path ends up with different MR-BSs, the subordinate station starts a handover process. If the subordinate station is an RS, then station migrates to the new subnetwork with all its subordinates. If the prevailing path is also ends up with the current BS, then RS sends SBC-REQ to BS with rates of the candidate stations. Extra information will be carried with TLV encoded part of the management packets.

After selecting a new path for a subordinate, statistical data analysis may change in favor of the previous path. This could happen repetitively if the costs of paths are nearly same. To prevent this situation (i.e., oscillations between alternative paths), to change the path, candidate path must be better than the current path by more than a threshold value.

Metrics defined in this section was grouped into two. First group of metrics are for selecting the best path to deliver packets between MS and MR-BS. The second group of metrics is for selecting paths which are more suitable for shortcut routing scenarios. Second group of metrics is meaningful for a scenario where the subscribers registered to the same MR-BS are heavily communicating with each other. If majority of the traffic is flowing inside the network (i.e., internal traffic, then the second group of metrics becomes more important. As a result, in our path rating function (i.e., path cost function), every metric has a coefficient. These coefficients may change according to the usage scenarios. Every metric value is scaled to a floating point number between 0 and 1 to normalize the weight of each metric. Our path cost formula is given in Equation 2.

$$P_{r-BS} = c_{MCS}.MCS_{ms-BS} + c_H.H_{ms-BS} + c_L.L_r + c_S.S_r + c_R.R_r \quad (2)$$

where c_{MCS} coefficient is for MCS metric, MCS_{ms-BS} is minimum MCS level on the path between MS and MR-BS. H_{ms-BS} is hop count between MS and MR-BS and c_H is the coefficient of the metric. L_r is the load of the candidate access relay r. S_r and R_r are sent and received data statistics to and from Relay r.

Declarations

B_{max}	:	Maximum available UL Bandwidth for an RS or MR-BS.
MCS_{max}	:	Maximum assignable MCS level for a link.
$CIDR_{max}$	:	Maximum assigned CIDR block in the network (MR-BS's CIDR block).
c_{MCS}	:	Coefficient for MCS metric.
c_{hop}	:	Coefficient for hop count metric.
c_{load}	:	Coefficient for load metric.
c_{sent}	:	Coefficient for sent traffic metric.
c_{recv}	:	Coefficient for received traffic metric.
N	:	Neighbor RS and MR-BS stations.
t	:	Path rate threshold multiplier.

Algorithm of path selection method is given in Algorithm 1. This algorithm runs of each MS and MR-BS periodically. Complexity of the algorithm is $O(n)$ for MS where n is the number of neighbor stations. For MR-BS complexity of algorithm is $O(n^2)$ since for every subordinate station, MR-BS runs path selection algorithm.

Algorithm 1 $path\_select(N)$

$max\_rate \leftarrow rate\_station(curr\_access)$

$candidate\_access \leftarrow curr\_access$

for all $S \in N$ **do**

  if $rate\_station(S) > max\_rate \times tres$ **then**

    $max\_rate \leftarrow rate\_station(S)$

    $candidate\_access \leftarrow S$

  end if

end for

if $candidate\_access \neq curr\_access$ **then**

  $connect\_to(candidate\_access)$

end if

Algorithm 2 $rate\_station(S)$

$aval\_bw\_ratio \leftarrow S.aval\_bw/B_{max}$

$load\_ratio \leftarrow 1/Log_{100}S.Congestion$

$mcs\_ratio \leftarrow MCS(S)/MCS_{max}$

$hop\_ratio \leftarrow (32 - S.IpAddressBlock.Subnet)/CIDR_{max}$

$sent\_ratio \leftarrow SentAnalysis[S]/SentAnalysis.Total$

$recv\_ratio \leftarrow RecvAnalysis[S]/RecvAnalysis.Total$

$rate \leftarrow c_{MCS} \times mcs\_ratio + c_{hop} \times hop\_ratio + c_{sent} \times sent\_ratio + c_{recv} \times recv\_ratio + c_{load} \times load\_ratio$

return $rate$

CHAPTER 5

SIMULATION RESULTS

This chapter describes the simulation environment used and gives the simulation results. Various network simulators were studied to simulate the work proposed in this thesis. First part of the work done, simulated in NCTUns and rest of the work simulated on WiMAX Multihop Relay Simulator which is developed by author. In simulations, end-to-end throughput, queue lengths of superordinate stations and latency of network packets are measured. Chapter consists of three sections;in first section simulation environment will be discussed with their capabilities and simulation models. Topologies that are simulated and simulation results are mentioned in Section 2. Comments on simulation results are given in last section.

5.1 Simulation Environment

Simulation environment is crucial to evaluate performance of the proposed solution. There are various network simulators which have implemented commonly used network protocols. Since IEEE 802.16j amendment released at late of 2009, many of the simulators do not have official support. WiMAX Forum gives support to ns2 simulator and there is only IEEE Std 802.16-2009 plug-in available. OPNET has also have WiMAX specialized model but gives support to IEEE 802.16-2004 and IEEE 802.16e-2005. The only network simulator that gives IEEE 802.16j support is NCTUns 6.0. Currently NCTUns have full support for transparent 802.16j. Non transparent 802.16j module of NCTUns 6.0 is still under development.

5.1.1 NCTUns 6.0

NCTUns is a powerful simulation tool that runs on Linux. Two key features distinguish it from other well known simulators. One of them is, it uses operating systems TCP/IP stack. Second one is, since it uses OS TCP/IP stack, any real world network applications can be used on NCTUns for testing. NCTUns provides GUI for constructing topologies and traffic models. After simulation completed, with playback option data flow could be seen.

To enable use of Linux's TCP/IP stack, kernel reentering methodology is used. Pseudo device interfaces named tunnel interface, are used each node in simulation. Every tunnel interface is recognized as a regular network interface by operating system. Figure 5.1 depicts the communication of two hosts in NCTUns simulation. Simulation engine manages only the links between two host and rest of the TCP/IP protocol details are handled with kernel functions.

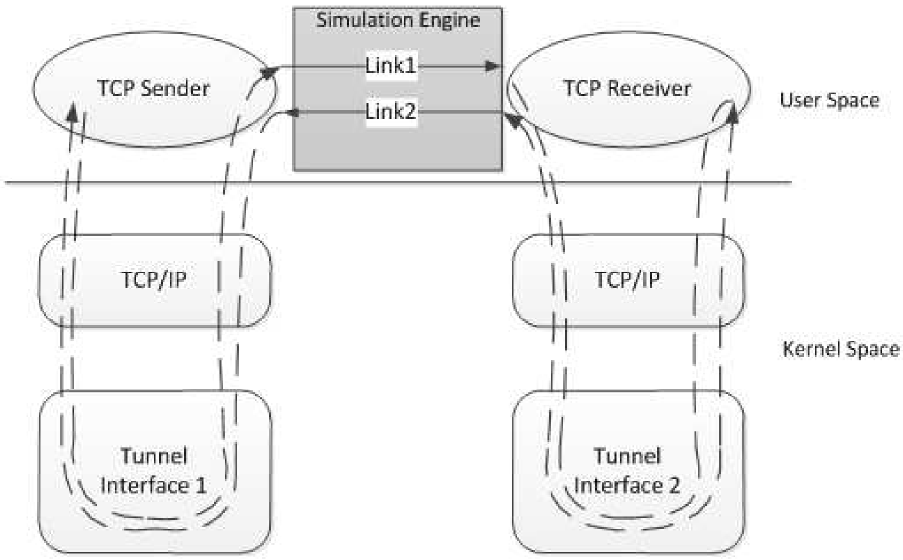

Figure 5.1: Kernel Reentering Methodology

NCTUns has a modular architecture. Each layer in protocol stack has a corresponding module in NCTUns. Each module may have one precessedor and one successor. Modules communicate with each other with common interfaces. Figure 5.2 shows how two TCP hosts connected via switch on NCTUns. Modules have send() and recv() interfaces for communicating modules on up and down. For calling upper layer module, recv() interface is used and for calling lower layer module send() interface is used. In Figure 5.2 flow of a packet from Host1 to Host2 is shown. Top layer is inteface module which is visible to application layer and the bottom layer is link layer where the packets forwareded to link layer of the next hop . Here Host1's link layer forwards the packet to Switch's link layer by calling recv() interface and packet propagated up to switch module. Since the purpose of switch is forwarding the packet right destination, switch module selects the next interface and sends the packet down. Packet is pushed down by calling send() interfaces and pushed up by calling recv() interface in Host2.

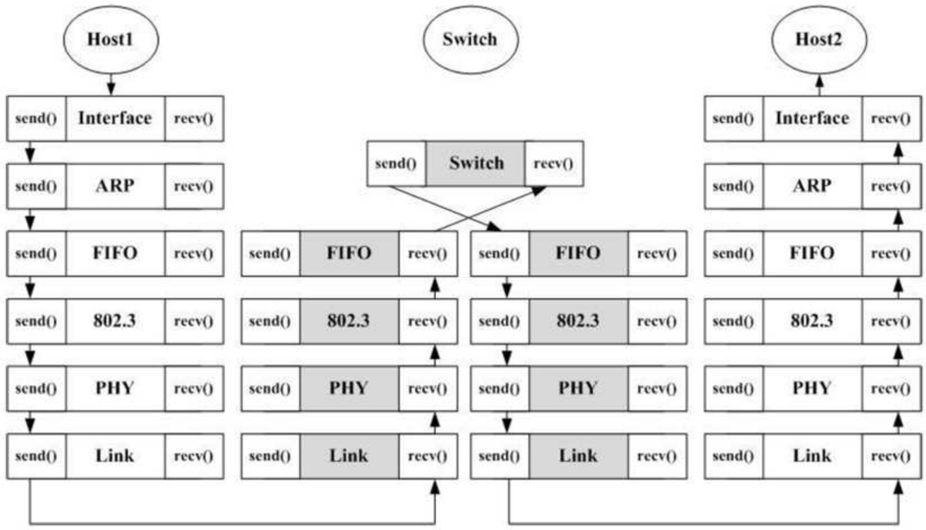

Figure 5.2: NCTUns Module Based Platform [17]

5.1.1.1 Protocol Stacks of IEEE 802.16j Non-transparent Mode Networks in NCTUns

A device can have more than one interface in NCTUns like Switch. MR-BS has also two interfaces which are IEEE 802.16j interface and Ethernet interface. Protocols stacks of MR-BS, RS and MS are shown in Figure 5.3. Ethernet interface of MR-BS is for communicating with backbone wired network. Responsibilities of MAC and PHY modules of MR-BS, RS and MS are different so for each device type separate modules are designed.

