

WIRELESS COMMUNICATION FOR SPARSE AND RURAL AREAS

by

Mingliu Zhang

A dissertation submitted in partial fulfillment
of the requirement for the degree

of

Doctor of Philosophy

in

Engineering

MONTANA STATE UNIVERSITY
Bozeman, Montana

July 2007

©COPYRIGHT

by

Mingliu Zhang

2007

All Rights Reserved

APPROVAL

of a dissertation submitted by

Mingliu Zhang

This dissertation has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the Division of Graduate Education.

Dr. Richard Wolff

Approved for the Department of Electrical and Computer Engineering

Dr. James N. Peterson

Approved for the Division of Graduate Education

Dr. Carl A. Fox

STATEMENT OF PERMISSION TO USE

In presenting this dissertation in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this dissertation should be referred to ProQuest Information and Learning, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute my dissertation in and from microform along with the non-exclusive right to reproduce and distribute my abstract in any format in whole or in part."

Mingliu Zhang

July 2007

ACKNOWLEDGEMENTS

I am grateful to acknowledge and thank all of those who assisted me in my graduate study at Montana State University. First, I would like to thank Dr. Wolff, my academic advisor here at MSU. His nice guidance, great support, patience and personal time throughout my years as a graduate student have been deeply appreciated. Special thanks are given to my other graduate committee members for their valuable suggestions and comments. Special thanks also go to Wenhao Lin, my colleague student, for his time and valuable discussions on program design and implementation in the OPNETTM. I would also like to thank our MSU wireless research group for their extensive discussions and comments to help improve my research work. Most importantly, I would like to thank my family and friends for their support throughout all the years.

TABLE OF CONTENTS

1. INTRODUCTION	1
Worldwide Internet and Broadband Usage.....	1
U.S. Internet and Broadband Usage	2
Why Wireless Communication for Sparse and Rural Areas?.....	4
Technologies for Broadband Access in the U.S.....	4
Potential of Wireless Communications in Rural Areas.....	5
Fixed Wireless.....	5
Mobile Wireless	6
History of Wireless Communications in Rural Areas	7
Wireless Voice Service.....	7
Wireless Data Service.	8
Challenges and Opportunities	12
Recent Progress in Wireless Technologies	12
802.11x.....	12
Wi-Fi	13
WiMAX	14
MANET	14
Need for the Research	15
Scope of the Study.....	16
Organization of the Dissertation.....	17
References	19
2. BASELINE NETWORK MODEL FOR BROADBAND FIXED WIRELESS ACCESS IN RURAL AREAS	22
Introduction	22
Approach and Assumptions.....	23
Approach and Wireless Network Design Tool.....	24
Geographical Areas Chosen for Baseline Network Model Study.....	26
Data Traffic and User Penetration Assumptions	27
Market Share for Wireless ISPs	28
Customer Premises Equipment (CPE) Cost.....	29
Baseline Network Model.....	29
Baseline Network Model Structure.....	30
Access Point (AP).	31
Wireless Ethernet Bridge.....	31
X Base T.	31
Model Parameters and Technology Specifications	32
Simulation Results.....	34
Case A: Baseline Case	34
Case B: Baseline Case with High Gain Antenna Applied.....	38
Case C: Baseline Case with Switched Array Antennas	38
Case D: Multi-hop Connections.....	40

Discussion	42
Summary	46
References	47
3. MULTIHOP NETWORK FOR BROADBAND FIXED WIRELESS ACCESS IN RURAL AREAS	48
Introduction	48
Multi-hop Model and Analysis	51
The Basic Multi-hop Model (SMN model) for WLAN	52
The SMN Model with Sector Antennas (SMNS) for WLAN	56
Multi-hop Model Selection	58
Demographic Information about the Study Area	58
Data Traffic Requirements and Technology Specification	59
Solutions Based on the SMN Model	60
Solutions Based on SMNS Model	61
Simulations and Results	62
802.11b vs. 802.11g	62
SMN2 vs. SMNS2	64
SMNS1, SMNS2 and SMNS3 vs. Population Density ρ	65
Sensitivity Analysis of the Critical Population Density	68
Summary	70
References	71
4. TOWARDS MOBILE WIRELESS COMMUNICAITON IN RURAL AREAS	72
Introduction	72
Fixed Multi-hop Wireless Networks and Mobile Ad hoc Networks	73
Routing Challenges for Mobile Ad hoc Networks in Rural Areas	76
Overview of the Existing MANET Routing Protocols	76
Characteristics of MANETs in Rural Areas	79
Needs for New MANET Routing Protocols	80
Mobility Models for Mobile Ad hoc Networks in Rural Areas	81
Summary	85
References	86
5. GEOGRAPHIC AND TRAFFIC INFORMATION BASED MOBILITY MODEL FOR MOBILE AD HOC NETWORKS IN SPARSE AND RURAL AREAS	88
Introduction	88
Related Work on Mobility Models	91
Existing Synthetic Mobility Models	91
Traces for Mobile Ad Hoc Networks in Rural Areas	93
Geographic and Traffic Information Based (GTI) Mobility Model	93
Fundamentals of GTI Mobility Model	94
The Generation of Traces for Mobile Nodes	95
Selection of the Sampling Time Interval in GTI Mobility Model	98
Application Example of the GTI Mobility Model	100
Application of GTI for Evaluation of a Routing Protocol	102

The Ideal Routing Protocol.....	103
Simulation Environment and Parameters.....	104
Simulation Results	105
Summary	111
References	112
6. BORDER NODE BASED ROUTING PROTOCOL FOR MANETS IN RURAL AREAS.....	113
Introduction	113
BBR Routing Protocol Overview.....	116
Assumptions.....	117
Neighbor Discovery Algorithm.....	118
Neighbor Discovery Process.	118
Distributed Border Node Selection Algorithm	122
Broadcast Based Routing Algorithm.	122
Heuristic for the Selection of Border Nodes.	124
Border Node Selection Scheme and Rules.....	128
Border node function.....	131
Conceptual Data Structures	132
Border Node Selection Table	132
Forward Table	133
BBR Options Header Format	134
Fixed Portion of the BBR Options Header	135
Neighbor Discovery Option (Option 1)	136
Data and Neighbor Destination Option (Option 2).....	137
Data and None Neighbor Destination Option (Option 3).....	139
Data Acknowledgement Option (Option 4)	140
Detailed Operation of the BBR Routing Protocol.....	142
General Packet Processing.....	142
Originating a Packet.	142
Add a BBR Options Header to a Packet.....	143
Originating a Hello Packet.	144
Originating a Data Packet.....	145
Originating a Data Acknowledgement Packet.....	146
Handling an Unknown BBR Option.	146
Processing a Received Packet.....	147
Processing a Received Neighbor Discovery Option.	147
Processing a Received Data and Neighbor Destination Option.	147
Processing a Received Data and None Neighbor Destination Option.	148
Processing a Received Data Acknowledgement Option.	151
Flowchart for Received Packet Processing.	151
Protocol Constants and Configuration	153
Some Features of BBR Routing Protocol.....	153
Scheduled Forwarding for Mitigating the Hidden Node Problem.....	153
Redundant Transmission Prevention.....	156
Effects of the Hello_Interval Length on the Delivery Ratio.....	158

Border Node Selection Optimization for a Fully Connected Network.....	161
Summary	164
References	165
7. BBR ROUTING PROTOCOL IMPLEMENTATION AND PERFORMANCE EVALUATION	167
Introduction	167
BBR Routing Protocol Process Model Development in OPNET™ Modeler	167
MANET Protocol Modeling in OPNET	167
Node Model for a Mobile MANET Station.....	168
Manet_mgr Architecture.....	169
Adding a Custom MANET Protocol.	170
BBR Routing Protocol Development in OPNET™	171
BBR Protocol Process Model.....	171
BBR Protocol Packet Format. The	173
BBR Protocol Packet Flow. The.....	174
Performance Evaluation	179
General Simulation Scenario	180
Packet Delivery Ratio vs. Simulation Time.	182
Packet Delivery Ratio vs. MaxRebroadcast Number. The	184
Packet Delivery Ratio vs. HelloInterval.....	187
Packet Delivery Ratio vs. Radio Transmission Range.....	188
Average Packet Delivery Delay vs. Radio Transmission Range.....	190
Comparisons between BBR Routing Protocol and DSR Routing Protocol.	191
Highway Scenario	197
BBR vs. Random Node Selection (RNS)	202
Summary	206
References	207
8. CONCLUSIONS.....	208

LIST OF FIGURES

Figure	Page
1.1. World Internet users and broadband subscribers 1999-2010[1.1], [1.2].....	2
1.2. High-Speed Internet access in US, as of December 31, 2005 [1.5].....	3
1.3. High-Speed lines by technology as of December 31, 2005 in the U.S. [1.7].	5
2.1. SWAT parameters and modeling flow.....	25
2.2. The Fisher-Pry (S-shaped) model for data service subscriber penetration.	28
2.3. The baseline network structure for BFWA in rural areas.....	30
2.4. Capital and expense flows for the composite area in the ten-year study period, (a) Baseline case, (b) Case B and (c) Case C.....	37
2.5. Discounted cash flow graph of the composite area for the three cases.	40
2.6. Part of the network with the multi hop network structure applied.	41
3.1. A simplified multi-hop wireless network.....	49
3.2. Star Mesh Network (SMN) model for the multi-hop WLAN.....	52
3.2. Performance of SMNS models vs. population density ($R_1 = 2.2$ miles).	66
3.3. Performance of SMNS models vs. population density ρ ($R_1 = 1.1$ miles).	66
3.4. Critical population density vs. first hop transmission range of SMNS1.....	68
3.5. ρ_{c12} vs. average traffic load at a fixed first-hop radio range of SMNS1.	69
4.1. A simplified fixed multi-hop wireless network.....	74
4.2. A sample of a simplified mobile ad hoc network.....	75
4.3. Cellular signal coverage test along a highway in Montana. Courtesy: Doug Galarus, Western Transportation Institute.....	84
5.1. Verizon cellular coverage map in the northwest area of U.S. [5.1].	89

LIST OF FIGURES-CONTINUED

Figure	Page
5.2. Initial node distribution on an edge.	96
5.3. Direction choosing rule at a cross point with $k = 3$	97
5.4. GTI mobility model trajectory generation flow diagram.....	98
5.5. Geographic information of YNP [5.12].	100
5.6. Mobile nodes and their trajectories in YNP.	102
5.7. Probability density function of average neighbor node degree.	108
6.1. Packet format of Hello message.	119
6.2. Table header for Neighbor Table.....	119
6.3. Hello message information recording procedure.....	120
6.4. Pseudo code for updating a Neighbor Table.	121
6.5. A typical broadcast and node distribution in a connected network.....	126
6.6. A typical broadcast and node distribution in a partially connected network.	127
6.7. A typical network with a broadcast source node s	131
6.8. The fixed portion of the BBR Options Header format.....	135
6.9. The IPv4 Header.	135
6.10. The Neighbor Discovery Option header.	136
6.11. Data and Neighbor Destination Option header.	137
6.12. Data and None Neighbor Destination Option header.	139
6.13. Data Acknowledgement Option header.	141
6.14. BBR Options Header position in a regular IP packet.	144
6.15. Flowchart for received packet processing.....	152

LIST OF FIGURES-CONTINUED

Figure	Page
6.16. Node distribution with possible hidden nodes problem.....	154
6.17. Hidden node problem in BBR routing.	154
6.18. Redundant transmission when two nodes meet each other.....	157
6.19. Illustration of short Hello_Interval problem.	159
7.1. Node model architecture for a mobile MANET station node [7.1].	169
7.2. Manet_mgr architecture in a MANET station node model [7.2].	170
7.3. Adding a custom routing protocol [7.3].	171
7.4. Finite state machine representation of the BBR process model.	172
7.5. BBR packet format.	174
7.6. Control packet from routing protocol [7.4].	175
7.7. Control packet from lower layer [7.4].	176
7.8. Data packet flow when a data packet is from application layer.....	177
7.9. Data packet flow when a data packet is from a lower layer.....	179
7.10. Packet delivery ratio vs. simulation time.	184
7.11. Packet delivery ratio vs. the MaxRebroadcast number.	185
7.12. Packet delivery ratio vs. HelloInterval.....	187
7.13. Packet delivery ratio vs. radio transmission range.....	189
7.14. Average delay per delivered packet vs. radio transmission range.....	190
7.15. Packet delivery ratio comparison between BBR and DSR protocols.	192
7.16. Average packet delivery delay comparison between BBR and DSR protocol.	195
7.17. Packet delivery ratio comparison between BBR and DSR protocol.....	200

LIST OF FIGURES-CONTINUED

Figure	Page
7.18. Packet delivery delay comparison between BBR and DSR protocol.	202
7.19. Packet delivery ratio comparison between BBR and RNS.....	203
7.20. Packet delivery delay comparison between BBR and RNS.....	204

LIST OF TABLES

Table	Page
2.1. Demographic Information for the Cities and Towns of Gallatin County	27
2.2. Network Components Cost and Expense Assumptions.....	35
2.3. NPV and Years to Break Even of Each CA and the Composite Area for the Three Different Cases	36
2.4. Impact of Increased Traffic Load on the Financial Performance.....	43
2.5. Effect of Raising the Discount Rate on the Financial Results	44
3.1. Simulation Result with 802.11b and g Technology Applied Using Conventional Single-Hop Point-to-Multipoint Topology	63
3.2. Cost Assumptions for Antennas Applied	64
3.3. Simulation Results Comparison between SMN2 and SMNS2 Models.....	64
5.1. Simulation Parameters	105
5.2. Transit Time vs Ratio α (Node Velocity: 12.1~14.7 m/s)	106
5.3. Transit Time vs Ratio α (Node Velocity: 24.2~29.4 m/s)	106
5.4. Transit Time vs Initial Velocity Randomness	109
5.5. Transit Time vs Initial Position Randomness	110
6.1. Comparison of Multipoint Relay Scheme, Cluster-Based Broadcast Scheme & BBR Routing Scheme.....	124
6.2. Broadcast Data Packet Format.....	128
6.3. Rebroadcast Packet Format.....	130
6.4. Border Node Selection Process Illustration	131
6.5. Border Node Selection Table	132

LIST OF TABLES-CONTINUED

Table	Page
6.6. Forward Table Header.....	133
6.7. Some Configurable Parameters and Their Default Values.....	153
7.1. Simulation Parameters	182
7.2. Simulation Parameters for Highway Scenario	198

Abstract

Wireless technology experienced a fast development in the past few decades. However, research and investment in wireless communication so far has been focused mainly on high-density domains or fully connected networks. The technologies/solutions developed for above domains do not readily apply to rural and sparse domains. The users in rural and sparse areas are still served predominantly by either low-speed dialup access or have no data service available at all. This research work explores the largely overlooked rural and sparse domains, where distance, rough terrain and low node density are the key parameters driving system design and performance, from the perspectives of fixed wireless applications to mobile wireless applications.

For fixed wireless applications, a baseline wireless network structure for rural and sparse areas is defined and the potential for improved high-speed fixed communication services in rural and remote areas is examined. The potential of using multi-hop network topologies in very sparse areas is explored. The cost benefits of several other emerging technologies and approaches are also investigated with the objective of finding cost-effective and affordable high-speed broadband communications solutions for rural and remote areas.

Mobile ad hoc networks (MANETs) are also examined as an approach to providing connectivity under conditions where fixed communications infrastructure is non-existent. The unique characteristics of MANETs in rural areas are analyzed and a Geographic and Traffic Information based mobility model (GTI mobility model) is proposed to model mobile node movement under real world constraints. A Border node Based Routing (BBR) protocol is designed specifically for MANETs in rural areas. Simulation results show that the BBR routing protocol yields better performance than Dynamic Source Routing (DSR) when the network is partially connected and has comparable performance to DSR when the network is fully connected. BBR also has better performance than a Random Node Selection algorithm (RNS) when the network is partially connected. The results show that BBR can be applicable to a range of rural area applications including sparse vehicular ad hoc networks (VANETs) for public safety uses as well as a broader class of sparse ad hoc network applications where the nodes may exhibit a more random mobility pattern.

CHAPTER 1

INTRODUCTION

The internet has become an essential part to more and more people's daily life. The demand for broadband access has grown steadily as users experience the convenience of high-speed responses combined with "always on" connectivity. Among the different kinds of telecommunication technologies, wireless communication has become one of the most attractive ways to connect information highways with people either staying at home or on the move. In this chapter, an overview is given on wireless communication technologies used for fixed and mobile communication services provided in urban, rural and remote areas. Recent progress and development trends in wireless communication technologies are also reviewed in this chapter. Finally, the need for this research work is addressed and the scope of the research is also defined.

Worldwide Internet and Broadband Usage

The number of worldwide Internet users reached about 1,093 million by the end of 2006 [1.1] and is forecast to grow 64.7% to approach 1,800 million in 2010[1.2]. Figure 1.1 shows the strong growth of the worldwide Internet usage in the past several years and its projection over the next several years. Also shown in Figure 1.1 is the number of worldwide broadband subscribers, which reached about 215 million at the end of 2005 and is expected to grow 132.5% and reach 500 million by the end of 2010 [1.2]. Nevertheless, despite the higher growth rates of broadband subscribers, the number of narrowband users is still much higher than the number of broadband users worldwide and

is expected to remain so in the coming several years. Many Internet users who can't get broadband service are still using narrowband service. Moreover, as the cost for broadband access is more than narrowband access, the situation that narrowband users outweigh broadband users might not change any time soon.

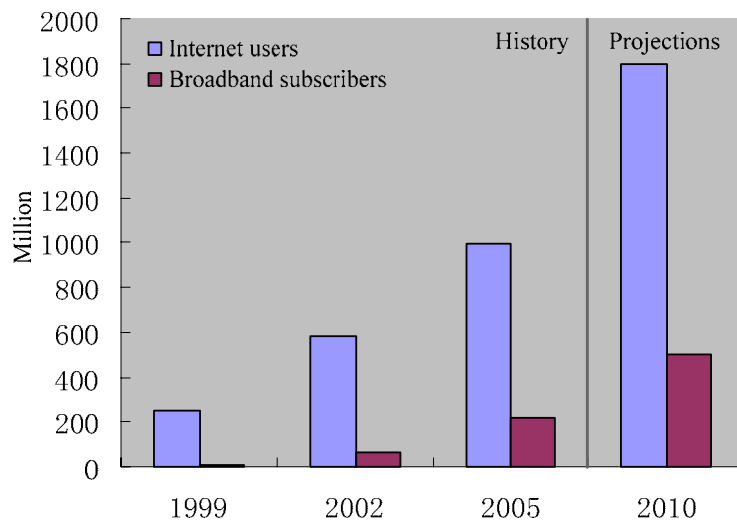


Figure 1.1. World Internet users and broadband subscribers 1999-2010[1.1], [1.2].

U.S. Internet and Broadband Usage

In the United States, by the end of year 2005, there were about 204 million Internet users, of which 46.9 million were broadband users, a 22.9 percentage of the total [1.3]. Broadband, as defined by the Federal Communications Commission (FCC), is a data service with transmission speeds exceeding 200 kilobits per second (Kbps), in at least one direction: downstream or upstream.

As compared to the whole world, the USA leads in the total broadband population. It is anticipated that about 61.5 million users in the United States will have broadband service by the end of 2008 [1.4], provided by a combination of cable modem, Digital

Subscriber Line (DSL), broadband wireless, satellite, first mile fiber, and Broadband over PowerLine (BPL). Figure 1.2 shows the high speed provider (ISP) geographic distribution in the United States, according to FCC data [1.5]. The blue color indicates those areas with 7 or more high speed ISPs. Light brown represents those areas where there are 1-3 high speed ISPs, while the green color indicates those areas that have no high speed ISP. The figure indicates that there are still large areas in the middle and north western states where there are very few or even no high speed ISPs. By the end of 2006, home broadband penetration in rural areas was only 25%, while in urban and suburban areas the number was 44% and 46%, respectively [1.6].

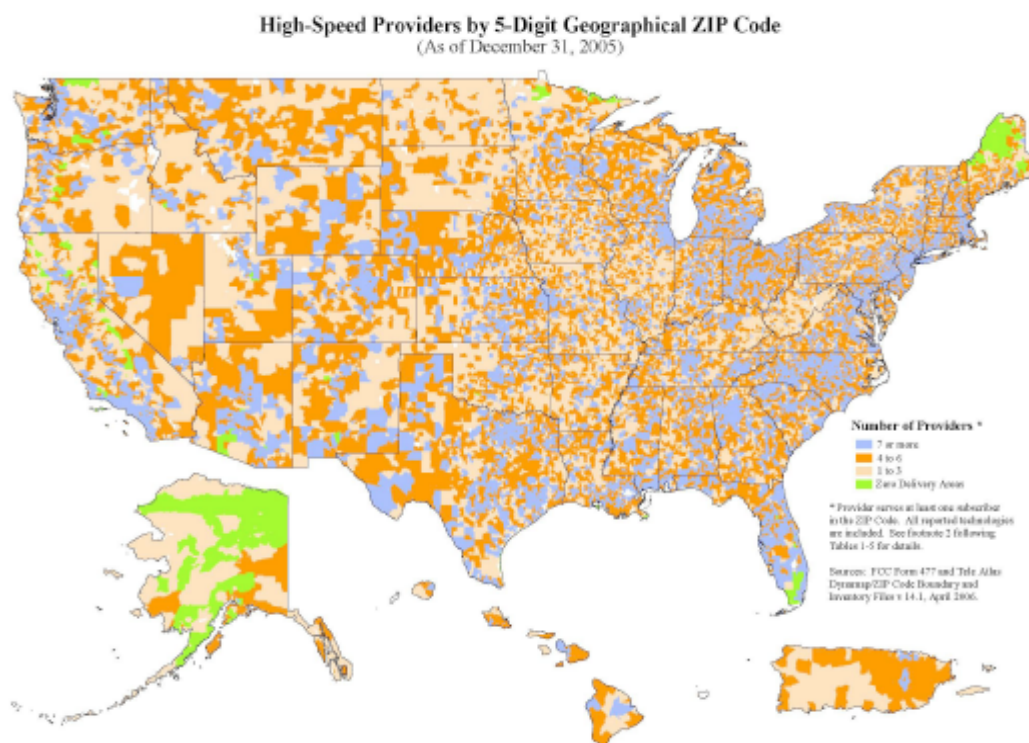


Figure 1.2. High-Speed Internet access in US, as of December 31, 2005 [1.5].

Why Wireless Communication for Sparse and Rural Areas?

The availability of broadband access to high quality Internet service is one of the factors improving people's quality of life, especially for those people living in rural and remote areas. Broadband will provide new opportunities and information resources for communities in rural and remote areas in many aspects, such as in education, culture, entertainment, tele-health and tele-medicine. However, because of the relatively low population density, longer geographical distance, and larger topographical barriers, broadband in rural and remote areas is generally much more difficult to obtain as compared to urban and suburban areas. Due to a combination of technology and economic factors, there are many places, particularly in rural and remote areas, where Internet users are still using narrowband connections such as low-speed modem dialup access, or have no Internet service available at all.

Technologies for Broadband Access in the U.S.

In the United States, broadband services are provided by a combination of technologies. Figure 1.3 gives the percentages of high-speed lines in the U.S. by technology as of December 31, 2005[1.7]. The figure indicates that wireless communication contributes only a small portion, which is less than 7.58%, to the whole high speed services. The main reason behind this is that broadband users are predominantly located in urban or suburban metropolitan areas, where a fixed wired infrastructure, such as coaxial cable or ADSL, is available and has been upgraded to support high-speed Internet services. However, in rural areas, wired infrastructures are generally not available. Broadband deployment with any of the wired solutions is

generally prohibitively expensive and not cost effective in rural areas.

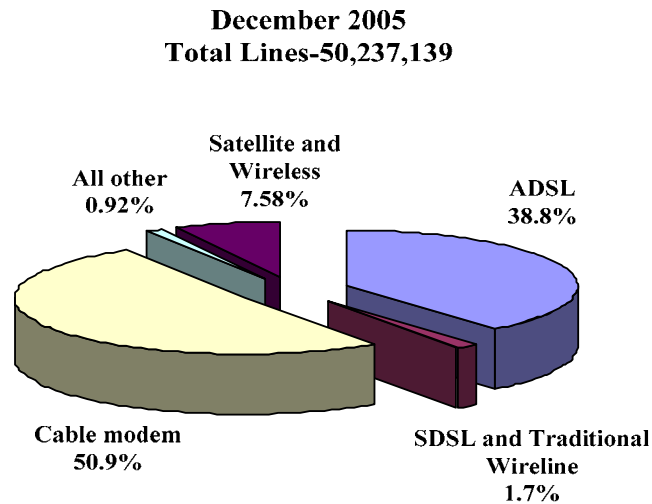


Figure 1.3. High-Speed lines by technology as of December 31, 2005 in the U.S. [1.7].

Wired solutions that need fixed infrastructure are not readily applicable to rural and sparse areas. This provides a chance for wireless communications. With the wireless technology advances in recent years, wireless communications has the potential to play a more important role in providing broadband access in rural and sparse areas. And, sooner or later, it will become tomorrow's main technology for high-speed transmission, perhaps not only in rural and remote areas, but also in urban and suburban areas.

Potential of Wireless Communications in Rural Areas

Fixed Wireless. The potential of using fixed wireless for broadband access has been widely discussed [1.8-1.10] and is being realized in some urban areas for specialized applications using local multipoint distributed systems (LMDS) [1.11] or multi-channel multi-point distribution systems (MMDS) [1.12]. Advances in wireless technology and

system-level applications for data communications, embodied in the IEEE 802.11 standards and commercialized by numerous suppliers as Wi-Fi, have stimulated the wide deployment of Wireless LANs on campuses, in enterprises, metro areas, and public places. IEEE 802.16-2004, the leading fixed wireless broadband standard and technology today [1.13], also referred to as Worldwide Interoperability for Microwave Access (WiMAX), provides an exciting addition to current broadband options for metropolitan area networks (MAN). It is expected that the combination of Wi-Fi with WiMAX will also open possibilities for broadband fixed wireless access in rural and sparse areas where no broadband access is currently available.

Mobile Wireless. While considerable efforts have been given to fixed wireless broadband access, the technology for wireless communication among mobile nodes, referred to as mobile ad hoc networking (MANET) technology, has also attracted increased attention. A mobile ad hoc network is a self-configuring network of mobile nodes connected by wireless links and may take an arbitrary topology, in which nodes communicate peer to peer through multi-hop routing. MANET technology was originally designed for military applications. With rapid technology development in data processing and smart array antennas, now the applications of ad hoc technology have expanded to a variety of areas, such as mobile sensor networks, military operations, public safety and disaster recovery [1.14-1.17]. Ad hoc networks are also used as a communication infrastructure in inter-vehicle communication systems such as in CarNet [1.16] and FleetNet [1.17]. Compared with the traditional single hop mobile cellular phone network, the ad hoc network promises convenient infrastructure-free communication with instant

deployment. It is a distributed dynamic wireless network without centralized management and uses multi-hop techniques to expand radio transmission range. As in a broadband fixed wireless domain, the applications of ad hoc networks have also been largely focused on high user density (urban) areas. Little of this research has been directed to rural areas with low population density. The overarching question that this research addresses is the potential for improved high speed fixed and mobile wireless communication services in rural and remote areas.

History of Wireless Communications in Rural Areas

Bringing wireless communication to rural areas is not a new idea. There is a long history of providing wireless telephone service to rural areas and there are successful stories. However, the history of providing data service such as today's broadband Internet access to rural areas is relatively short and has experienced significant difficulties.

Wireless Voice Service. As early as the 1960's, the Bell System introduced Improved Mobile Telephone Service (IMTS), and some IMTS systems were implemented for fixed telecommunication service in rural areas. In the middle of the 1980's, Basic Exchange Telephone Radio Service (BETRS) technology was developed and used as the last segment of the local loop to provide wireless telephone service to subscribers in remote areas [1.18]. As IMTS gradually became an obsolete technology it was replaced with BETRS systems for fixed wireless access to the public switched telephone network (PSTN). These early wireless systems provided voice and voice band data services. Later, the passage of Telecommunication Act of 1996 ("1996 Act") eliminated the historical barriers for competition, allowing common carriers other than the Incumbent Local

Exchange Carriers (ILECs) to be designated as an Eligible Telecommunication Carrier (“ETC”) for the purpose of providing universal service support. This legislation greatly encouraged competition in telecommunications in rural areas as numerous new companies began to offer fixed and mobile services. Western Wireless serves as one example. By the end of 2004, Western Wireless had 1.4 million subscribers primarily in rural areas in 19 western states [1.19] under the brands Cellular One and Western Wireless. Although mainly offering cellular phone service, Western Wireless also provides residential phone service by using its cellular network infrastructure. Compared with the local wireless telephone service provided by new entrants such as Western Wireless and others, the old BETRS system exhibits significant disadvantages in terms of the relatively high maintenance and investment costs and gradually lost its appeal in market. Furthermore, the growth in Internet use has changed the demand for telecom services from voice to voice plus high speed data. The demand for high-speed Internet access could not be met by the older wireless systems. Solutions such as BETRS, analog cellular and even second-generation digital cellular are not capable of meeting Internet demands for bandwidth.