MR-BS is connected to backhaul network via Ethernet interface. This connection should have enough bandwidth to carry the whole IEEE 802.16j network traffic. MAC802_16J_NT_PMP_BS module performs MAC operations like scheduling, connection management. OFDMA_PMPBS_MR module performs physical layer operations which use OFDMA technology. CM module simulates the channel model like signal power attenuation, shadowing and multi-path fading effects [17]. MAC802_16J_NT_PMP_RS module performs scheduling and relaying functionalities. OFDMA_PMPRS_MR module encodes and decodes the data transferred to MR-BS and MSs. MAC802_16J_NT_PMP_MS also performs MAC functionalities including sending receiving message from MR-BS and RS.

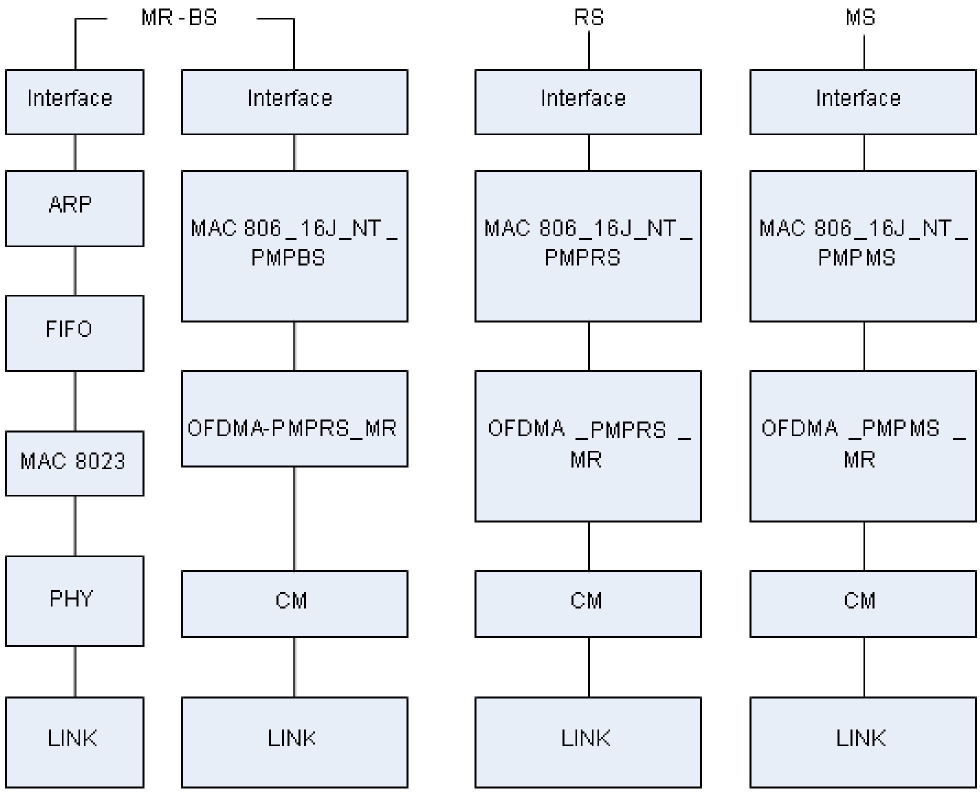

Figure 5.3: Protocol Stacks of IEEE 802.16j Non-transparent Mode in NCTUns

Main functionalities that are implemented in MAC layer are the initial ranging procedure, the network entry procedure, management message negotiation, network management, connection management and packet scheduling. In initial ranging procedure, MR-BS, RS and MSs synchronize with each other to decode received frames. After successful ranging process, during network entry MR-BS assigns CIDs to RSs and MSs attached to itself. Two management connections and one transport connection are established for each subordinate. Details of RS and BS packet scheduling are shown in Figure 5.4.

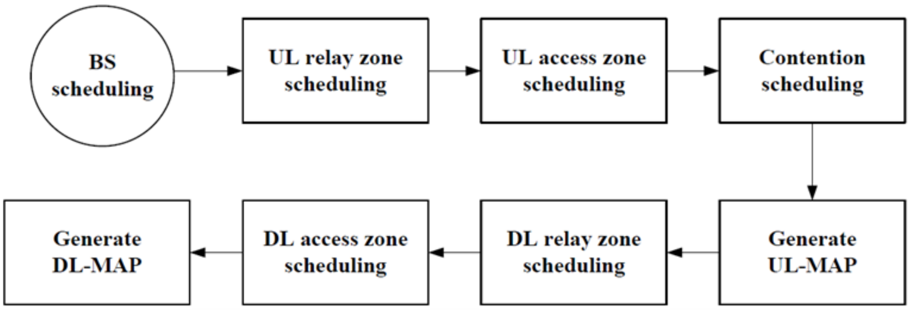

Figure 5.4: Procedure of packet scheduling in superordinate stations [17]

5.1.1.2 Modification for Shortcut Routing

For proposed shortcut routing mechanism, only MAC802_16J_NT_PMP_RS module is modified. Since the mechanism is routes the packets according to the IP address, it is not appropriate to do this operation in MAC module. Therefore a new module is attached on top of MAC802_16J_NT_PMP_RS as in Figure 5.5.

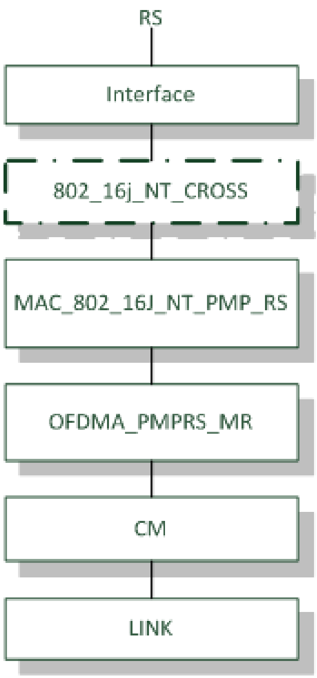

Figure 5.5: Protocol Stack of RS after modification

As stated in the previous chapter, responsibility of 802_16J_NT_CROSS module is to keep IP-CID mappings for shortcut routing. Before modification MAC_802_16j_NT_PMP_RS was forwarding data packets received from subordinate station to access link. In opposite direction, RS was checking for the CID of the MAC header, and then forwards the packet to link associated with CID. With modification when a RS receives a data packet, it is sent to 802_16J_NT_CROSS module with CID and UL/DL information. 802_16J_NT_CROSS module modifies the MAC header if shortcut routing possible or directly sends down packet as described in Figure 4.4.

5.1.2 WiMAX Multihop Relay Simulator

NCTUns is a powerful tool but have some limitations. IEEE 802.16j_NT module of NCTUns has not released yet. Some of the crucial features of IEEE 802.16j MR networks do not exist. NCTUns 6.0 does not support mobility, multihop connection more than two hops. Automatic frequency selection does not work in PHY module so RSs and BSs interfere with each other. Consequently a simulator which has stated features is needed to verify the proposed model. A discrete time event simulator is developed for simulating multihop relaying and burst based network traffic. Simulator is composed of two main parts which are simulation and topology editor. Topology editor is used for creating, viewing and modifying IEEE 802.16j network topologies. Network traffic description is also generated with this tool. Generated topology and traffic description are passed to simulation engine as parameter. During simulation, connection and modification of wireless links can be viewed. Simulation results are printed on standard output. Simulation model is shown in Figure 5.6.

Simulation logic of SimulationEngine is given in Figure 5.7. Each station has interfaces for sending and receiving burst. Stations keep outgoing burst in a queue for each link. All BSs have same total UL bandwidth during simulation. MSs request UL bandwidth before connecting a superordinate. If granted UL bandwidth requests reach to total UL bandwidth, BS does not accept connection requests. Link quality is determined according to distance between source and destination. Adaptive Modulation and Coding Scheme selects MCS level based on link quality. For an allocated bandwidth if higher level of MCS is chosen, then bitrate of link increases. If the one of the relay links of in a path have lower MCS level than burst queue of that link may become congested. To evaluate the performance of proposed

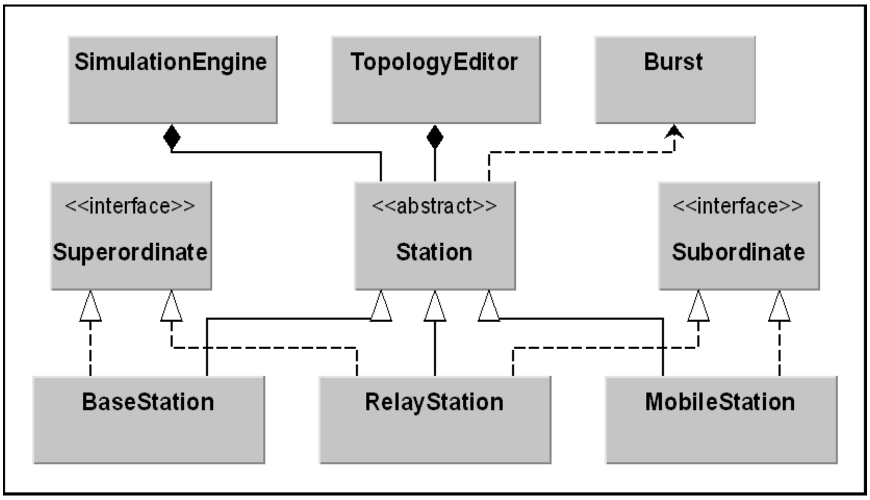

Figure 5.6: Simulation Model

algorithms and methods, throughput and congestion are used.