Wireless Data Service. In 1998, the FCC auctioned the spectrum around 28GHz for Local Multipoint Distribution Services (LMDS), which can provide broadband wireless point-to-multipoint service with a range of 3~4 miles and a maximum aggregate bandwidth of 1.5 Gbps downstream and 200 Mbps upstream [1.20]. LMDS is a licensed spectrum solution, adding costs to the deployment. Due to a lack of standards and low volume production, the average customer premises equipment (CPE) cost can be as

expensive as \$5,000 and a base station as costly as \$100,000. Combined with the range limitations, all these factors inhibited LMDS from widespread use even in urban regions. The bankruptcy of LMDS carriers such as WinStar, Teligent and Advanced Radio Telecom in September 2001 made this situation even worse. LMDS is still in use in various places, but mainly in urban areas as an easily installed wireless backhaul to other networks. Clearly, LMDS is not a solution to wireless broadband for rural areas either.

Another wireless broadband technology, multi-channel multipoint distribution systems (MMDS), often referred to as "wireless cable," operates in the licensed 2.5GHz to 2.69GHz spectrum. MMDS was intended for use as pay television service. In 1998 the FCC allowed the use of MMDS for two-way digital transmission. Operating at a lower frequency than LMDS, the range of MMDS can reach as far as 35 miles, with a speed ranging from 512 Kbps to 5Mbps. Sprint and WorldCom were once the two principal MMDS license holders. MMDS was once conceived as the solution to bridge the long last mile. However, the line-of-sight requirement of MMDS systems inherently limited the number of customers that could get service and also increased the installation cost, as it is necessary to install an antenna outside. The average CPE cost for MMDS can't go down without technology standardization and multiple vendors volume building on one standard. MMDS was projected to grow quickly in Latin America, Mexico, and in some Asian countries, but mainly as a substitute for cable television systems. In the U.S., the competition from DSL, cable and satellite services is very strong. Hence, MMDS-based broadband in the U.S. has achieved very limited market penetration up to now. Unless non-line-of-sight (NLOS) MMDS equipment can be developed in the near future and more standardization work can be done, the fate of MMDS is not optimistic.

Recently, cellular systems with third-generation (3G) or 3.5G technologies have begun to provide high-speed broadband data service overlaid on the voice networks. The two most representative technologies/standards are EV-DO (Evolution Data Only/Evolution Data Optimized) and HSDPA (High Speed Downlink Packet Access).

EV-DO is based on CDMA2000 1xRTT (single channel Radio Transmission Technology) standard. CDMA2000 1xRTT is a version of CDMA2000 radio technology operating in a pair of 1.25-MHz radio channels, sometimes also called 2.5G technology, which has a peak data rate of 144Kbps. EV-DO is an evolution of CDMA2000 1xRTT with high data rate capability added. EV-DO Rev 0 is currently available in most metropolitan U.S. cities with a peak downlink rate of 2.0Mbps and a peak uplink data rate of 154 Kbps. Verizon and Sprint are the two big carriers whose cellular networks use this technology and they are both currently upgrading their service to Revision A, which has a peak rate of 3.1Mbps on the down link and a peak rate of 1.8Mbps on the uplink.

HSDPA, sometimes called 3.5G, increases the data rates and capacity of cellular systems based on WCDMA/UMTS (Wideband CDMA / Universal Mobile Telecommunications System) technologies. From 2G GSM (Global System for Mobile Communications), 2.5 G EDGE (Enhanced Data rates for GSM Evolution) or EGPRS (Enhanced General Packet Radio Service) added support for high speed data packet service, with a downlink data rate of 384Kbps. WCDMA is an evolution of GPRS with a downlink data rate up to 2Mbps. HSDPA achieves an even higher data rate by using adaptive modulation and coding and fast packet scheduling. It also introduces antenna array technologies such as beamforming and Multiple Input Multiple Output (MIMO).

Current HSDPA deployments support 1.8 Mbps, 3.6 Mbps, 7.2 Mbps and 14.4 Mbps on the downlink. As of April 15 2007, 104 HSDPA networks have commercially launched mobile broadband services in 54 countries [1.21]. AT&T/Cingular has provided HSDPA service to major metropolitan areas in U.S. and is expanding rapidly its HSDPA networks with real world transmission speeds of 400-700Kbps. HSDPA supports simultaneous voice and data and has better international roaming capability than EV-DO technology.

Currently 3G/3.5G cellular systems are a very promising approach to provide high speed broadband service, especially to mobile users. However, high speed cellular data service is with high price. In the U.S., AT&T/Cingular and Verizon Wireless offer unlimited high speed data plans for \$80 per month. And that price does not include cell phone voice service and the cost for a cell data card needed at the customer site. Sprint offers unlimited laptop connection service for \$60 per month. T-mobile offers a total Internet package plan (unlimited data plus unlimited Wi-Fi hotspot service) for \$50 per month, but its data speed is a little bit lower, with a 144Kbps effective downlink connection speed. Its networks now are still mainly based on the 2.5G EDGE standard, but it will also roll out its HSDPA service later in 2007.

3G/3.5G cellular data services are now generally available in metropolitan areas and major cities in the U.S. Generally 3G operates at different frequencies from 2G, such as in Europe. The cost to upgrade 2G cellular systems to 3G/3.5G systems is very high because the upgrading of equipment is expensive and also because of the possible licensing fees for additional spectrum. In rural and sparse areas, the coverage of existing cellular systems is limited. Extending 3G cellular data service to rural and sparse areas where there is no existing 2G infrastructure is cost prohibitive. For an example, to cover the

whole state of Maine, it will cost roughly \$55 million just to build enough cell towers [1.22]. Without a broad subscriber base as in metropolitan areas to provide the revenue needed to offset the investment, broadband through 3G cellular system in rural and sparse areas is not going to happen in the near future without state or government support.

Challenges and Opportunities

Bringing broadband wireless service to rural areas is financially challenging. The cost of delivering broadband services increases substantially outside of large urban areas, which in turn greatly increases the risk of investment in broadband wireless for rural areas. However, today's new wireless technologies embodied in such commercial products as Wi-Fi, WiMAX and smart antennas show considerable promises for cost effective wireless broadband service. Technology standardization stimulates competition and also drives the equipment costs down. The use of unlicensed spectrum saves operators a substantial amount of money in spectrum license costs. All these factors contribute to the feasibility of a higher success rate in bringing broadband service to rural and sparse areas.

Recent Progress in Wireless Technologies

In the following section, some of the representative new wireless technologies, including 802.11x, Wi-Fi, WiMAX, MANET, are reviewed.

802.11x

802.11x standards refer to a family of specifications developed by the IEEE for wireless LAN technology [1.23]. 802.11 standards specified a medium access control

(MAC) layer and three different physical (PHY) layers: Infrared (IR) operates in the optical spectrum using infrared light, two radio-based technologies for operation at 2.4GHz band, namely, Direct Sequence Spread Spectrum (DSSS) and Frequency Hopped Spread Spectrum (FHSS). All physical layers support 1-2 Mbps data rate. Later, several extensions were made to the earlier 802.11 standards, the most popular ones are those defined by the *a*, *b*, *g* amendments made to the original standard. The extension IEEE 802.11a applies orthogonal frequency division multiplexing (OFDM) modulation as part of the PHY specification, operating at the 5 GHz U-NII frequency band and the data rates range from 6 Mbps to 54 Mbps [1.24]. The second extension, IEEE 802.11b, also referred to as high rate DSSS (HR/DSSS), defines a set of PHY specifications operating within a 2.4 GHz ISM frequency band with a data rate up to 11 Mbps [1.25]. Another extension, IEEE 802.11g, works in the same 2.4 GHz band as 802.11b but employs the OFDM modulation scheme as used in 802.11a to obtain a maximum raw data rate of 54 Mbps, and is backwardly compatible with 802.11b [1.26].

Wi-Fi

Wi-Fi, short for wireless fidelity, refers to a wireless local area network based on the 2.4GHz 802.11b/g standards. Applications of Wi-Fi include private using homes and offices, public use as hotspots deployed in hotels, airport, campus and other public places. The potential of using Wi-Fi as a last mile solution, providing users with wireless broadband access to the Internet in metropolitan areas, has already been shown in numerous local fixed networks offering Wi-Fi wireless access. And the possibility of using Wi-Fi to create a truly unwired network in metropolitan-scale areas by using

dense-cellular Wi-Fi mesh architecture has also been discussed [1.27].

WiMAX

WiMAX, Worldwide Interoperability for Microwave Access, is a wireless technology based on IEEE 802.16 standards that provides high-throughput broadband connections over long distances. It specifies the air interface and PHY layer compliance between 10 and 66 GHz, which is ideal for the point to multipoint (PMP) access with data rates in excess of 120Mbps over a radius of up to 50 km [1.28]. The extension 802.16a adds support for the 2-11 GHz range. One very attractive part of the extension is its ability to support near-LOS and non-LOS (NLOS) performance in the license-exempt frequencies below 11 GHz (primarily 5–6 GHz) [1.28]. WiMAX has the potential of being used as a wireless backhaul link for rural and mountainous areas.

MANET

MANET (mobile ad-hoc networking) technology enables an autonomous system of mobile nodes. The challenges on MANET technologies are routing, security and reliability, quality of service (QoS), internetworking and power consumption [1.29]. Most research in MANET has been focused on efficient routing protocol design for fully connected ad hoc networks. The IETF Mobile Ad Hoc Network Working Group has been formed with the purpose to “standardize IP routing protocol functionality suitable for wireless routing application within both static and dynamic topologies with increased dynamics due to node motion or other factors” [1.30]. However, the routing strategies for partially connected ad hoc networks in rural and sparse area are less explored.

Need for the Research

The development of 802.11 standards, the widespread vendor adoption, and the low cost and convenience of use, have driven Wi-Fi hotspots networks to sprout like mushrooms in big, urban high user density areas for broadband wireless access. The standardization of 802.16 (WiMAX) provides a powerful alternative for the once expensive wired backhaul link with long range, high capacity wireless radio link. It is highly possible that metropolitan-scale Wi-Fi mesh networks combining Wi-Fi local access with WiMAX wireless backhaul link will be built as the broadband wireless access in urban high density areas. However, these solutions do not readily apply to sparse and rural domains where distance, rugged terrain and low node density are key factors that affect the system design and performance. The conventional single hop, point-to-multipoint hotspot model is suitable for networks that are capacity limited, while the multi-hop network model might be more cost effective when the networks are coverage limited. The trade-off between node density and traffic per node achieved under different cell models needs to be explored before a reasonable solution can be offered to rural and sparse areas. Specifically, more research work is needed to figure out the cost benefits of applying the emerging technologies and architectures, such as high gain antennas, adaptive array antennas and multi-hop networking.

The improvement in MANET technology and the ever-increasing safety requirements have turned vehicle safety communications into a hot research topic [1.31-1.36]. Vehicle safety communications can be broadly categorized into inter-vehicle communication and vehicle-with-infrastructure communications. In rural areas, there is

generally no conventional communication infrastructure. Mobile communication systems, such as a traveler information system designed for rural and sparse areas, may be based on inter-vehicle communications rather than fixed infrastructure as in denser areas. The existing routing algorithms designed for mobile ad hoc networks consider either the high node density case or where fixed infrastructure is taken as part of the network. In some cases, location information, derived using GPS is leveraged to help the routing procedure. It is challenging to provide communications services for a mobile ad hoc network in rural areas with low node density, non-existent communication infrastructure, no GPS information system and the nodes moving at high speed. Routing algorithms that take all these factors into account haven't been defined yet. Therefore, more research is needed to study the effects of mobility and node density on routing protocol design and performance. In addition to taking advantage of the existing routing algorithms, the characteristics of mobile ad hoc network in rural areas and their effects on routing protocol design need to be further explored before a routing protocol specifically suitable for this high mobility and sparse node density scenario can be designed.

Scope of the Study

The general purpose of this research is to explore the potential offered by emerging wireless techniques and to develop new networking approaches for fixed and mobile high-speed communications services suitable for rural areas. One focus is to develop network models based on existing and emerging technologies that are suitable for fixed wireless broadband access for rural and sparse areas, including theoretical baseline network model development, realization and multi-hop cell model development for

wireless LAN. The cost benefits of several emerging technologies and innovative approaches are investigated as well. Another focus is to develop a routing protocol suitable for very sparse networks where end-to-end connectivity may be intermittent. Under these conditions, conventional ad hoc network protocols perform poorly. This work has resulted in the creation of a new Border node Based Routing (BBR routing) scheme that can be effectively applied to mobile ad hoc networks for rural and remote areas with low node density. In addition to the routing protocol design, a Geographic and Traffic Information based mobility model (GTI mobility model) for mobile ad hoc networks is developed as well for routing protocol performance simulations. The emphasis in this work has been to address vehicular-based nodes where motion is constrained to roadways, but the BBR method is generally applicable. The routing method performs similarly to conventional ad hoc network protocols such as dynamic source routing (DSR) as the node density increases.

Organization of the Dissertation

After the introduction and the need for the research is explained in Chapter 1, the following two chapters (Chapters 2 and 3) focus on fixed wireless communication and its application in rural and sparse areas.

In Chapter 2, a baseline network model is developed and then the cost benefits of emerging technologies and approaches are studied and determined. Chapter 3 presents the multi-hop cell model development for wireless Local Area Networks (LAN). The network capacity for a multi-hop network is explored and the domain where the multi

hop has advantage over one hop model is investigated as well.

Chapter 4 briefly reviews the current situation of mobile wireless communication in rural areas. Also, it highlights the challenges for providing wireless communication to mobile users in rural areas.

Chapters 5, 6 and 7 focus on mobile wireless communication and its applications in rural and sparse areas. The GTI Mobility model for mobile ad hoc networks is developed in Chapter 5. The BBR routing algorithm designed for partially connected mobile ad hoc network is described in Chapter 6. The routing scheme performance evaluation under different network connectivity scenarios using different mobility models are given and discussed in Chapter 7.