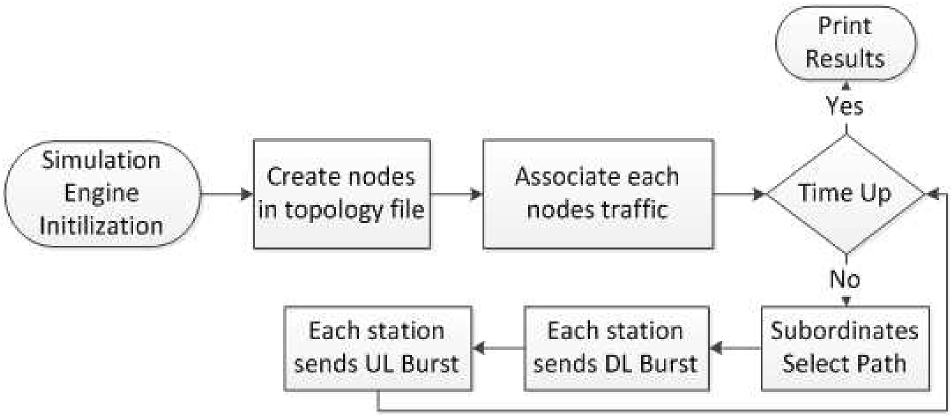

Figure 5.7: Discrete Time Simulation Logic

5.2 Topologies and Simulation Results

In this section various topologies and traffic scenarios are tested in NCTUns and WiMAX Multihop Relay simulator. Topologies are selected for both indicating the difference in algorithms and real world conditions. As stated in previous sections, NCTUns has limitations for IEEE 802.16j Non-transparent mode networks. For this reason complex topologies cannot be tested on NCTUns. For NCTUns the only metric for evaluating the performance is total throughput of the network. In WiMAX Multihop Relay simulator latency and queue length of RSs are also used for evaluation.

In all simulations data traffics are started at the beginning of the simulation and ended when simulations are ended. Throughput is calculated as sum of total data received at destinations during simulations. Uplink and downlink queues of RSs are infinite so packet drops do not occur. Queue lengths are used as indication for congestion on nodes. Hop count and queuing delays affect the latency in Multihop Relay Simulator.

5.2.1 Scenario 1

This scenario indicates how shortcut routing affects latency. Simulation was run for a short duration and total throughput evaluated. MS1 and MS2 generate 1Mbps TCP traffic to each other. Without shortcut routing each packet is first delivered to MR-BS which means each packet traverses one hop more.

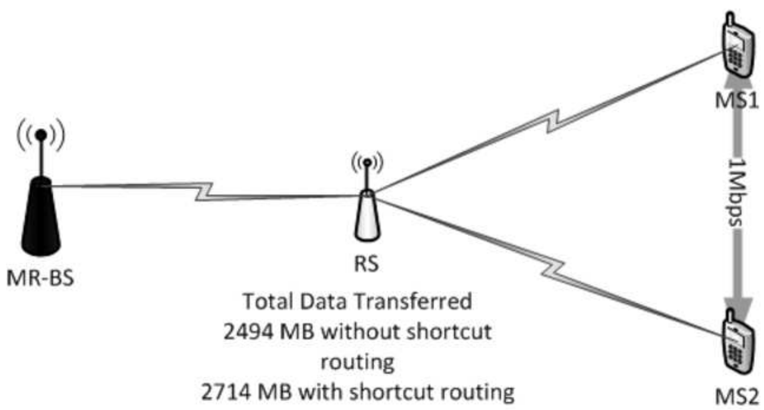

Figure 5.8: Topology and Traffic description for scenario 1

5.2.2 Scenario 2

In NCTUns 6.0 only supported QoS class is UGS. MR-BS allocates bandwidth for each MS upon request. In this case if all the subscribers communicating with each other with shortcut routing, the only benefit will be decrease in latency. However, for downlink traffic, there is no bandwidth reservation for MSs. This scenario show, with shortcut routing downlink queue of MR-BS will be less congested and network throughput of system will increase. Shortcut routing performs 30% better than actual routing mechanism.

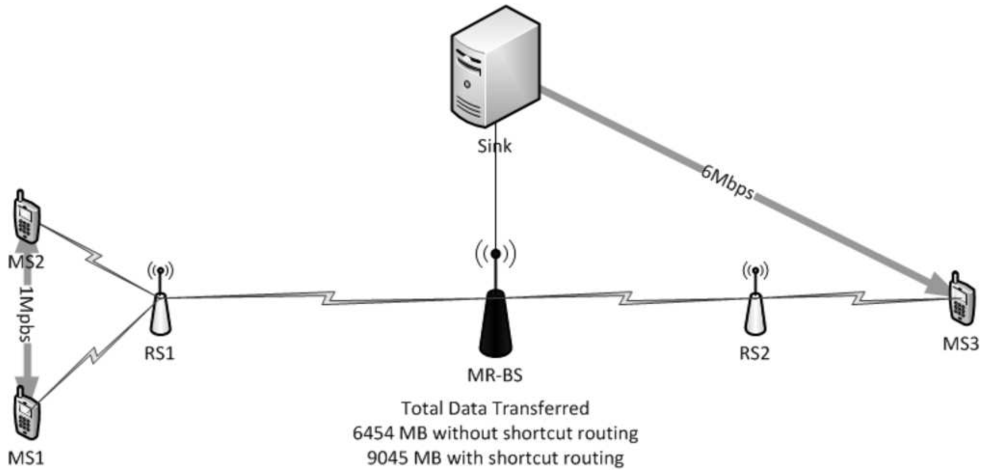

Figure 5.9: Topology and Traffic description for scenario 2

5.2.3 Scenario 3

Motivation behind the shortcut routing is to decrease the load of superordinate stations at higher level. Since IEEE 802.16 Non-transparent network forms a tree topology, nodes at higher level becomes heavy loaded. This scenario indicates how decreasing the load of MR-BS increases the throughput.

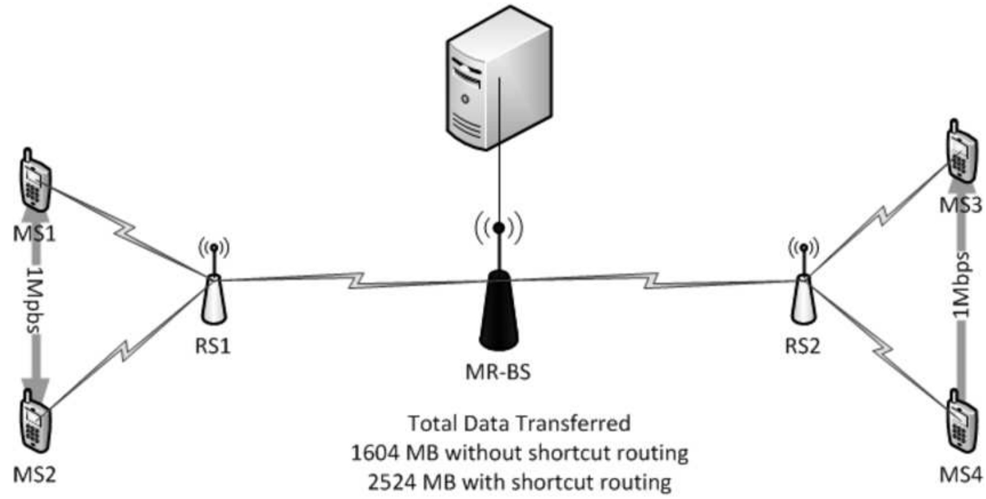

Figure 5.10: Topology and Traffic description for scenario 3

5.2.4 Scenario 4

Previous scenarios were designed for observing the effects of shortcut routing. The scenario below has hybrid traffic types where some of them are suitable for shortcut routing and some of are not. Only traffic MS2 to MS1 and MS4 to MS3 will be forwarded from RSs.

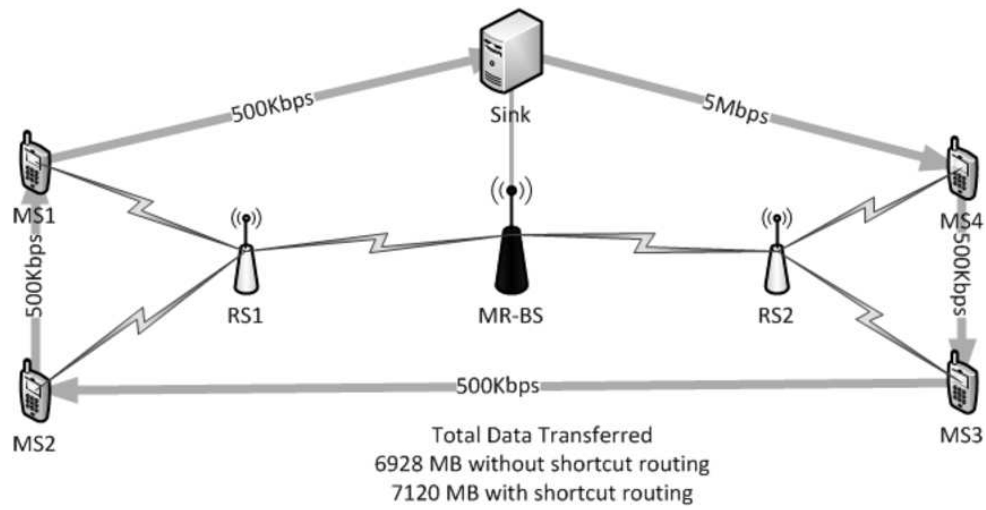

Figure 5.11: Topology and Traffic description for scenario 4

7Mbps TCP traffic is generated in overall network and 1Mbps of this traffic is suitable for shortcut routing. The simulation results do not differ so much since the major traffic route is same in both cases.

5.2.5 Scenario 5

This scenario aims to indicate benefit of proposed metric in path selection. MS1 and MS2 generate TCP traffic destined to each other. MCS and Hop Count metrics directs path selection method to connect directly to base station. At the beginning of the simulation MS1 selects MR-BS1 as access station and MS2 selects MR-BS2. After a while MS1 and MS2 sent and received traffic analysis metric dominates other so that MS1 and MS2 choose RS as access station. Simulation results without traffic analysis metric and without traffic analysis metric are shown in Figure 5.12 and Figure 5.13 respectively.

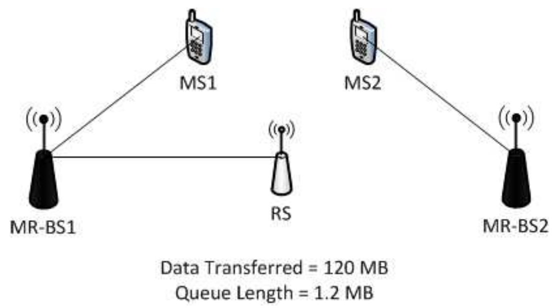

Figure 5.12: Path selection results with traditional metrics for scenario 5

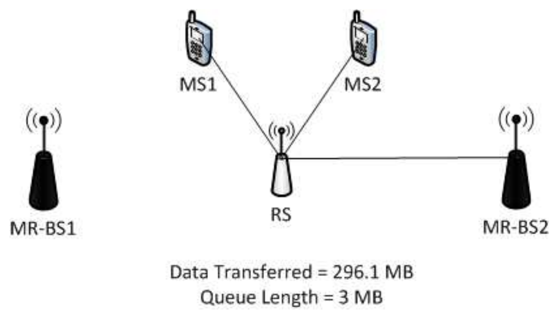

Figure 5.13: Path selection results with traffic aware metrics for scenario 5

5.2.6 Scenario 6

As in Scenario 5, this scenario also depicts the benefits of traffic aware routing algorithm. Compared to previous scenario, there are more MSs deployed and some of them have longer distance to RS in this scenario. Path selection results with traditional metrics are shown in Figure 5.14. Traffic aware path selection result is shown in Figure 5.15. One of MS's access station is not changed because selecting RS3 as access station may degrade access link quality. Sending traffic over backbone network may be more efficient. Throughput improvement is lower than previous scenario since MCS levels of access links of MS which are far away from RS are low.

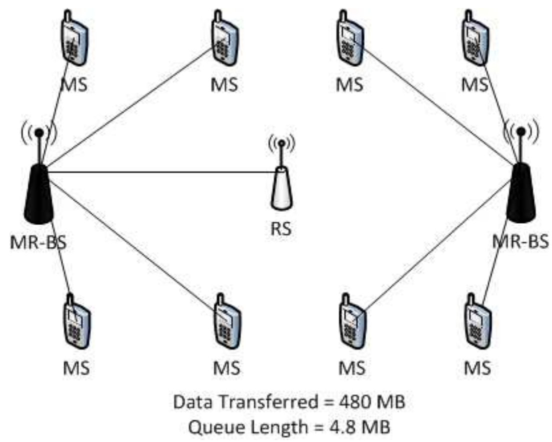

Figure 5.14: Path selection results with traditional metrics for scenario 6

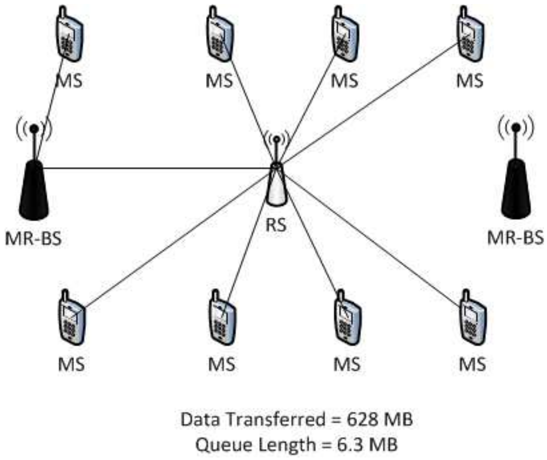

Figure 5.15: Path selection results with traffic aware for scenario 6

5.2.7 Scenario 7

Since NCTUns is not mature simulator to test complex topologies, random and complex topologies are tested on author's simulator. A base topology (Figure 5.16) is constructed with dense RS deployment to put forward the Multihop Relaying concept. Topology is logically divided into four regions to distribute MSs uniformly. Each region has one MR-BS for serving subordinate stations.

A uniform random distribution of 80 MSs is shown in Figure 5.17. Traffic scenario is semi-random where the probability of destination being in the same region of source node can be declared. By this way, how communication ratio of the nodes in same region affects the proposed path selection method and shortcut routing can be observed. Each MS selects two traffic destinations and has total 300 KHz UL bandwidth. Total UL bandwidths of MR-BSs are 7000 KHz and total DL bandwidth of RSs and MR-BSs are 8000 KHz.

Proposed path selection method uses four metrics as stated in previous chapter. Each metrics coefficient is fixed with best values obtained from different simulation runs. Utilization of dedicated bandwidth to each MS is increased in different simulation runs. With low utilization, intermediate RSs will not become congested even if its access link quality is lower than its subordinate's access link. For observing the behavior of proposed methods on local communication, probability of choosing traffic destination in same region is increased