References

- [1.1] Online, <http://www.internetworldstats.com/emarketing.htm>.
- [1.2] Online, <http://www.c-i-a.com/pr1105.htm>.
- [1.3] Online, <http://www.internetworldstats.com/am/us.htm>.
- [1.4] Enrique De Argaez, "Broadband Usage in USA", <http://www.internetworldstats.com/articles/art030.htm>.
- [1.5] Federal Communications Commission, Wireline Competition Bureau, Industry Analysis and Technology Division, "High-Speed Services for Internet Access: Status as of December 31, 2005", Page(s): 24, July 2006.
- [1.6] John B. Horrigan, "Home Broadband Adoption 2006", <http://www.pewinternet.org/>.
- [1.7] Online, http://www.fcc.gov/Bureaus/Common_Carrier/Reports/FCC-State_Link/IAD/hspd1205_tables.xls
- [1.8] W. Webb, "Broadband Fixed Wireless Access as a Key Component in the Future Integrated Communications Environment," *IEEE Communications Magazine*, Volume 39, Issue 9, Page(s): 115-121, September 2001.
- [1.9] D. Gesbert, L. Haumonte, H. Bolcskei, R. Krishnamoorthy and A. Paulraj, "Technologies and Performance for Non-Line-of-Sight Broadband Wireless Access Networks", *IEEE Communications Magazine*, Volume 40, Issue 4, Page(s): 86-95, April 2002.
- [1.10] K. Nidaira, T. Shirouzu, M. Baba, K. Inoue, "Wireless IP access system for broadband access services," *2004 IEEE International Conference on Communications*, Volume 6, 20-24 June, 2004, Page(s): 3434 – 3438.
- [1.11] L. A. Ims and H. Loktu, "Extending the Broadband Service Footprint Beyond the Urban and Highly Competitive Areas by Hybrid Broadband Access Solutions", ISSLS-2002, 72-81.
- [1.12] B. Kim, N.K. Shankaranarayanan, P.S. Henry, K.P. Schlosser, T.K. Fong, "The AT&T Labs broadband fixed wireless field experiment," *Wireless Communications and Networking Conference, WCNC*, 21-24 Sept. 1999, vol.2, Page(s): 747 – 751.

- [1.13] IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems, IEEE Std 802.16-2004 (Revision of IEEE Std 802.16-2001).
- [1.14] Nojeong Heo, P.K. Varshney, "Energy-efficient deployment of Intelligent Mobile sensor networks," Part A, *IEEE Transactions on Systems, Man and Cybernetics*, Volume 35, Issue 1, Jan. 2005 Page(s): 78 – 92.
- [1.15] J. M. Kahn, R. H. Katz and K. S. J. Pister, "Mobile Networking for Smart Dust", *ACM/IEEE Intl. Conf. on Mobile Computing and Networking (MobiCom 99)*, Seattle, WA, August 17-19, 1999.
- [1.16] R. Morris, J. Jannotti, F. Kaashoek, J. Li and D. S. J. De Couto, "CarNet: A scalable ad hoc wireless network system," *9th ACM SIGOPS European Workshop: Beyond the PC: New Challenges for the Operating System*, Kolding, Denmark, Sep. 2000.
- [1.17] A. Ebner and H. Rohling, "A self-organized radio network for automotive applications," *Conf. Proc. ITS 2001, 8th World Congress on Intelligent Transportation Systems*, Sydney, Australia, and Oct. 2001.
- [1.18] Online, <http://wireless.fcc.gov/basicex/>.
- [1.19] Online, <http://biz.yahoo.com/ic/51/51331.html>.
- [1.20] Online, <http://www.tech-faq.com/lmds-local-multipoint-distribution-service.shtml>.
- [1.21] Online, <http://www.gsacom.com/news/statistics.php4>
- [1.22] Katie Zezima, "Rural living, high-speed Internet access?", *The New York Times*, December 19, 2005
- [1.23] IEEE Std. 802.11, "Wireless LAN Media Access Control (MAC) and Physical Layer (PHY) Specifications," 1999.
- [1.24] IEEE Std. 802.11a-1999, "wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: high-speed physical layer in the 5 GHz band."
- [1.25] IEEE Std. 802.11b-1999, "Wireless LAN Medium Access Control (MAC) And Physical Layer (PHY) Specifications: Higher-speed Physical Layer Extension in the 2.4 GHz Band."
- [1.26] IEEE Std. 802.11g-2003, "Further Higher-Speed Physical Layer Extension in the 2.4GHz Band."

- [1.27] M. Audeh, "Metropolitan-scale Wi-Fi mesh networks," *Computer*, Volume 37, Issue 12, Dec. 2004, Page(s): 119 – 121.
- [1.28] IEEE Std. 802.16-2004 (Revision of IEEE Std 802.16-2001), "IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems."
- [1.29] Jun-Zhao Sun, "Mobile ad hoc networking: an essential technology for pervasive computing", *2001 International Conferences on Info-tech and Info-net, Proceedings. ICII 2001 - Beijing*. Volume 3, 29 Oct.-1 Nov. 2001, Page(s): 316 – 321.
- [1.30] Online, <http://www.ietf.org/html.charters/manet-charter>.
- [1.31] Jason LeBrun, Chen-Nee Chuah, Dipak Ghosal, Michael Zhang, "Knowledge Based Opportunistic Forwarding in Vehicular Wireless Ad Hoc Networks," *IEEE VTC Spring 2005*, Volume 4, Page(s): 2289-2293.
- [1.32] B. Xu, A. Ouksel, and O. Wolfson, "Opportunistic resource exchange in inter-vehicle ad-hoc networks," *IEEE Int. Conf. Mobile Data Management*, Berkeley, CA, Jan. 2004, Page(s): 4-12.
- [1.33] L. Wischhof, A. Ebner, H. Rohling, "Information dissemination in self-organizing intervehicle networks", *IEEE Transactions on Intelligent Transportation Systems*, Volume 6, Issue 1, March 2005, Page(s): 90 – 101.
- [1.34] L. Briesemeister, G. Hommel, "Role-based multicast in highly mobile but sparsely connected ad hoc networks Mobile and Ad Hoc Networking and Computing," *First Annual Workshop on MobiHOC*, 11 Aug. 2000, Page(s): 45 – 50.
- [1.35] R.A. Santos, R.M. Edwards, A.E. Edwards, D. Belis, "A novel cluster-based location routing algorithm for inter-vehicular communication," *15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 2004. PIMRC 2004. Volume 2, 5-8 Sept. 2004 Page(s): 1032 – 1036.
- [1.36] C. Lochert, H. Hartenstein, J. Tian, H. Fussler, D. Hermann, M. Mauve, "A routing strategy for vehicular ad hoc networks in city environments," *Intelligent Vehicles Symposium, 2003*. Proceedings. IEEE, 9-11 June, 2003, Page(s): 156 – 161.

CHAPTER 2

BASELINE NETWORK MODEL FOR BROADBAND
FIXED WIRELESS ACCESS IN RURAL AREASIntroduction

The use of wireless, DSL and cable for broadband access has become increasingly prevalent in metropolitan areas. While these technologies are being successfully utilized both in terms of service quality and economics in densely populated areas, there are still vast geographic regions where broadband services are either prohibitively expensive or simply unavailable at any price.

Broadband adoption in rural and remote areas is relatively low. Recent data shows that among farmers and ranchers in Montana, one of the least densely populated states, in 2005 70% have Internet access and over 75% have computers [2.1]. However, in spite of demand for higher grades of service, these users and others in rural areas are served predominantly by low-speed dialup access. Dialup was the most common method of accessing the Internet with 69 percent of U.S. farms [2.1].

Advances in wireless technology and system-level applications for data communications, now embodied in the 802.11 standards and commercialized by numerous suppliers as “Wi-Fi”, open new possibilities for broadband fixed wireless access [2.2]. Public use of Wi-Fi is emerging as “hot spots” are deployed in hotels, airports, coffee shops and other public places. The potential of using Wi-Fi to extend Internet access to less densely populated and isolated areas is being explored by several entrepreneurial service providers. These companies typically utilize Wi-Fi for last mile

access and some form of radio link for back haul as well. The proliferation of Wi-Fi has resulted in significant reductions in equipment costs, with the majority of new laptop computers now being shipped with Wi-Fi adapters built in.

Based on the above facts, Wi-Fi is taken as a viable technology for fixed wireless Internet access and the possibilities of using Wi-Fi cost-effectively in sparsely populated and rural areas are explored in this chapter. A “life cycle” approach, considering both capital and operations costs over a prolonged study period is used. The baseline network model takes into account market demand and segmentation, customer adoption and retention (churn) rates, service pricing, as well as technology-based factors such as range, bandwidth, antenna gain, receiver sensitivities and signal strengths. To make the assessment realistic, a real rural area: the Gallatin County of Montana is modeled. The entire county, including the major towns, satellite communities, and widely scattered rural population, is modeled to have broadband access.

In the following sections of this chapter, first the approach and assumptions adopted to study the baseline network model are discussed in detail. Secondly, the baseline network model is proposed. Then the economic viability of this network is evaluated based on realistic geographic areas and economic parameters. After that, the model is varied by incorporating several technical and architectural alternatives, such as high-gain antennas, beam forming and packet switching technology, and multi-hop networking to assess the cost benefits of innovative approaches.

Approach and Assumptions

In this section, the approach and the network design tool used to study the baseline

network model is briefly reviewed. Then the geographic area chosen to study the baseline network model is presented with detailed demographic information. After that some assumptions made to model the network, including data traffic, user penetration, and market share for wireless service providers and customer equipment cost are discussed.

Approach and Wireless Network Design Tool

Strategic Analysis Tool (SWAT) for Wireless Internet Service Providers (WISPs) [2.6], simply called WirelessSWAT or SWAT, is used to model the network structure and evaluate the economics of wireless networks. SWAT is a useful tool to develop network and service plans including detailed technical and business analysis.

To build a network model in SWAT, several steps are necessary. First, the study area is selected and divided into one or multiple canonical areas (CAs). A CA is the smallest planning area of uniform characteristics (e.g., demographics, technology, propagation, terrain, etc.). Second, the customer segments are defined in terms of user traffic, population growth rate, market penetration, service types and CPE (Customer Premise Equipment) cost. Service pricing plans and the price for different services must also be defined. The network architecture must be defined and the network technologies, including details for node and link parameters must also be selected. Finally, study parameters, including study period, capital costs for licenses and network components, and recurring expenses for network maintenance and management must be established. Figure 2.1 shows the relationship among the modeling steps and parameters.

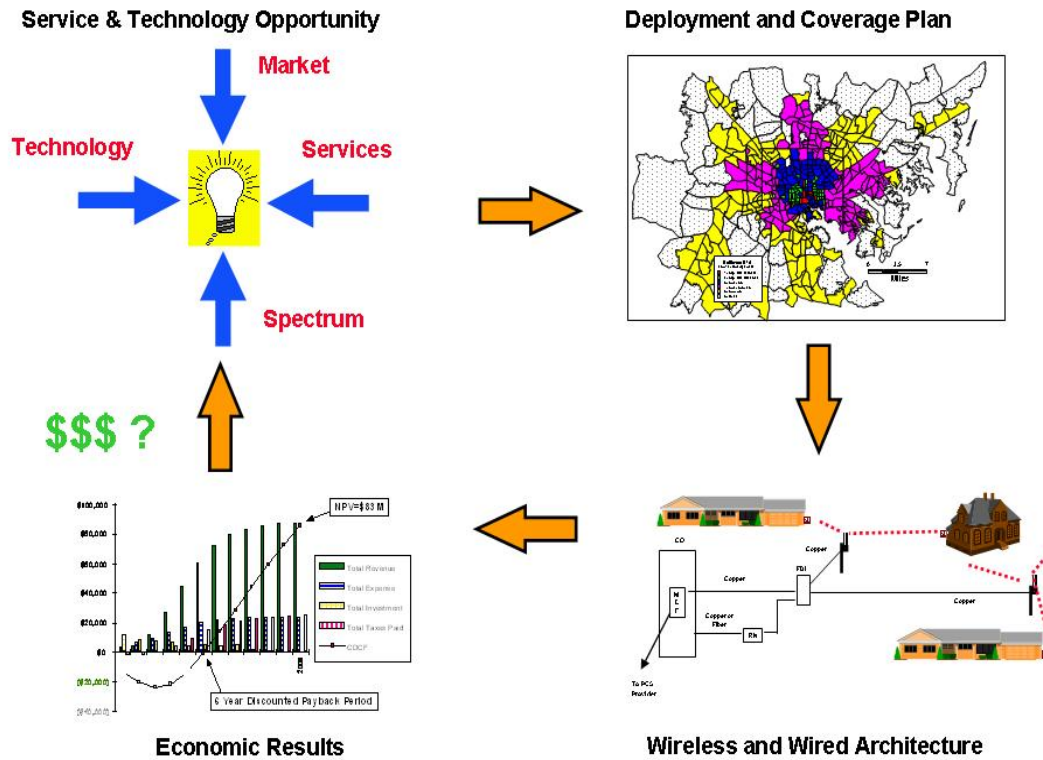


Figure 2.1. SWAT parameters and modeling flow.

The network model built in SWAT combines the capital, installation and operations costs for the wireless network elements, the backhaul links, switches, routers and connection to the Internet. It also includes the cost and expense of the operations support systems (OSS) necessary to provision, maintain and manage the network. The model calculates the engineering parameters: e.g., the total subscriber traffic, the number of nodes, links and facilities needed to cover the whole service area with sufficient capacity to meet the service demand. The model then calculates the multi-year financial results, e.g., net present value (NPV), return of investment (ROI), cash flow, years to break even, etc., over the study period using the equipment capacities, demographics, service

definitions and assumed values for market share, market penetration and pricing. The results are not a detailed engineering plan for the network. Rather, the model demonstrates the feasibility and rationality of the wireless network structure based on the financial and market assumptions combined with technology based engineering calculations. Different technologies, services, penetration rates, pricing, etc. can be used to evaluate the sensitivity of the results to the overall network design and to the various service, usage, and financial assumptions.

Geographical Areas Chosen for Baseline Network Model Study

Gallatin County, Montana has a population of 71,000 and a land area of over 2600 square miles, roughly twice the size of the state of Rhode Island [2.3]. In this study Gallatin County is chosen as a typical area, as it includes a densely populated urban area, suburban areas, small communities and a large sparsely populated rural area consisting of farms and ranches. The total service area is categorized into 8 canonical areas (CAs) according to their demographic characteristics: each city or town is taken as an independent CA, all the other rural areas excluding the land covered by national forest are grouped as another CA. Current census data and projected growth rates are given in Table 2.1[2.4].

Table 2.1. Demographic Information for the Cities and Towns of Gallatin County

Canonical area (CA)	Resident population 2000 census	Land area, excluding national forests (sq mi)	Population density (persons/mi ²)	Average annual growth rate (%) 1990 to 2000
Amsterdam-Churchill	727	4.1	178	1.9*
Belgrade city	6893	7.2	958	5.0
Bozeman city	31591	13.1	2410	1.9
Four Corners	1828	10.2	179	1.9*
Manhattan town	1396	0.6	2289	3.0
Three Forks city	1728	1.3	1361	3.7
West Yellowstone	1177	0.8	1453	2.6
Rural area	22491	1580.2	15	1.9*
Total	67831	1617.5	41.9	1.9

(*: These data were not given in the Census 2000, the average annual growth rate for the whole Gallatin County was used instead.)

Data Traffic and User Penetration Assumptions

In each CA, two customer segments are defined according to different data service types provided. One customer segment is assumed to use low-speed data service and the other customer group is assumed to use high-speed data service. The two wireless data service categories have asymmetrical upstream and downstream throughputs. Low-speed data service: 512 kbps asynchronous, 5% duty down, 1% duty up, 5 sessions/hour, 100 seconds /session; High-speed data service: 1 Mbps asynchronous, 5% duty down, 1% duty up, 5 sessions/hour, 100 seconds /session. The resulting average daily Internet traffic per user is about 0.37 Giga bits and about 0.72 Giga bits for low-speed and high-speed respectively. There is very little data available characterizing general public Internet usage. Dartmouth College reports that daily usage of their campus-wide Wi-Fi wireless

network averaged 0.32 Giga bits per user in 2001 [2.5]. The assumption for the traffic load is reasonable and conservative, anticipating that people will develop increased Internet use as time goes on. The high penetration of computers in rural areas also suggests that higher levels of Internet access demand and usage should be anticipated. The percentages of high or low-speed data service required by the potential users are assumed to be 20% and 80% respectively, and this ratio is varied in a sensitivity analysis. The ratio of total customers who subscribe to a particular service to the potential users (total population) at a point of time is described using a Fisher-Pry (S-shaped) model as indicated in Figure 2.2, assuming an initial penetration of 20%, final penetration of 80%, and a period to half of the final penetration of 3 years. This adoption rate is similar to that of cell phone users.

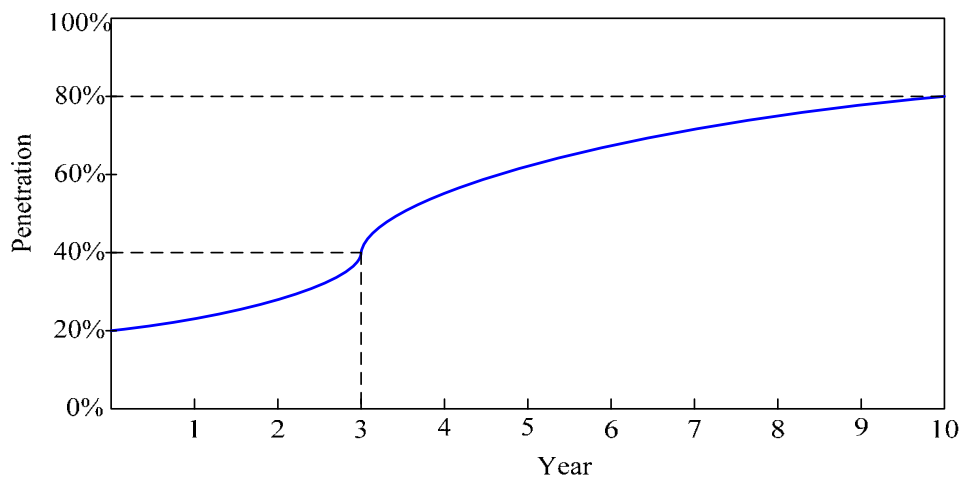


Figure 2.2. The Fisher-Pry (S-shaped) model for data service subscriber penetration.

Market Share for Wireless ISPs

These services could potentially be provided by several competing wireless Internet service providers (ISPs). The portion of service provided by a particular wireless ISP is

modeled using a parameter to characterize market share. It is likely that few wireless ISPs would compete in an area with rural demographics. Hence a relatively high market share of 75% is assumed in the small communities and rural areas. In densely populated areas such as Bozeman, DSL may be available and the wireless ISP market share may be less. The sensitivity of the results to this factor is tested later in the discussion section. Also assume that in each year, 5% of the subscribers leave the system (churn) and are immediately replaced in addition to any overall market growth.

Customer Premises Equipment (CPE) Cost

The Customer Premises Equipment (CPE) cost for data services is for the Wi-Fi network adapters, which are now readily available for about \$50~\$150, and potential external antennas. In the urban areas, end users may be able to access the service without external antennas. In the more rural areas, however, it is anticipated that some users will need antennas mounted outside the premises, which would add approximately \$200 to the CPE cost. In the analysis, it is also assumed that all CPE costs are borne by the consumers.

Baseline Network Model

This section gives the proposed baseline network model structure for broadband fixed wireless access (BFWA) in rural areas. Some basic components of the network model are briefly reviewed. And some technologies used in the baseline network model are further specified.

Baseline Network Model Structure

The proposed baseline network model structure is indicated in Figure 2.3. In this network model, the Access Point (AP) is a device that "connects" wireless communication devices together to create a wireless network. End users connect to APs using wireless link based on 802.11 technologies, which is widely tested and applied in local wireless area networks in urban areas. The connection of the APs to Ethernet access switches can be via wireless backhaul links. Wired or wireless links can be used to connect the Ethernet switches to a central router, and finally to the Internet. Additionally, an AP management system is included (not shown in the figure) to support the management and operation functions of APs. In the following, some basic components of the network model, including AP, wireless Ethernet bridge and x Base T lines are briefly reviewed.

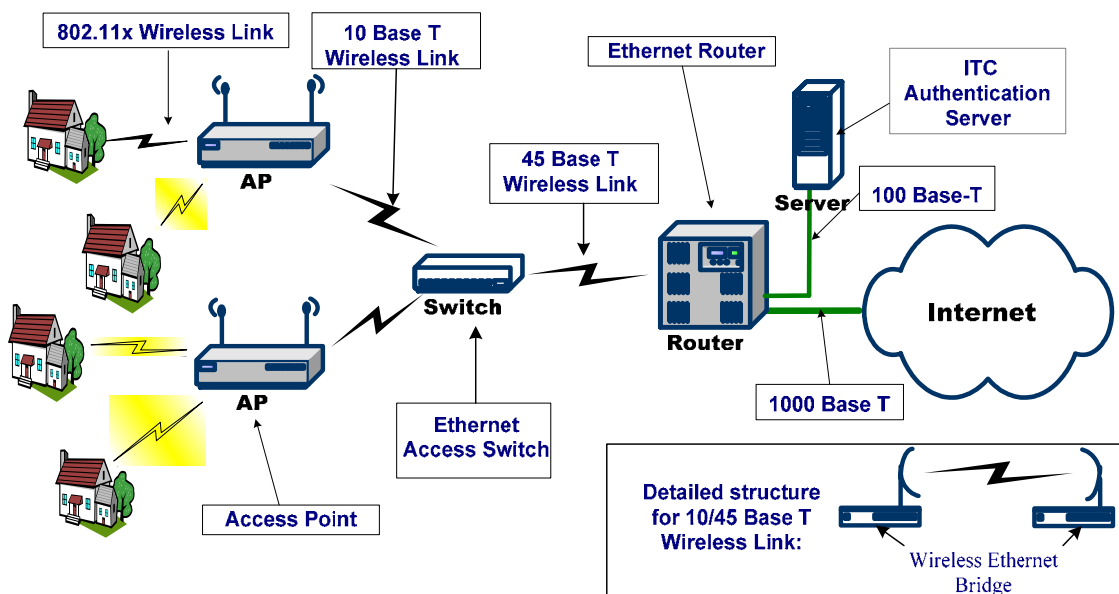


Figure 2.3. The baseline network structure for BFWA in rural areas.

Access Point (AP). AP is defined in the IEEE 802.11 standard as “Any entity that has station functionality and provides access to the distribution services, via the wireless medium (WM) for associated stations” [2.7]. The AP in a basic service set (BSS) is analogous to the base station in a cellular phone network. When an AP is present, stations do not communicate on a peer-to-peer basis: all traffic goes through the AP. APs are generally fixed and form part of the wired network infrastructure. In the baseline network model, APs serve the end users in a point to multi-point configuration, interconnected to Ethernet switches or routers using wireless Ethernet bridges.

Wireless Ethernet Bridge. The wireless Ethernet bridge provides a way to connect wireless LANs to larger, wired Ethernet networks. In the baseline model, two wireless Ethernet bridges work as a pair through point to point wireless link. One wireless Ethernet bridge at wireless LAN side connects with the AP, which interfaces with wireless end users. Another wireless Ethernet bridge at the wired side connects directly to an Ethernet hub or switch that further connects to the wired Internet. These wireless Ethernet bridges are quick to set up and easy to configure, make them a practical way to extend the range of the existing wireless LANs.

X Base T. According to 802.3 standards, X Base T refers to a series of twisted-pair cables used for Ethernet connection in Local Area Networks (LANs). X refers to different level of data transmission speed. 10 Base T refers to 10 Mbps baseband Ethernet over twisted pair cables with a maximum length of 100 meters. 100 Base T refers to 100 Mbps baseband Ethernet over twisted pair cable. 1000 Base T refers to 1000 Mbps baseband Ethernet over four pairs of Category 5 unshielded twisted pair cable. These wired lines

are generally used for short range connections. The term “10/45 Base T wireless link” refers to the use of 10 Mbps or 45 Mbps baseband Ethernet over a wireless link respectively. The wireless links generally provide much longer communication range than wired lines.

Model Parameters and Technology Specifications

In the study, it is assumed that users are uniformly distributed in each CA and adopt an average of 2Mbps link throughput. A Wi-Fi wireless link has a maximum throughput of 11 Mbps, but this high data rate requires a high signal-to-noise ratio, thus leading to a relatively small AP coverage area. Propagation models with maximum range specified according to each of the CAs are adopted. In addition, it is also assumed that each AP will serve a single site, requiring one RF channel and providing an average per user throughput of 2Mbps. These assumptions are appropriate where system requirements are dictated by the need to cover a large geographical area and when the user demand does not exceed the AP capacity. The validity of these assumptions is explored later.

In the baseline network model, APs connect to Ethernet access switches over a 5.8 GHz U-NII wireless link. Other technology, such as WiMAX, is also an attractive alternate for this backhaul link. An 802.3/100Base-T wired link is used to connect the Ethernet switches to a central router. From the router, a Gigabit wired Ethernet link is used to connect to the Internet. Additionally, an AP Management System is included (not shown in the figure) to support the management and operations functions.

The baseline network model is specifically applicable to a CA where there is an Internet point of presence (such as Bozeman). CAs without local ISP access are modeled

with the same network structure from the end users to the Ethernet access switches, and then 5.8 GHz U-NII wireless links are used to connect the access switches directly with the router located in Bozeman. A separate router in each of these areas is not necessary due to the low total user traffic load. For the rural area another layer of switches for data aggregation after the first layer of Ethernet access switches is included to span the large coverage area.

Several common financial metrics including Net Present Value (NPV), Earnings before Interest, Taxes, Depreciation and Amortization (EBITDA), Return of Investment (ROI) and years to break even are used to evaluate the long-term financial performance of the project. These metrics are further explained in the following:

NPV is the present value of net cash flows. It can be calculated using the equation (2.1) [2.8]:

$$NPV = \sum_{t=1}^n \frac{C_t}{(1+r)^t} - C_0 \quad (2.1)$$

Where

t - the time of the cash flow.

n - the total time of the project.

r - the discount rate.

C_t - the net cash flow (the amount of cash) at time t .

C_0 - the capital outlay at the beginning of the investment time ($t = 0$).

EBITDA is a measure of the pure operating cash flow of a plan. It is calculated in each period as revenue minus capital minus expenses [2.6].

ROI is a measure of the financial reward of a plan. It is calculated using the equation

(2.2) [2.6].

$$ROI = \frac{\text{Cumulative_Discounted_EBITDA}}{\text{Cumulative_Discounted_}(Capital + Expense)} \quad (2.2)$$

The performance parameters and costs of the network equipment, including both capital and expense, are based on typical, currently available commercial network-grade products suitable for outdoor use. Additional costs (e.g., land, tower for the access points, site preparation, installation, electrical power, etc.) are modeled either as one-time capital investment included at the beginning of the study period or as an annual expense.

Simulation Results

Four cases are defined and examined based on the study area defined above. Case A is the baseline case. Based on the baseline case, the cost benefits of several emerging technologies and architectures, including high gain antennas, dynamically steerable beam-forming antennas and multi-hop routing are studied in Case B, Case C and Case D, respectively.

Case A: Baseline Case

For APs in the baseline case, the outdoor propagation range at maximum transmit power setting with a 2.2 dBi gain diversity dipole antenna is 800 ft (244 m) @ 11 Mbps and 2000 ft (610 m) @ 1 Mbps. Hence an AP coverage area radius of 244 meters is used for the more densely populated CAs with high average user traffic, and a coverage range of 610 meters is used for the sparsely populated rural areas with low average traffic. Monthly per subscriber service prices of \$50/month for 512 kbps and \$75/month for 1 Mbps are assumed in the baseline case, comparable to those currently offered by local

ISPs. In the baseline model, it is assumed that coverage for the entire region is provided in the first year. However, alternatives, such as building the network in steps over the study period, are considered later in the discussion. Table 2.2 summarizes the basic cost assumptions of network equipment and the other expenses used in the simulations. The cost assumptions for access points, switches and routers are based mainly on Cisco products. The costs assumed for the wireless link equipment based mainly on Proxim Tsunami products.