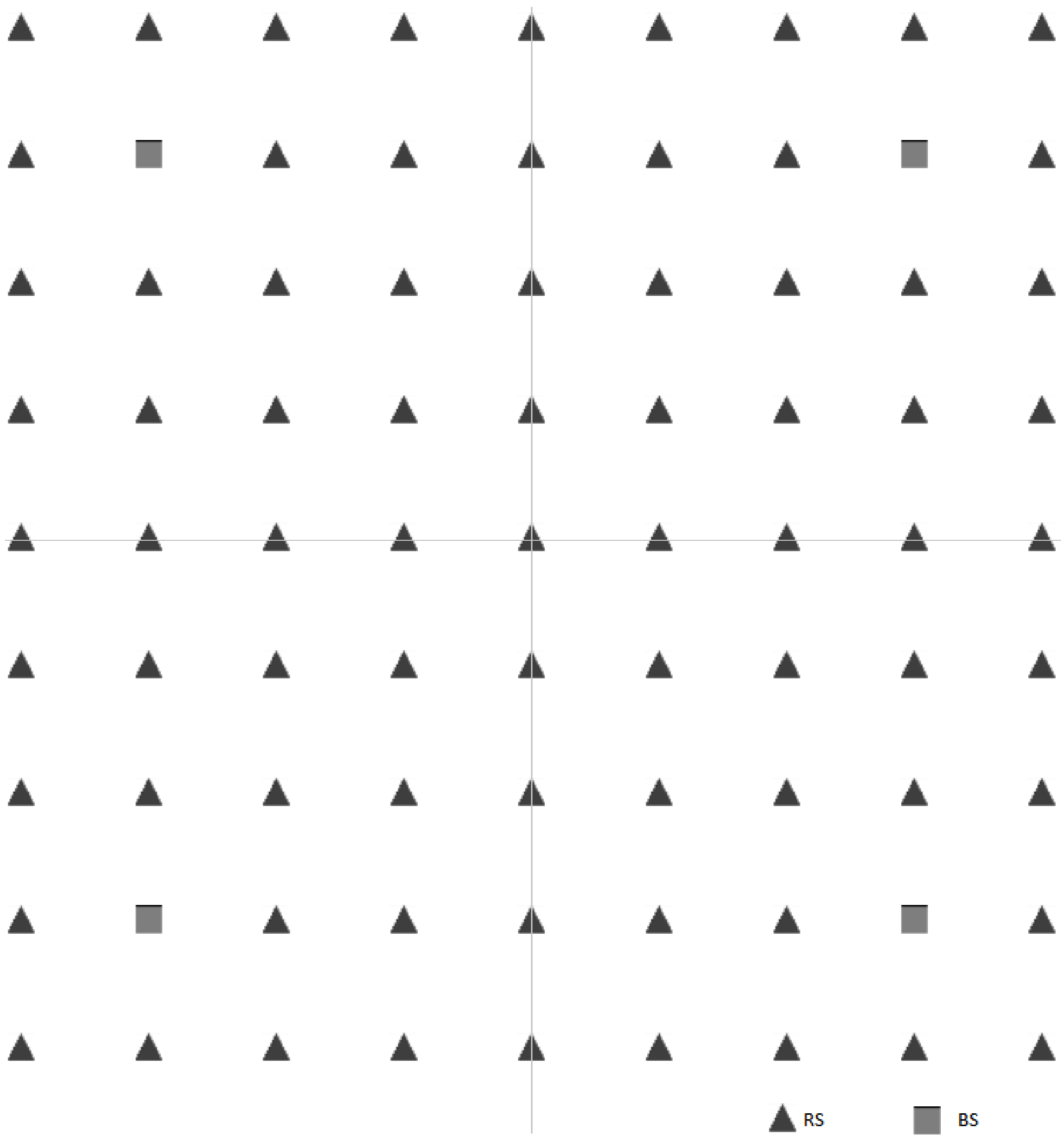

Figure 5.16: Base Topology

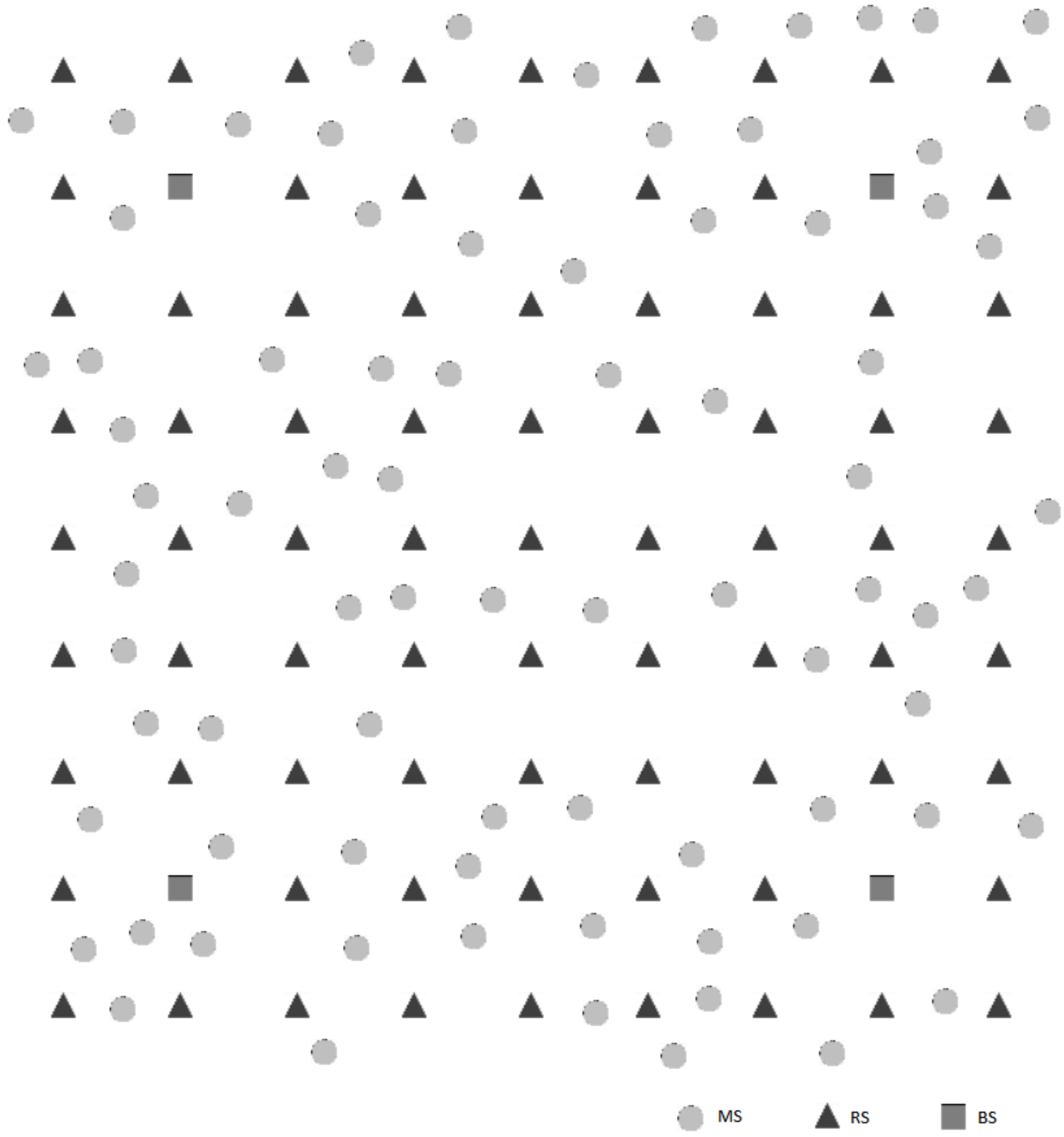

Figure 5.17: Uniformly Distributed MSs on Base Topology

in different simulation runs. The results of simulations are show on following graphs. The abbreviations that are used in graphs, is given in Table 5.1.

Table 5.1: Abbreviations for Simulation Results of Different Methods

Abbreviation	Description
T1	Throughput of standard routing and path selection
Q1	Queue Length of standard routing and path selection
L1	Latency of standard routing and path selection
T2	Throughput of shortcut routing and standard path selection
Q2	Queue Length of shortcut routing and standard path selection
L2	Latency of shortcut routing and standard path selection
T3	Throughput of shortcut routing and traffic aware path selection
Q3	Queue Length of shortcut routing and traffic aware path selection
L3	Latency of shortcut routing and traffic aware path selection

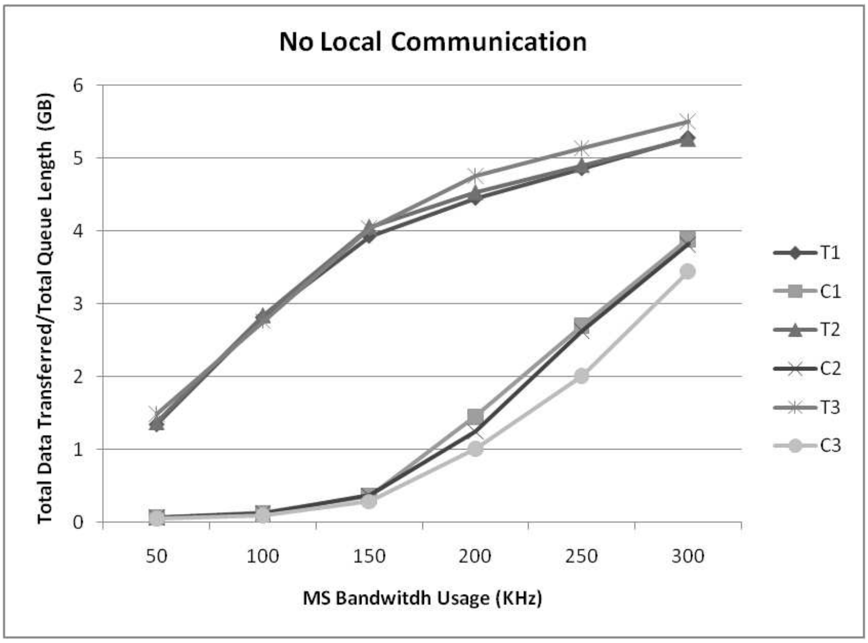

Figure 5.18: Simulation results for all traffic destinations are outside the region

For the scenario where there is no traffic between subscribers in the same region, our expectation is all algorithms should produce similar results. Since there is no shortcut routing opportunity, routing paths are almost same for all schemes. As shown in Figure 5.18 throughput, queue length of RSs and average latency of a packet are almost same for all algorithms. There is slight improvement for traffic aware routing algorithm. Average latency

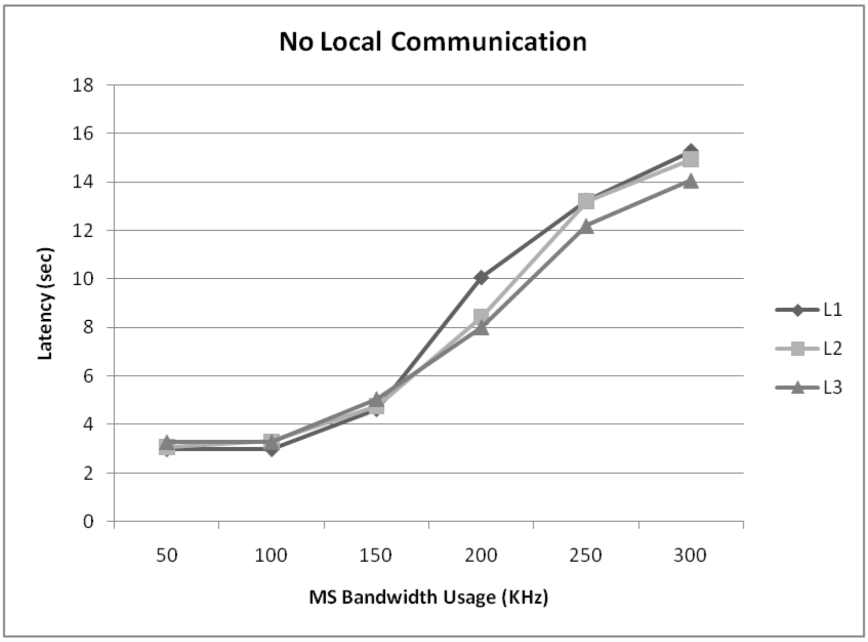

Figure 5.19: Average latency of a packet where probability of selecting destination from same region is 0

of a packet is also similar for all algorithms as shown in Figure 5.19. For the traffic between subscriber from neighbor regions, subscribers may select access station from neighbor RS. Consequently source and destination nodes are connected to same network where shortcut routing possibility exists.

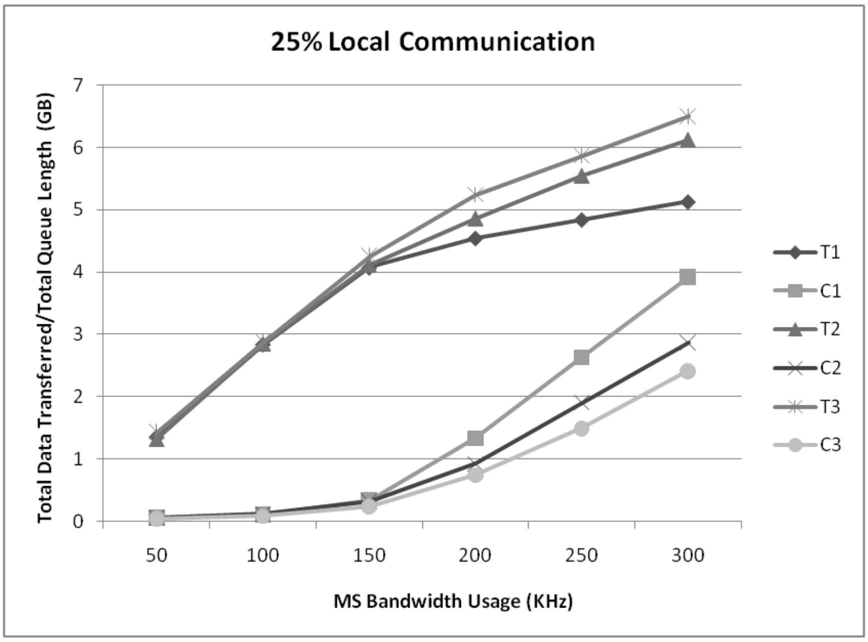

Figure 5.20: Simulation results for probability of traffic destination being in same region with source is 0.25

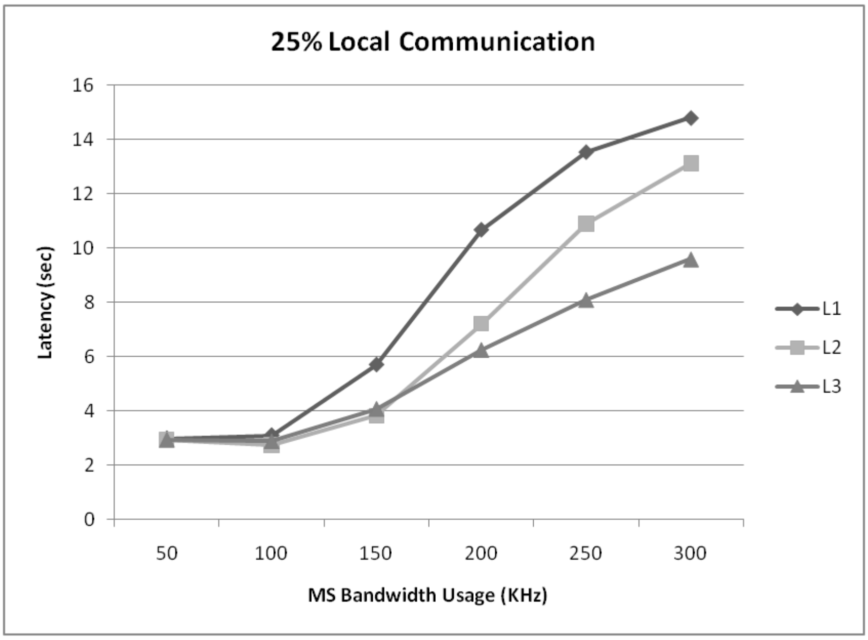

Figure 5.21: Average latency of a packet where probability of selecting destination node from same region is 0.25

Increasing the probability of choosing traffic destination from the same region, throughput of network improves for the shortcut routing enabled routing schemes. According to Figure 5.20, throughput and queue length of standard routing algorithm remains same. Performances of algorithms are almost equal when channel utilization is under 50%. When channel fully utilized, enabling shortcut routing increase throughput 20% and with using sent & received data statistics metric throughput improvement reaches 25%. Latency of a packet increases as incresing the utilization of dedicated bandwidth. Using shortcut routing and traffic aware path selection decreases latency in higher utilization levels as shown in Figure 5.21.

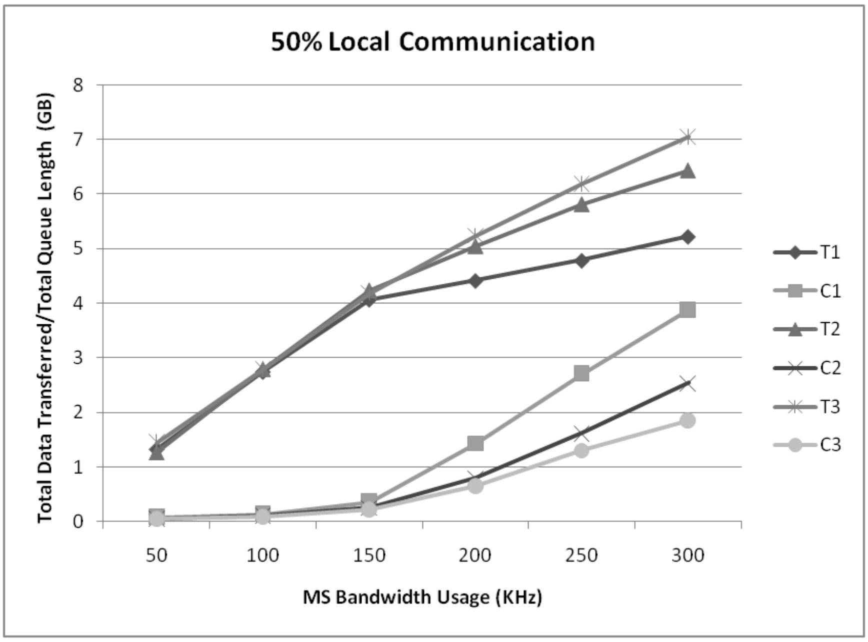

Figure 5.22: Simulation results for probability of traffic destination being in same region with source is 0.5

Increasing probability of traffic destination being in the same region with source node to 0.5, throughput is 28% improved with traffic aware routing for higher channel utilization. Moreover, latency is 50% decreased compared to standard algorithm.

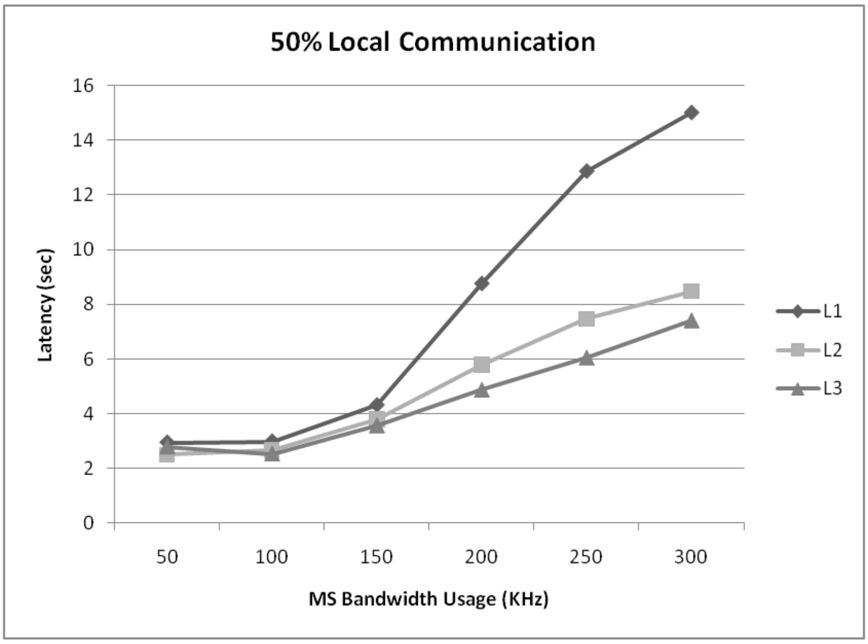

Figure 5.23: Average latency of a packet where probability of selecting destination from same region is 0.5

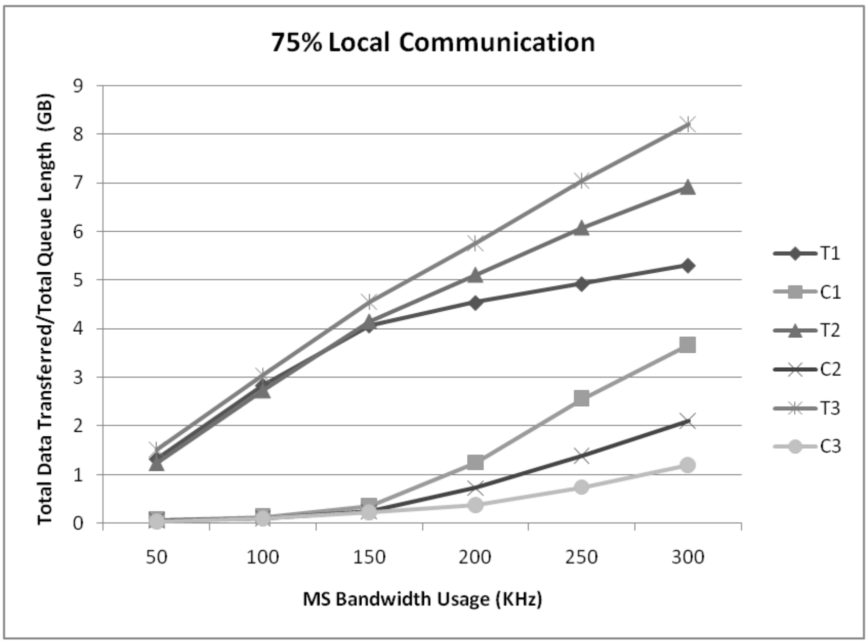

Figure 5.24: Simulation results for probability of traffic destination being in same region with source is 0.75

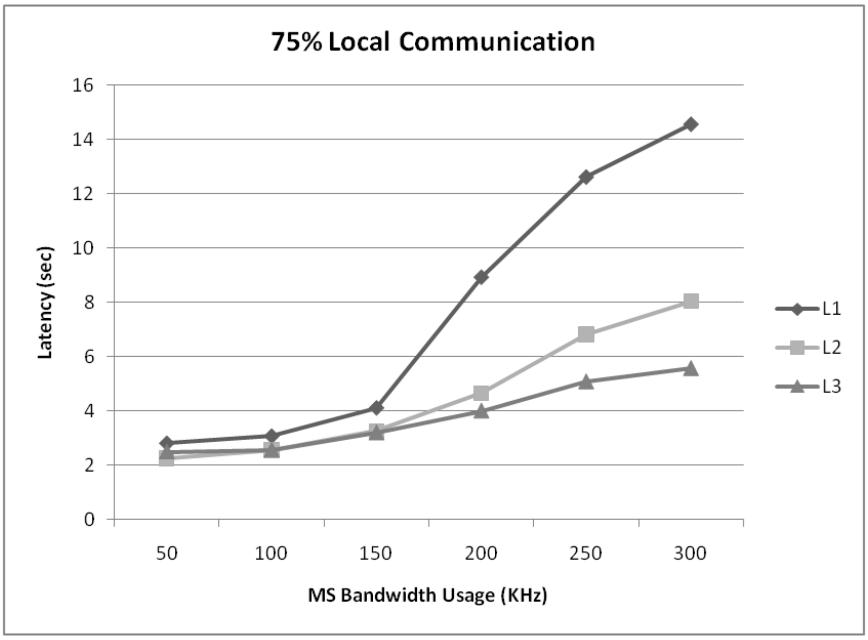

Figure 5.25: Average latency of a packet where probability of selecting destination from same region is 0.75