Table 2.2. Network Components Cost and Expense Assumptions

System Component	Cost	
	Capital Cost/Unit (\$K)	Operating Expense (\$)
Access Point	1.275	12/month
10 Base T wireless link	23.5	20/month
45 Base T wireless link	31.5	30/month
Ethernet Access Switch	2	10/month
Ethernet Router	17.5	100/month
AP Management System	6.1	90/month
Internet Service POP connection	0	420/month
Land/Access Site	0	1000/year
Tower/Access Site	1	0
Miscellaneous/Access Site	0.5	0

Table 2.3 Case A gives the NPV for the 10-year study period and the years to break even, for each CA and for the composite area. Densely populated areas have a high ROI and a short period to break even, while the rural area has a negative NPV, indicating that it will not break even in this ten-year study period. The NPV of the composite area is seriously reduced by the investment required to serve the rural area. Figure 2.4 (a) shows the capital and expense flows of the composite area over the study period.

Table 2.3. NPV and Years to Break Even of Each CA and the Composite Area for the Three Different Cases

Case & items CAs	Case A		Case B		Case C	
	NPV (\$K)	Years to break even	NPV (\$K)	Years to break even	NPV (\$K)	Years to break even
Amsterdam Churchill	966	3	1134	1	1013	2
Belgrade	13015	1	13638	1	13629	1
Bozeman	30486	2	33475	1	33616	1
Four Corners	2272	3	2723	1	2640	1
Manhattan	2173	2	2332	1	2219	1
Three Forks	2850	2	3074	1	2972	1
West Yellow Stone	1746	2	1876	1	1755	2
Rural area	-46165	N/A	26176	3	28254	3
Composite area	7345	10	84430	2	86101	1

Figure 2.4(a) indicates that the dominant capital investment is in the APs and the wireless links from the APs to the access switches. The large number of APs means that there is also a large investment in wireless links, further raising the capital cost. After the first year, operating costs, which include leases for AP sites, dominate the expense flow, and the large number of APs drives this number as well. These results indicate that as a whole the wireless network is “coverage-limited” over the entire study period. That is, the initial investment is needed to provide the coverage for the whole area, but on average, the demand per unit area never exceeds the capacity of an AP. Architectures and deployment strategies that would reduce the number of APs, particularly in the rural areas would make the network more financially attractive.

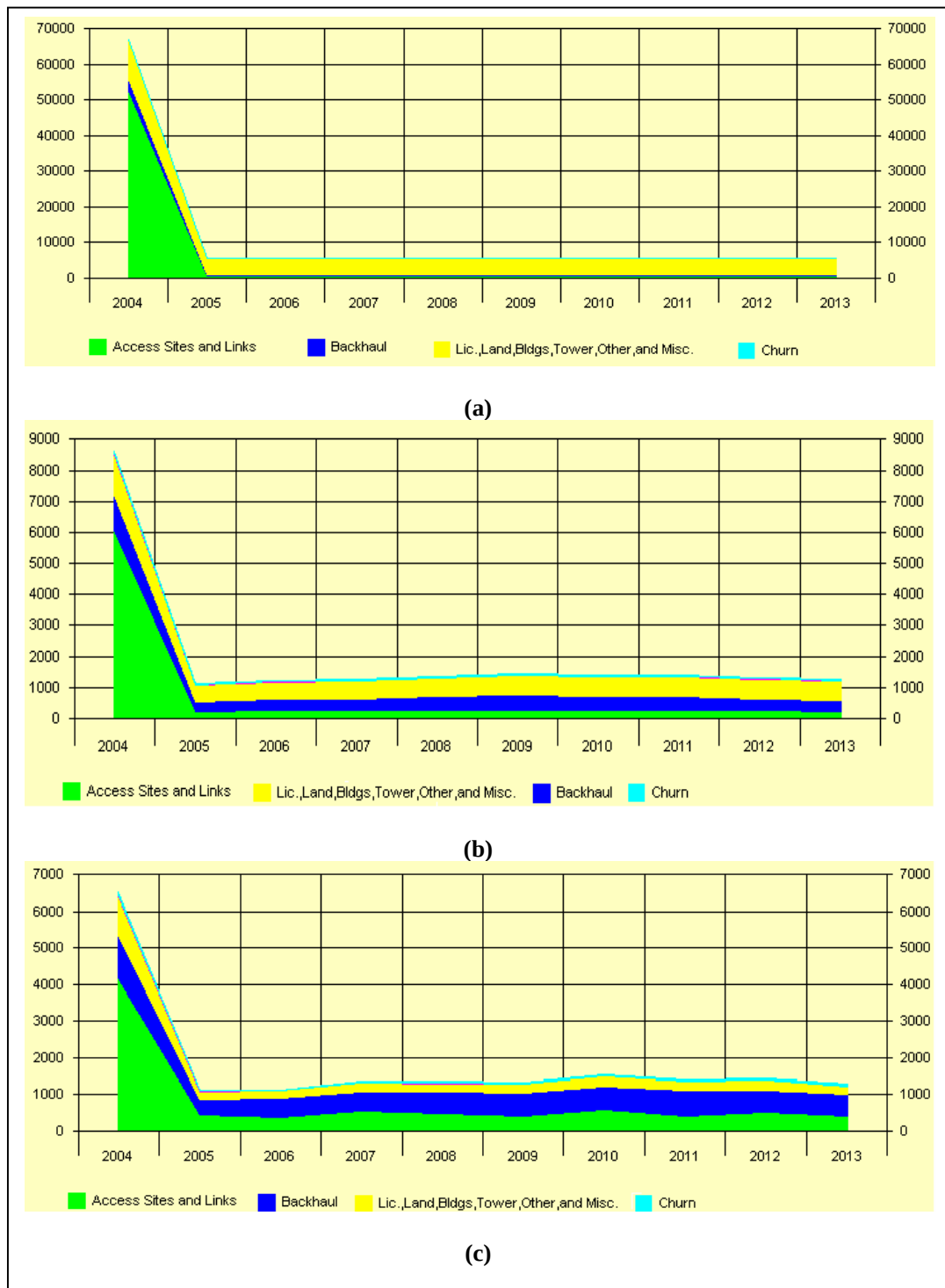


Figure 2.4. Capital and expense flows for the composite area in the ten-year study period, (a) Baseline case, (b) Case B and (c) Case C.

Case B: Baseline Case with High Gain Antenna Applied

The baseline case results show that the wireless network cost is dominated by the large number of APs needed to cover the rural area. Therefore, the alternative of providing the APs with high-gain antennas is considered to extend the coverage range, while the maximum effective isotropic radiated power (EIRP) for each AP is still limited to that prescribed by FCC regulations. For this case, a 12 dBi antenna is used with each AP, and all other assumptions are left unchanged. The cost assumptions are almost the same as indicated in Table 2.2. The additional cost assumption is for the high-gain antenna, which is assumed to be \$250. Assuming outdoor, line-of-sight propagation, the average AP range is almost tripled (The exact number would be adjusted according to terrain effects in a detailed engineering design.). Table 2.3 Case B and Figure 2.4 (b) give the financial results for this alternative case. The break-even points for the rural and composite areas are 3 years and 2 years respectively. The associated NPVs are \$26.2M and \$84.4M for the rural and composite areas over the ten-year period.

By comparing Figure 2.4 (a) with (b), it is noted that the capital investment, still dominated by APs and links, is reduced by a factor of 8. There are additional capital savings in backhaul due to the fewer number of APs as well. Even with the larger AP coverage area, the network is still coverage limited over the study period.

Case C: Baseline Case with Switched Array Antennas

The results for case B indicate the advantages of increasing the AP coverage area in this coverage-limited situation. Here a second alternative is considered by using recently introduced phased array antenna technology, combined with beam switching to increase

the range and capacity of an AP. So called “packet beam” or PacketSteering™ is being introduced for Wi-Fi networks by at least one vendor, Vivato, and shows promise in outdoor and indoor applications [2.9]. APs equipped with these switched array antennas have a maximum outdoor line-of-sight transmission range of up to 4.2km at 11 Mbps and 7.2km at 1Mbps. Each switched array antenna covers up to 100° in the azimuthal plane and up to 3 concurrent Wi-Fi beams provide connections on a packet-by-packet basis with a maximum throughput per channel of 11Mbps. Hence both increased coverage and increased capacity are achievable.

Case C is defined based on the assumptions of the baseline case. And the APs with diversity antennas in Case B are replaced with APs integrated with switched array antennas (AP/switch). Four AP/switches are needed per site and all 4 AP/switch connect to a common Ethernet hub. Wireless links are used to connect to the access switches and the remainder of the network infrastructure is unchanged, as indicated in Table 2.2. The additional cost assumption is for the AP/switch, which is assumed to be \$15,100 per AP/switch. Table 2.3 Case C and Figure 2.4 (c) give the financial results for this case. The break-even point for the rural area alone is in 3 years, while the composite area is in 1 year. The NPVs at the end of 10th year for rural and composite areas are \$28.3M and \$86.1M respectively. For the composite area, this is a factor of 10 times improvement over the baseline case and a little higher than the NPV of Case B. By comparing Table 2.3 Case B and C, it is noted that the use of high-gain antennas and switched array antennas yields similar cost benefits, but Case C has an advantage over Case B in that it needs less initial capital investment and the capital investments are more uniformly distributed over the whole study period. As the AP coverage area for cases A, B and C are

increasingly larger, the maximum subscriber density served decreases accordingly. For the assumed traffic and usage characteristics, it is found that the subscriber densities where an AP becomes capacity-limited are 53, 7 and 5 subscribers per square mile for cases A, B and C, respectively.

Figure 2.5 compares the discounted cash flows for the composite area for the three cases. The data indicate that cases B and C have similar and very attractive financial outcomes. The further increase of the capacity in APs of Case C makes it slightly more cost effective than Case B. From the point view of a real engineering design, case C requires fewer AP sites than case B, which offers further operations and management cost advantages. Hence, for this specific wireless network, the use of switched array antennas is the most cost-effective and favorable approach.

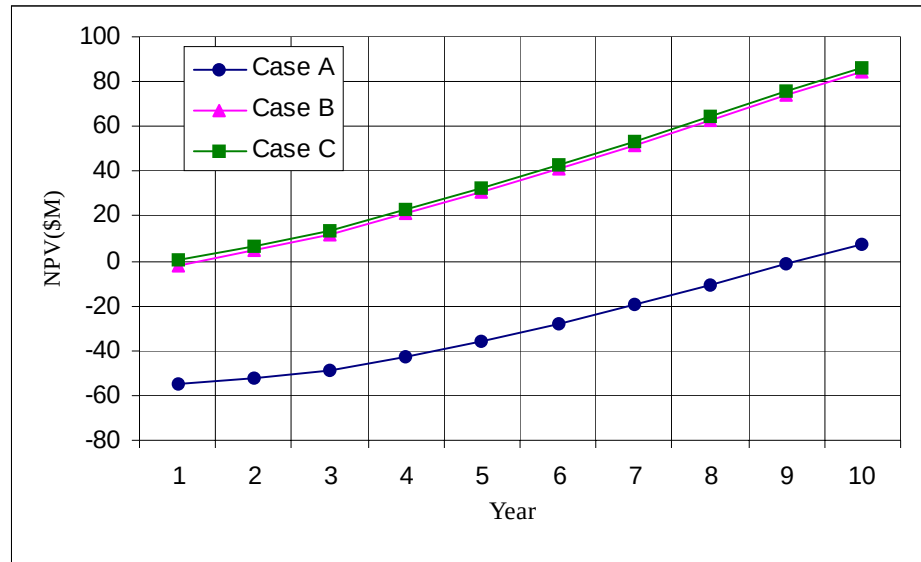


Figure 2.5. Discounted cash flow graph of the composite area for the three cases.

Case D: Multi-hop Connections

For the baseline case, the large investment in APs and associated backhaul links to

cover the low-density rural area limits the financial viability of the network. Every AP is connected to the access switch by an independent wireless link. Here an alternative to point-to-point backhaul is explored by considering the use of APs in a multi-hop configuration to handle the backhaul traffic. This approach is particularly attractive in the coverage-limited domain where the backhaul bandwidth requirement is low.

In a multi-hop network structure, two APs instead of one are installed at each AP site as defined in the baseline case. Part of the network configuration is shown in Fig 2.6. One AP is used for the usual conventional function of serving subscribers in the coverage area, while the second AP is configured to operate as a wireless Ethernet bridge to communicate with another AP or connect to the access switch over a wireless link. The tradeoff is that the number of APs is almost doubled while the number of wireless links decreases by a factor of two, and the number of access switches is accordingly lower. As the cost of a point-to-point wireless backhaul link is almost ten times higher than that of the pair of APs, the use of a multi-hop network will significantly reduce the capital investment.

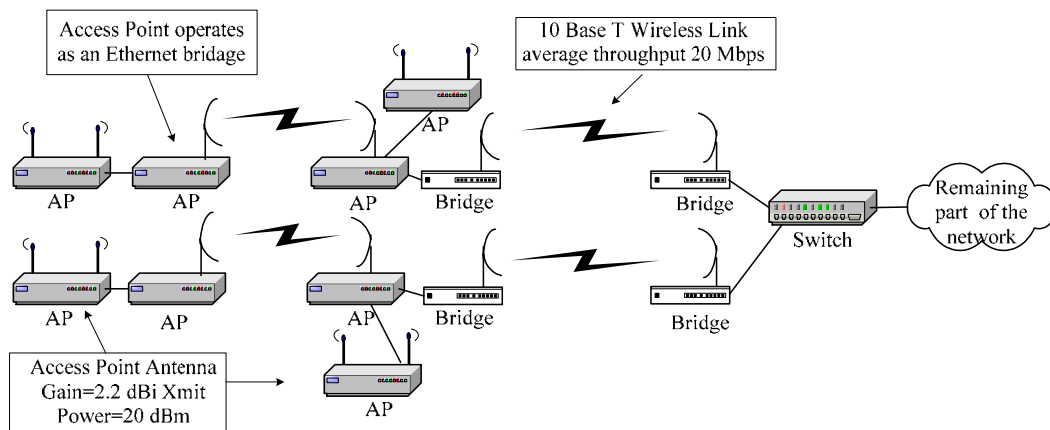


Figure 2.6. Part of the network with the multi hop network structure applied.

The AP chain can be extended (e.g., increase the number of hops), but the total number of hops will be constrained by the capacity of an AP. Alternate routing may be needed to avoid congestion and may also be used to increase network reliability [2.9]. Excessive delay (and delay variation if routing is variable) is an additional concern and may limit the applicability of this approach for real-time services. Further study is needed to establish the technical and economic viability of this approach, which is carried out in detail at Chapter 3.

Discussion

The baseline case and alternative case results indicate considerable promise in cost-effectively providing affordable, Wi-Fi based high-speed Internet access in rural areas. Here the sensitivity of these conclusions to several of the study assumptions is considered. First it is noted that the previous analysis assumes that all CPE costs are borne by the end user. This position is substantiated by the high percentage of households and businesses that already have network-ready computers and the trend to include Wi-Fi technology in new products. It is also noted that consumers are accustomed to paying directly, or through long-term contracts with rebates, for cable and DSL modems. Therefore it is reasonable to surmise that a wireless ISP can expect that consumers will agree to similar pricing approaches.

The robustness of the results to several of the major assumptions has been tested. Network equipment costs used in the model are “typical retail prices”, available from vendor web sites, and are an upper bound, as vendors normally offer discounted prices in volume sales to service providers. Static prices are also used, which ignores the trend of

more performance at lower cost that is driven by a combination of volume production. Also tested is the sensitivity of results to assumptions regarding user preference for low-speed and high-speed services. If a larger percentage of users were to subscribe to high-speed (1 Mbps) access, the traffic load would increase and more network equipment might be required, but at the same time, the revenues would increase. The impact of this trend is calculated on Case B for several alternative scenarios: 50% and 80% usage of high-speed service. The results shown in Table 2.4 indicate that the network remains coverage-limited. The increased traffic generates more revenue without requiring additional investment.

Table 2.4. Impact of Increased Traffic Load on the Financial Performance

Percentage of User \ Items		NPV (\$K)	Years to break even
1Mbps	512Kbps		
20%	80%	84430	2
50%	50%	99580	2
80%	20%	107261	2

The initial assumption that Wi-Fi access can compete effectively where DSL is available might be optimistic and skew the results. The impact of DSL is tested by excluding densely populated areas such as Bozeman from the study area. For Case A, the resulting 10-year NPV would be negative. However, for Cases B and C, the NPVs would be \$51.4M and \$52.5M respectively, indicating providing wireless Internet access to small communities and rural areas is still economically viable.

The network roll out strategy may be a key factor in assuring successful deployment.

In all our cases it is assumed that the network would be deployed quickly to provide service in the entire coverage area at the beginning of the study period. This led to large initial capital investment that might not be feasible for a start up operator. The results suggest that a phased deployment, where the network is constructed first in the densely populated areas where the break-even point is short, would be attractive, and revenues could then be used to provide the capital needed to offer coverage in the less profitable rural areas. The results for case C, where the higher population density areas become capacity limited, indicate the benefits of this approach. Such scenarios require further assessment but are beyond the scope of this study.

Finally, the sensitivity of previous findings to a key financial parameter, the discount rate is also tested. The value of this factor is determined by a combination of effects including the prevailing interest rates and the risk associated with the investment. The latter is a highly subjective consideration, and one might question the viability of a rural Internet access business and ascribe a higher discount rate to the analysis. A comparison is made on the results for Case B, by varying the discount rate between 10% and 25%. Results in Table 2.5 indicate that though the years to break even are the same, the NPV decreases significantly as the discount rate rises, but still remains positive.

Table 2.5. Effect of Raising the Discount Rate on the Financial Results

Items Discount Rate	NPV (\$K)	Years to break even
10%	84430	2
15%	63590	2
20%	48980	2
25%	38490	2

Since these models were run, there has been huge growth in the Wi-Fi area. A large number of public Wi-Fi networks have been built in mainly metro areas and service is provided by competing ISPs. According to MuniWireless's list of public broadband initiatives [2.10], by the end of March 2007, there are 340 cities or counties with citywide Wi-Fi service available. This number does not include the 49 cities and counties that are just thinking about deploying a city- or countywide Wi-Fi network. If included, the total number rises to 389. These city-wide Wi-Fi networks not only provide Internet access or web surfing, they also are used for public safety (police, wireless camera surveillance, etc.), traffic management, parking control, voice over Wi-Fi and so on. The robust growth of Wi-Fi networks in hundred of cities and towns also validates the studies discussed above and the observation that Wi-Fi based Internet access is feasible in areas where there is sufficient population density.

In metropolitan areas, Wi-Fi mesh networks are being deployed to avoid bottleneck problems by using alternative routing. However, for very sparse areas, mesh networks are not applicable because there would not be enough nodes to form a mesh network. However, for very sparse areas, where the node density is very low, the bottleneck problem does not arise and multi-hop networks can be used. The population density where multi-hop technology can be used without resulting in a bottleneck will be explored in Chapter 3.

Summary

In this chapter, the baseline network model for broadband fixed wireless access in rural areas has been proposed and investigated. The simulation results show that with reasonable assumptions for equipment costs, customer adoption rates, services prices and market share, a Wi-Fi based broadband Internet access network is financially viable in a rural area. Technology and architecture alternatives, as evidenced by study cases B, C and D can improve the cost effectiveness and financial performance.

The simulation study has been restricted to 802.11b Wi-Fi technology. Other 802.11 technologies, such as 802.11a at 5 GHz, 802.11g at 2.4 GHz and 802.16 WiMAX either between 10-66 GHz or 5-6 GHz are now standards and becoming available. The increased capacity offered by these alternatives would be applicable in high-demand areas and could be explored in a similar way as well.

The simulation results also indicate that alternative architectures, such as multi-hop networks, have promise in the rural domain but need further examination to be effectively exploited. Using multi-hop for broadband fixed wireless access in rural areas will be explored in Chapter 3.

References

- [2.1] <http://www.nass.usda.gov/mt/pressrls/misc/compuse.htm>.
- [2.2] P. Henry and H. Luo, "Wi-Fi –What's Next?", *IEEE Communications Magazine*, Vol. 40, No. 12, Page(s): 66-72, December 2002.
- [2.3] US Census Bureau Data, <http://www.census.gov/main/www/cen2000.html>
- [2.4] Census and Economical Information Center (CEIC), U.S.Census 2000(SF3).
- [2.5] D. Kotz and K. Essien, "Analysis of a Campus-wide Wireless Network", *MOBICOM'02*, Sept.23-26, 2002, Atlanta, GA.
- [2.6] Rsoft Design Group, Strategic Analysis Tool, SWAT, Version 4.1.8
- [2.7] IEEE Std. 802.11, "Wireless LAN Media Access Control (MAC) and Physical Layer (PHY) Specifications," 1999.
- [2.8] Online, http://en.wikipedia.org/wiki/Net_present_value#Formula
- [2.9] Online, <http://www.vivato.net>
- [2.10] Online, <http://www.muniwireless.com/article/articleview/13/1/24/>

CHAPTER 3

MULTIHOP NETWORK FOR BROADBAND FIXED WIRELESS ACCESS IN RURAL AREAS

Introduction

The studies in Chapter 2 indicate that in rural applications of Wi-Fi technology the system is typically coverage limited, and the key challenge is to improve the AP coverage area. Backhaul link and access point costs tend to dominate the overall system cost. Hence, cost-effective solutions are typically characterized by few APs, large coverage areas and the use of wireless backhaul. Technologies such as high gain directional antennas or packet steering beam antennas can be applied to improve the coverage range of an AP, and the cost benefits of these approaches have been studied in Chapter 2. This chapter studies an alternative approach of using multi-hop technology, where user nodes operate in the peer-to-peer ad hoc mode and rely on their neighbor nodes to forward traffic to an AP. A simplified multi-hop wireless network structure is shown in Figure 3.1. In the figure, it is not necessary for a user node to be in the coverage range of an AP to reach the outside network. This reduces the number of APs and associated backhaul links and increases the coverage area as a result.

A multi-hop wireless network is an application of ad hoc network technology in a fixed wireless network topology. By “fixed”, it means that the user nodes in the network are stationary, but the possibility of alternative routing is left open. There have been several studies of applications of multi-hop wireless networking for broadband access in urban, high-density areas [3.1]-[3.2]. Wireless LANs with a wireless multi-hop backbone

network were proposed in [3.3]-[3.4]. Merging IP and wireless networks using cell hopping was proposed in [3.5]. Test beds have been built to demonstrate the effectiveness of the approach, and real multi-hop networks have been deployed in some high-density community areas [3.6], [3.7]. However, the potential of using multi-hop in very sparse areas to extend the reach and lower the cost of more conventional point-to-multipoint access system architectures has been less explored.

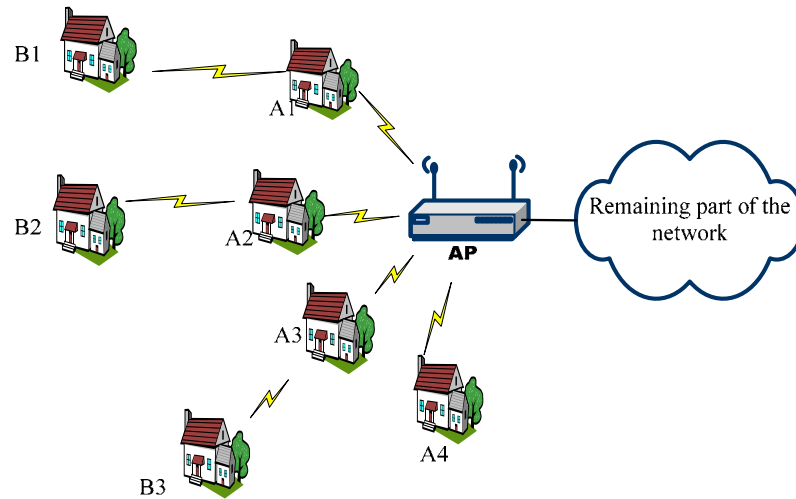


Figure 3.1. A simplified multi-hop wireless network.

A multi-hop wireless network is differentiated from a pure ad hoc network by its traffic pattern. In a typical multi-hop topology, the user traffic goes to or from a centralized AP, where the AP operates as the gateway to the outside fixed Internet. In a pure ad hoc network, the user traffic under control of a routing protocol spreads more geographically and uniformly among peer nodes, resulting in efficient spatial reuse and potentially higher capacity. Mesh networks are one type of ad hoc network, with nodes inside the network generally not mobile. In a mesh network, each node can find a path to any other node through one or multiple hops. Mesh networks are reliable and use

dynamic routing to route around broken or blocked links by finding alternative routes. In metropolitan areas, mesh networks are deployed to avoid the bottleneck problem by using alternative routing. However, for very sparse areas, mesh networking is not applicable as there are not enough nodes to form a mesh network.

For the multi-hop case, the traffic is always centralized, and spatial frequency reuse provides no obvious throughput advantage. The average user throughput is always limited by the constraints of the bottleneck link nearest to the AP or gateway, as pointed out in [3.8]. The authors explored the nominal capacity of a wireless mesh network based on the concept of the “bottleneck collision domain” and pointed out that the existence of an AP or gateway throttles the average throughput of each node, which decreases as $O(1/n)$, where n is the total number of nodes in the multi-hop WLAN. This would be a concern as the user density grows, however, studies in this chapter indicate that there are still large user density regimes typical of rural demographics where the bottleneck problem is not a limiting factor and the cost-benefits of multi-hop can be realized.

One issue associated with a multi-hop network is the extra delay added due to multi-hop relays. The further a node is away from a central gateway node, the more delay is incurred due to additive link transmission time and queuing in intermediate nodes. For real time applications, such as voice over IP, multi-hop networking is not a good choice. There are also privacy concerns of having intermediate user nodes serving as relays for other user’s traffic. In this study, it is assumed that intermediate nodes are willing to forward traffic for edge nodes. Also there are fairness concerns regarding how different nodes in a multi-hop network can be assigned the same amount of bandwidth. However, the algorithms used to maintain fairness among nodes at different positions in a multi-hop

network are beyond the scope of this study. It is assumed here that there are some algorithms that can be applied to address the fairness problem to ensure that nodes at the end of the chain will get the same amount of bandwidth as nodes closest to the gateway nodes and will not be starved for bandwidth.

In the following sections of this chapter, first a basic area coverage model based on multi-hop technology is proposed and its capacity limitations are analyzed by applying the “bottleneck collision domain” concept discussed in [3.8]. To mitigate the bottleneck problem and improve the average throughput of each node, an alternative model using sector antennas at the AP is considered. In the situation where the user node density is low and nodes are widely separated to form a multi-hop network, it is assumed that a high gain omni-directional antenna will be used to improve the transmission range to make a multi-hop network possible. An omni-directional antenna is necessary to enable the media access control (MAC) protocol to perform properly. A MAC protocol that exploits the characteristics of directional antennas has been recently proposed [3.9], but it is premature to apply it to a realistic scenario. Furthermore, 802.11 protocols and radio technology is used to enable a direct comparison with the results presented in Chapter 2. The possibilities of using multi-hop in rural areas are explored by using the same approach as in Chapter 2. And the domains where it can be cost-effectively utilized to provide broadband Internet access are established.

Multi-hop Model and Analysis

In this section, two types of area coverage models are proposed for wireless local area networks (WLANs). One is the basic multi-hop model, star mesh network (SMN)

model. The other is a SMN model with sector antennas (SMNS). The network capacity for these two types of models with different number of hops is analyzed. In the model development, it is assumed that there are always line of sight paths between neighboring nodes. Also no terrain effects are included and it is assumed that there is no blockage between neighboring nodes.

The Basic Multi-hop Model (SMN model) for WLAN

The basic model is also called a Star Mesh Network (SMN) model, as depicted in Figure 3.2, with only some of the nodes shown. The user nodes are fixed and assumed to be uniformly distributed according to node density ρ_h (nodes/square mile). Here a node is generally a house unit and might have several users. For simplification, the SMN model with a maximum of 3 hops from the end node to the center AP is analyzed. For models larger than 3 hops, a similar analysis can be applied.

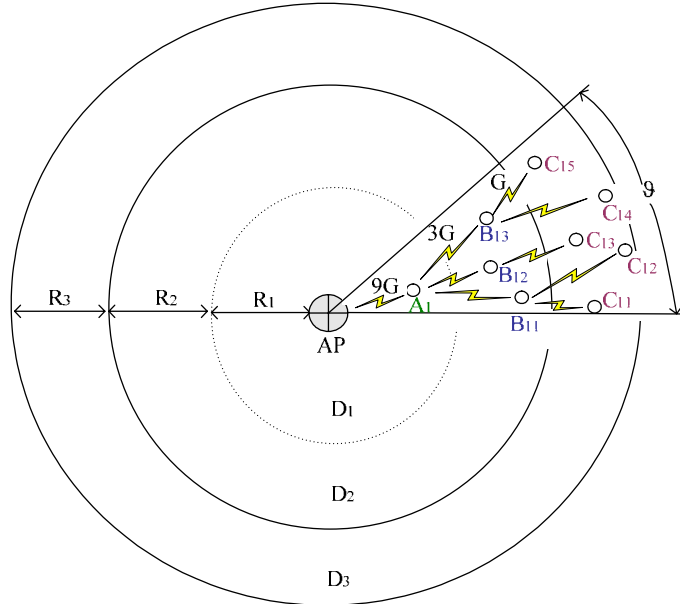


Figure 3.2. Star Mesh Network (SMN) model for the multi-hop WLAN.