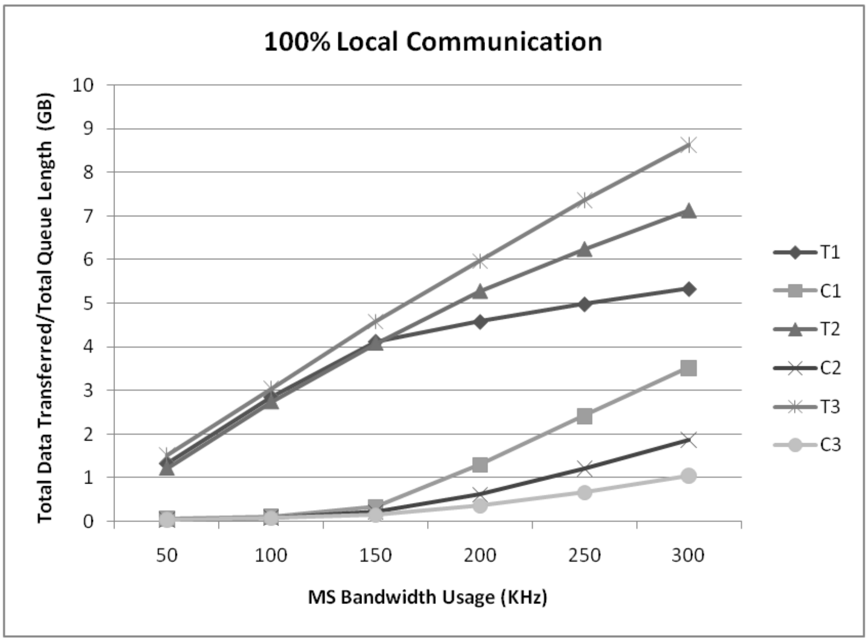

Figure 5.26: Simulation results for probability of the traffic destination being in the same region with the source is 1

By inspecting the graphs in Figure 5.22, Figure 5.24 and Figure 5.26, increasing the probability of traffic destination being in the same region with source node, throughput is

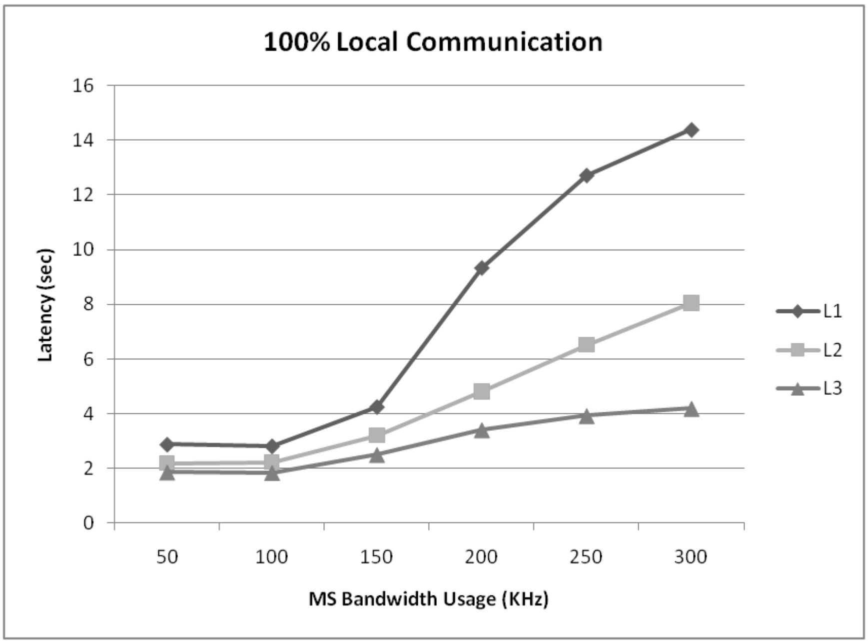

Figure 5.27: Average latency of a packet where probability of selecting destination from same region is 1

improved for proposed traffic aware routing algorithm and shortcut routing with standard path selection. For standard routing scheme queue length and throughput results almost remain same. Slight differences may result because of traffic scenario changes at each simulation run. Figure 5.24 and Figure 5.26, throughput and queue length of shortcut routing with standard path selection scheme remains same. Increase in throughput decelerates after bandwidth utilization passes over 50%. For traffic aware routing, if the probability of traffic destination being in the same region with source node is 1 then there would not be any deceleration in the throughput increase according to Figure 5.26.

According to Figure 5.23, Figure 5.25 and Figure 5.27 average latency does not change for standard algorithm while increasing the local communication. For standard path selection algorithm with shortcut routing, latency decreases down to 8 seconds. The amount of decrease in latency with respect to local communication getting smaller after 50% local communication. With proposed traffic aware routing algorithm, average latency fall down to 4 seconds.

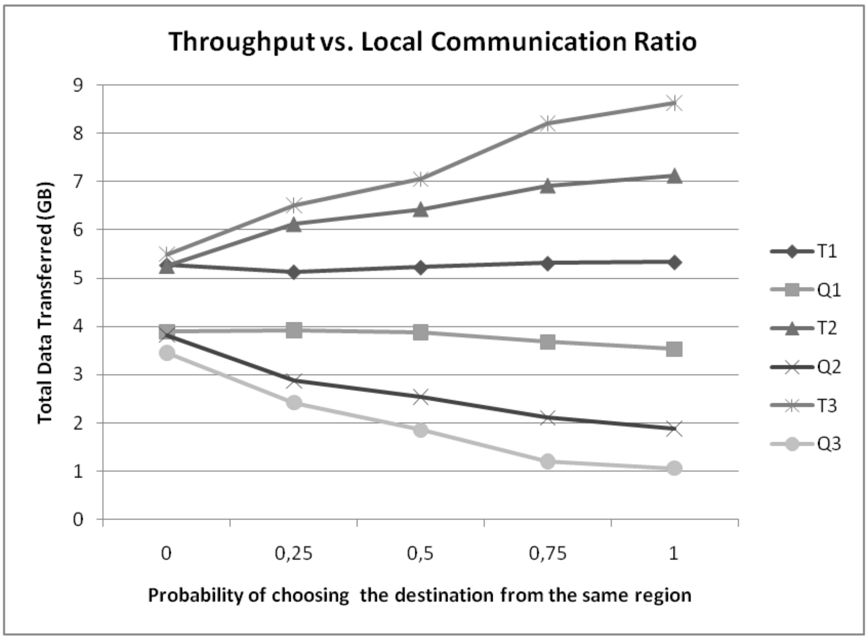

Figure 5.28: Behaviors of algorithms according to the rate of regional traffic

In Figure 5.28 throughput and queue length with respect to the probability of traffic destination being in the same region of source node is shown. The gap between the compared solutions is getting larger as the probability increases. Consequently proposed traffic aware routing solution performs better for topologies where subscribers in communication with peers in vicinity. For Internet usage proposed solution does not provide any benefits since the routing paths almost remain same.

5.3 Comments

Two simulators used to test different topologies and traffic scenarios. Since both simulators have different limitations and assumptions, effects of the proposed solution observed below expectations. In NCTUns, when RS become bottleneck, it serves only some of subscribers and the rest starves. If we categorize traffic types as inside and outside network, outside network traffic may cause inside network traffic starve. Inside network traffic is the targeted case for throughput improvement in proposed solution. Therefore, throughput improvement could not achieved as expected in such cases.UL and DL scheduler of MR-BS and RSs should be improved to serve each subscriber fairly.

Comparing the routing algorithms, performances of algorithms diverge according to the utilization of bandwidth dedicated to subscribers and regional communication. If the subscribers do not fill the dedicated bandwidth, then network resources are more than enough to meet the needs. Traffic aware routing performs slightly better than others because of reduction of hop count with shortcut routing. Regional communication is the targeted problem so decreasing the ratio of regional traffic, performance of algorithms converges to each other.

Worst and best cases for the proposed algorithm are tested in complex scenario where local communication ratio is changed between 0% and 100%. Results for totally random traffic scnearios should be in a place between worst and best cases according to local communication ratio.

CHAPTER 6

CONCLUSION AND FUTURE WORKS

WiMAX (IEEE 802.16) is the latest wireless communication technology which provides last mile broadband wireless access. The IEEE 802.16 Standard is accepted as the 4th generation communication technology since it satisfies today's communication demands like high bandwidth and mobility. WiMAX operates on high frequency spectrum and there is no uniform global licensed spectrum for it. Mobile Non-Line-of-Sight propagation channels require dense deployment of base stations. To overcome this problem, IEEE 802.16j Standard is proposed as an amendment to IEEE 802.16, which enables Multihop Relaying by introducing Relay Stations. Two operation modes are defined for Relay Stations, which are Transparent mode and Non-transparent mode. In the transparent mode, a relay station does not forward framing information, so mobile stations should be in the range of MR-BS. Deployment of transparent relay stations improves the CINR of access links which also increases the throughput. A non-transparent relay station forwards framing information, which extends the network coverage in addition to CINR improvement.

Using non-transparent IEEE 802.16j as a backhaul network for domain specific networks like a Tactical Area Communications Network, throughput can be further improved. In a tactical area, subscribers in the same field communicate with each other more than with nodes outside the field. However, according to the current specification, all uplink data of MSs is forwarded to MR-BS even if the destination MS is connected to the same RS with the source MS. Forwarding data directly from an intermediate RS to the destination MS instead of sending it to MR-BS first may increase the performance by decreasing latency and increasing throughput. Selected the path between MS and MR-BS is another factor that influences the network performance.

Our proposed solution for improving throughput of IEEE 802.16j networks consists of two parts. First part is integrating shortcut routing mechanism to NT-RSs. Second is traffic aware path selection method proposal and embedding it to the original standard. For enabling shortcut routing, a cross layer is added to the protocol stack. Cross layer keep destination CID and destination IP pairs. When an UL IP packet is received at RS, it is checked whether destination IP is in the routing cache list. If the destination IP address is in the list, it means that the destination is subordinate station of the RS, so it is possible to directly forward the packet to destination. Proposed path selection method employs four metrics which are MCS, hop count, load and sent & received traffic analysis. MCS metric is used to select a path which consists of high quality links. Number of hop counts on the path affects latency and network resources. More hop counts means consuming resources of more links. Heavily loaded intermediate stations may become bottleneck and this may results with packet drops. Selecting an access station with less loaded, provides fair load distribution and better network utilization. Sent & received traffic analysis is used to select access station according to traffic demand. Being in the range of more than one RSs and MR-BSs, subordinate station would choose the access station which is more close traffic pairs in tree topology.

To validate our proposed shortcut routing solution, NCTUns simulation tool is used. NCTUns 6.0 is the only simulation tool that has Non-transparent IEEE 802.16j plug-in. A new module is developed for NCTUns to add CROSS layer into NT-RS protocol stack. According to simulation results, shortcut routing neither improve nor degrade the throughput for Internet traffic. For traffic scenarios where subscriber connected to same MR-BS communicating with each other, throughput improvement changes according to the topology. If traffic pairs are connected to same RS and that RS is not overloaded, throughput improvement becomes more distinctive. When RSs become overloaded, scheduling becomes crucial factor for simulation results. If the scheduler serves only the first comer, then the other stations may starve. According to the results of complex topologies tested on NCTUns, UL and DL scheduler of MR-BS and RS do not fairly serve to subscribers.

Due to the limitation of NCTUns, proposed system cannot be tested on it. NCTUns 6.0 does not support topologies with depth more than 2 hops. Also mobility support is not given in current version of the tool. A custom simulator is implemented to test the overall system. Complex topologies are tested in our simulator to evaluate the performance of original standard, shortcut routing without traffic aware path selection and shortcut routing with traffic aware path selection. A topology is constructed by deploying RSs and MR-BSs as a grid and uniformly distributing MS inside grid. Topology is divided into four regions symmetrically and each region has one MR-BS. Simulation results show that increasing the probability of choosing MS's traffic destination in the same region, shortcut routing with traffic aware path selection outperforms the original standard. If the MSs do not entirely utilize the bandwidth that is allocated to them, RSs do not become congested so the only proposed solution performs slightly better since none of the links become bottleneck. The best case scenario which puts forward the shortcut routing efficiency is subscribers that are connected to RS, communicating with each other. Throughput improvement in this scenario is 37% according to original standard .For the hybrid traffic scenario throughput improvement is no more than 3% since shortcut routing possibility is low and scheduler does not serve to some stations.

Our proposed solution is entirely tested in our simulator. Shortcut routing, hierarchical addressing, MobileIP enhancement and traffic aware path selection are tested in the same environment. When shortcut routing and path selection method enabled together, throughput improvement reaches to 60%. Major factor that affects the throughput is the MCS level used for links on path. When calculating the path cost, minimum MCS level on path is used as metric. Two subscribers that are communicating with each other may have common intermediate RS. If the links on the selected paths for subscribers do not have low MCS levels, subscribers can transmit with higher bitrates. Traffic aware path selection method enforces communicating subscribers to have common intermediate RS. Comparing the original standard with proposed solution, our routing scheme performs better as the ratio of local commutations to Internet increases. Proposed system throughput always above the throughput of original system and RSs are less congested.

Various improvements can be further made to increase the throughput of IEEE 802.16j Non-transparent Networks. Each metric in the proposed path selection method has a coefficient. Weight of the metric can be changed according to topologies and traffic scenarios. Predetermining these factors may not be feasible in real world conditions. By using machine learning methodologies, coefficient of metrics could be selected according to observations.

In our proposed solution, improvements are made adhering to IEEE 802.16j standard. By adding new management messages and physical capabilities, more efficient solution can be obtained. One of the challenges that we try overcome was subscribers and RSs do not aware of locations of destinations. By using hierarchical addressing, problem was resolved but hierarchical addressing brings limitations to size of network. Protocol may be extended by adding new management messages for querying destination node's location to overcome addressing problem.

The metrics used for path selection in literature are calculated for selecting the best path between MS and MR-BS. These metrics may not be meaningful for shortcut routing paths. These metrics should be adapted for shortcut routing scenarios to evaluate paths correctly. Further throughput and latency improvements can be made by providing novel scheduling algorithms. By adding downlink queues for local communication data, and adjusting scheduling mechanism to give priority to local communication, throughput could be increased. Vice versa, by separating local communication queues, priority may be given to Internet traffic. In this thesis, we provide scheduler only for UGS traffic. Our solution is applicable for all traffic types but results may different. For each traffic type proposed solution could be tested and metrics can be adapted according to demands.

REFERENCES

[1] IEEE, ”IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems Amendment 1: Multiple Relay Specification,” *IEEE Std 802.16j-2009 (Amendment to IEEE Std 802.16-2009)* , pp.c1-290, June 12 2009

[2] Genc V., Murphy, S. Yang Yu, Murphy J. , ”IEEE 802.16J relay-based wireless access networks: an overview”, *IEEE Wireless Communications* , vol.15, no.5, pp.56-63, October 2008

[3] Roger B. Marks, ”The IEEE 802.16 Working Group on Broadband Wireless Access Standards”, *http://grouper.ieee.org/groups/802/16/*, Last update on April 11 2010

[4] ”WiMAX”, *http://en.wikipedia.org/wiki/WiMAX*, Last accessed on April 21 2010

[5] ZTE, ”Recent Progress of WiMAX Standards” *http://wwwen.zte.com.cn/en/products/wireless/wimax/200903/t20090324_171163.html*, Last accessed on April 21 2010

[6] ”IEEE 802.16”, *http://en.wikipedia.org/wiki/IEEE_802.16*, Last accessed on April 15 2010

[7] ”IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems”, *IEEE Std 802.16-2009 (Revision of IEEE Std 802.16-2004)* , pp.C1-2004, May 29 2009

[8] Bo Li, Yang Qin, Chor Ping Low, Choon Lim Gwee , ”A Survey on Mobile WiMAX [Wireless Broadband Access]”, *IEEE Communications Magazine*, vol.45, no.12, pp.70-75, December 2007

[9] Debalina Ghosh, ”Mobile IP”, *http://www.acm.org/crossroads/xrds7-2/mobileip.html*, Last accessed on April 21 2010

[10] ”Shortcut Routing for Relay-based Systems”, *FUTURE MOBILE COMMUNICATION FORUM Working Group Technical White Paper*, pp.237-244, 2007

[11] Xiaopeng Hu, Xuming Fang, Zhengguang Zhao, Yan Long, Yue Zhao, ”A Local Forwarding Implementation Scheme in OFDMA-Based Multihop Cellular Networks”, *International Conference on New Trends in Information and Service Science*, pp. 1263-1267, 2009

[12] Bo Wang, Mutka M. , ”Path selection for mobile stations in IEEE 802.16 multihop relay networks”, *International Symposium on a World of Wireless Mobile and Multimedia Networks*, pp.1-8, 23-26 June 2008

[13] Sojeong Ann, Kyung Geun Lee, Hyung Seok Kim , ”A Path Selection Method in IEEE 802.16j Mobile Multi-hop Relay Networks”, *Second International Conference on Sensor Technologies and Applications SENSORCOMM ’08*, pp.808-812, 25-31 Aug. 2008

[14] Kuei-Ping Shih, Sheng-Shih Wang, Chan-Ying Lien , ”A high spectral efficiency and load-aware metric for path selection in IEEE 802.16j multi-hop relay networks”, *IEEE Symposium on Computers and Communications ISCC 2009*, pp.61-66, 5-8 July 2009

[15] Yeong-Sheng Chen, Chih-Chieh Lin, Sheng-De Wang , ”A traffic-aware routing algorithm for IEEE 802.16j multihop relay networks”, *Fourth International Conference on Communications and Networking in China ChinaCOM 2009*, pp.1-10, 26-28 Aug. 2009

[16] ”Classless Inter-Domain Routing”, *http://en.wikipedia.org/wiki/Classless_Inter-Domain_Routing*, Last accessed on April 21 2010

[17] S.Y. Wang, H.Y. Chen, and S.W. Chuang, ”NCTUns Tool for IEEE 802.16j Mobile WiMAX Relay Network Simulations,” *A chapter of the ”Recent Developments in Mobile Networks Research” book*, (to be published by Nova Science Publishers in 2010)

[18] S.Y. Wang, C.L. Chou, C. C. Lin, and C.H. Huang, ”The Protocol Developer Manual for the NCTUns 6.0 Network Simulator and Emulator”

[19] S.Y. Wang, C.L. Chou, and C.C. Lin, ”The GUI User Manual for the NCTUns 6.0 Network Simulator and Emulator”

[20] ”WiMAX (802.16) specialized model”, *http://www.opnet.com/WiMAX/index.html*, Last accessed on April 21 2010

[21] WiMAX FORUM, ”NS-2 Simulator”, *http://members.wimaxforum.org/kws/ns2_download/*, Last accessed on April 21 2010

[22] ”Discrete event simulation”, *http://en.wikipedia.org/wiki/Discrete_event_simulation*, Last accessed on April 21 2010

[23] ”What is WiMAX”, *http://www.wimax.com/education*, Last accessed on April 21 2010

[24] M.Okuda, C.Zhu, D.Viorel, ”Multihop Relay Extension for WiMAX Networks - Overview and Benefits of IEEE 802.16j Standard”, *FUJITSU Scientific & Technical Journal*, Vol.44 No.3, p. 292-202, July 2008

[25] Fujitsu Limited, ”Fujitsu Scientific & Technical Journal Archives”, *http://www.fujitsu.com/global/news/publications/periodicals/fstj/archives*, Last accessed on April 21 2010

[26] Wang S.Y., Chou C.L., Huang C.H., C.C. Hwang, Yang Z.M., Chiou C.C., and Lin C.C., ”The Design and Implementation of the NCTUns 1.0 Network Simulator”, *Computer Networks*, Vol. 42, Issue 2, pp.175-197, June 2003.

[27] Yanlei S. and Shiduan C., ”An Enhanced Packet Scheduling Algorithm for QoS Support in IEEE 802.16 Wireless Network”, *Networking and Mobile Computing*, pp.652-661, 2005