To analyze the capacity characteristics of this model, the whole area is divided into three regions, indicated by D_1 , D_2 , and D_3 . The value of R_1 represents the communication range between AP to first hop user (e.g., AP to A_1), R_2 represents the distance from first hop user to the second hop user (e.g., A_1 to B_{11}), R_3 represents the distance from the second hop user to the third hop user (e.g., B_{11} to C_{11}).

In the models developed here, it is assumed that $R_2=R_3=R$, where R is the average distance between the neighboring user nodes determined by the user node density ρ_h . It will become clear later that the value of R_1 might be larger or less than R , depending on the value of the maximum load offered by each user node, the total link throughput provided and also the number of hops. Examples of the selection of the value for R_1 are given in the section discussing about “Multi-hop Model Selection”. The areas of these three regions are:

$$S_{D1} = \pi R_1^2 \quad (3.1)$$

$$S_{D2} = \pi(R^2 + 2RR_1) \quad (3.2)$$

$$S_{D3} = \pi(3R^2 + 2RR_1) \quad (3.3)$$

The numbers of nodes in each region are:

$$N_1 = \rho_h \cdot S_{D1}, N_2 = \rho_h \cdot S_{D2}, N_3 = \rho_h \cdot S_{D3} \quad (3.4)$$

On average each node in region D_1 has to forward traffic of $\frac{N_2}{N_1}$ nodes from region D_2 , while each node in region D_2 has to forward traffic of $\frac{N_3}{N_2}$ nodes from region D_3 .

The analysis of capacity limitation of the above model is based on the collision domain concept defined in [3.8]. A collision domain is formed by a set of links. At a particular time, if one link in the collision domain is active, all other links in the domain have to be inactive for the active link to have a successful transmission, as all links use the same frequency. A Bottleneck Collision Domain (BCD) [3.8] is a collision domain in one network that has to transfer the most traffic in the network. In this model structure, the BCD will exist among the links near the AP. For example, if R_1 is also equal to R , then in Figure 3.2 the BCD surrounding the link from the AP to A_1 generally includes all the links from the AP to A_i ($i = 1$ to N_1), all the links from A_i to B_{ij} ($i = 1$ to N_1 , $j = 1$ to 3), and also the links among B_{ij} ($j = 1$ to 3) and C_{1k} ($k = 1$ to 5).

The average load offered by each user node is noted as G . If a user has the following usage parameters: a data rate of K Mbps, N sessions per hour and P seconds per session, then on average the user load G can be calculated by using equation (3.5).

$$G = \frac{K(\text{Mbps}) \cdot N(\text{session/hr}) \cdot P(\text{s/session})}{3600\text{s/hr}} = \frac{KNP}{3600}(\text{Mbps}) \quad (3.5)$$

It is assumed that every user offers the same average load G . Then the traffic on each link, including the forwarded traffic and the originating traffic from each node, can be calculated. The total traffic load (L_T) on the links of the BCD is:

$$L_T = \left[\left(G + \frac{N_3}{N_2} G \right) \frac{N_2}{N_1} + G \right] \cdot N_1 + \left(G + \frac{N_3}{N_2} G \right) \cdot N_2 + \frac{N_3}{N_1} \cdot G \quad (3.6)$$

After simplification,

$$L_T = \left[2(N_2 + N_3) + N_1 + \frac{N_3}{N_1} \right] G \quad (3.7)$$

The nominal MAC layer capacity of each link is assumed to be B , which is the throughput that can be achieved at the MAC layer in a one hop traditional WLAN [3.8]. According to the theorem presented in [3.8], this BCD can't forward more than the nominal MAC layer capacity, which is B . Namely, $L_T \leq B$ and the maximum average offered load $G_{\max}$ is related to B as follows:

$$G_{\max} \leq \frac{B}{2(N_2 + N_3) + N_1 + \frac{N_3}{N_1}} \quad (3.8)$$

If we reduce the hop number from 3 to 2, the maximum average offered load $G_{\max}$ is related to B as follows:

$$G_{\max} \leq \frac{B}{2N_2 + N_1} \quad (3.9)$$

Hence, based on the required average user load G , the nominal MAC layer capacity B and the node density ρ_h , the value $R_{1\max}$ that satisfies the capacity requirement can be calculated. Depending on the actual values for above parameters, there might be or might not be a solution for $R_{1\max}$ that meets the requirement. If $R_{1\max}$ exists, then in a real deployment, R_l should be selected with a value not larger than $R_{1\max}$. Otherwise, the whole network will become capacity limited and the users will not be served with the required throughput.

For a given fixed node density ρ_h , the nominal MAC layer capacity B and the range of R_l , the maximum average load $G_{\max}$ that can be offered by each node before the whole network becomes capacity limited also can be calculated.

From the above equations for G ((3.8), (3.9)), it is clear that the maximum average

load per user node that can be served is much less than the nominal capacity for each link. As more links are included in the BCD, the capacity that can be offered by each node decreases. In this multi-hop model, the increase of the transmission range of the AP is not as efficient as that in the traditional one hop AP model, in which the increase of the transmission range is desirable in that it increases the coverage range for each AP in a coverage-limited network. In this multi-hop case, increasing the transmission range of the first hop adds more nodes to the direct coverage area of the AP, and due to the nature of the multi-hop network, where the nodes nearest to the AP need to forward the traffic from their corresponding outside nodes, the increase of nodes in the range of first hop (N_1) results in a substantial decrease in $G_{\max}$. This occurs because the first hop links are always within the bottleneck collision domain and increasing the transmission range of the AP reduces the spatial reuse ratio. The advantage is that with more nodes in the first hop range, the time delay per user decreases once they have access to the channel. Hence the proper selection of the transmission range of the AP, here the range R_1 , is very important, and is also related to the selection of the right kind of antenna for the AP.

The SMN Model with Sector Antennas (SMNS) for WLAN

If the network becomes capacity limited, there are two solutions that can be applied without changing the capacity for each link. One solution is to shrink the area of the model by reducing the hop number, e.g., reduce the hop number from 3 to 2. The relationship between $G_{\max}$ and B is already given in equation (3.8). The other solution is to put several APs at one site and use a sector antenna for each AP. In this SMNS model, the same assumptions as the previous basic model are made, except that there are now

three APs in the center of the coverage area, each with a sector antenna that covers 120 degree azimuthally, and each AP operates at a different frequency to reduce the inference among the neighbor nodes associated with different APs. It is also assumed that static routing is used for these fixed nodes to avoid the need to have nodes dynamically change frequencies to forward traffic from neighboring nodes in a different sector.

With the use of sector antennas, the original bottleneck collision domain can be separated into 3 independent smaller domains, and the total traffic load (L_T) on the links of the smaller BCD, after simplification, is:

$$L_T = \left[\frac{2}{3}(N_2 + N_3) + \frac{N_1}{3} + \frac{N_3}{N_1} \right] G \quad (3.10)$$

Accordingly, the maximum average offer load $G_{\max}$ is:

$$G_{\max} \leq \frac{B}{\frac{2}{3}(N_2 + N_3) + \frac{N_1}{3} + \frac{N_3}{N_1}} \quad (3.11)$$

For the 2-hop SMNS model, the maximum average offered load $G_{\max}$ is related to B as follows:

$$G_{\max} \leq \frac{B}{\frac{2N_2}{3} + \frac{N_1}{3}} \quad (3.12)$$

Hence with the use of sector antennas, $G_{\max}$ for each node increases approximately by a factor of 3 and the value $R_{1\max}$ that satisfies the capacity requirement can be calculated accordingly. It is noted that in general the increase will be a factor of M, where M is the number of sectors. The limiting factor will be the number of non-interfering frequencies available for use, which are 3 for Wi-Fi. Once an appropriate value of the

first hop range R_l that is less than $R_{l\max}$ is selected, the gain of the sector antenna needed to cover the specified range can be determined.

Multi-hop Model Selection

In the previous section, two types of multi-hop models for WLANs are proposed and the capacity limitations for each type are analyzed. Each type of multi-hop model can have different number of hops. In a realistic deployment, the average user traffic would be predefined, which means that the average load traffic that can be offered by each user is taken as a known design parameter. Not every multi-hop model can be applied, and the appropriate selection will be a function of user density. Based on the anticipated user data traffic, the appropriate model can be chosen to meet the user data traffic demand. As in Chapter 2, the rural area in Gallatin County, Montana is chosen as the study area. Based on demographic information and user data traffic requirements for this area, appropriate models are selected.

Demographic Information about the Study Area

The study area considered here is only the rural area in Gallatin County, Montana. The total population in the rural area is 22,491 with a land area of about 1580 square miles, so the population density is about $\rho = 15$ person/square mile, based on the demographic data obtained in the 2000 census [3.10]. It is assumed that on average there are approximately 3 persons per node, resulting in a node density $\rho_h = 5$ nodes/square mile. Considering a ten-year study period and applying an annual 1.9% percentage population increase, then at the end of ten years, the node density ρ_h will be about 6

nodes/square mile. It is also assumed that the nodes are distributed uniformly. Then the average distance between the neighboring user nodes R will be approximated by

$$\sqrt{2} \cdot \sqrt{\frac{1}{\rho_h}} = \sqrt{2} \cdot \sqrt{\frac{1}{6}} = 0.58 \text{ miles.}$$

To make a comparison, study results using the model for the case B in Chapter 2 are compared with results obtained with the SMN model or SMNS model in this chapter while all the other assumptions are kept unchanged.

Data Traffic Requirements and Technology Specification

As discussed in Chapter 2, the average traffic is $\frac{1}{12}$ Mbps per user and $\frac{1}{4}$ Mbps per node. Here 802.11g at 2.4GHz is considered. As compared with 802.11b at 2.4 GHz, 802.11g offers higher capacity with shorter radio range. Compared with higher frequency 802.11a systems at 5.8GHz, 802.11g systems provide greater range with the same capacity. It is assumed that 802.11g is used at user nodes with maximum transmit power of 50 mW and receive sensitivity of -86dBm at 18Mbps. For neighbor user nodes to be within range of each other, antennas that are omni-directional in the azimuthal plane with 8 dBi gain are assumed for each node. Additionally, an outdoor and line-of-sight propagation model is assumed with free space path loss. For the above performance parameters of 802.11g radio, reference is made to the Cisco Aironet 1200 for 802.11 a/b/g Access Point and the Cisco Aironet 802.11 a/b/g Wireless PCI Adapter. Using the approach described in [3.5], the maximum 802.11g MAC layer throughput B is 12 Mbps for wireless links with a data rate of 18 Mbps. Though 802.11g can have a maximum throughput of 54 Mbps, it will have only about 1/3 the coverage range as compared with operation at a data rate at 18Mbps with same antenna applied. Hence 802.11g with

moderate throughput of 18 Mbps is used to obtain an appropriate coverage range for a rural area with sparse node distribution.

Solutions Based on the SMN Model

With these data traffic and technology parameters, to guarantee that the network will not become capacity-limited until the end of ten-year study, the average hop distance R is approximated to be 0.58 miles, as shown in the previous section.

For the 3-hop SMN model, equation (3.13) is obtained by substituting $B=12$ Mbps, $G_{\max} = \frac{1}{4}$ Mbps, $\rho_h = 6$ nodes/square mile, $R = 0.58$ miles and N_1 , N_2 and N_3 from equations (3.1) to (3.4) into equation (3.8), and let $\frac{R_1}{R} = k$.

$$6.28 \cdot (8 + 8k + k^2) + \frac{3}{k^2} + \frac{2}{k} \leq 48 \quad (3.13)$$

A meaningful solution for k must meet the condition that $k > 0$, but with the assumed parameters, equation (3.13) has no such solution for k that meets this condition, indicating that the 3 hop SMN model is always capacity-limited under the assumed conditions. Hence the 3-hop SMN model is not an appropriate choice for the given network environment and will not be considered further in this study.

As known from the analysis of the multi-hop model, if the network is capacity-limited, one solution is to decrease the hop number. For the 2-hop SMN model, after substituting all the assumed parameters, equation (3.14) is obtained as follows:

$$6.28(2 + 4k + k^2) \leq 48 \quad (3.14)$$

Solving this equation for k , the maximum allowed value $k_{\max}$ is 1.1, which means

$R_{1\max} = 1.1R \approx 1.1 * 0.58 = 0.64$ miles. In other words, with the 2-hop SMN model, given the specific requirements for capacity, the maximum distance between the AP and the first node range can't exceed 0.64 miles. Otherwise more user nodes will be included in the first hop range and the average data traffic possible for each user will be lower than the required offered load.

Solutions Based on SMNS Model

Another solution is to use the SMNS model instead of SMN model to solve the capacity-limited problem. For the 3-hop SMNS model, after substituting all the assumed parameters, equation (3.15) is obtained:

$$\frac{6.28}{3}(8+8k+k^2)+\frac{3}{k^2}+\frac{2}{k} \leq 48 \quad (3.15)$$

Solving the equation for k , the maximum value $k_{\max}$ is 1.4, which means that $R_{1\max} = 1.4R \approx 1.4 * 0.58 = 0.81$ miles.

Similarly for the 2-hop SMNS model, it is found that $k_{\max} = 3$ and $R_{1\max} = 1.74$ miles.

The 1-hop SMNS model, which is noted as SMNS1, does not utilize multi-hop, but instead is the conventional model discussed in Chapter 2 with sector antennas applied to each AP to improve the coverage range and the capacity. Assuming characteristics of typically available 802.11g commercial products, with the assumption that there is clear line of sight and sufficient Fresnel zone clearance, and without consideration of other possible path losses, ideally, every increase of 6 dB in the link budget will double the transmission range, and the resulting ideal maximum distance possible between the AP

and user nodes is about 2.218 miles for SMNS1 with the use of a 20dBi sector antenna at the AP and 8dBi antennas at the user nodes. For 2-hop and 3-hop SMNS models, the antenna gains needed for each AP are calculated separately for each model based on the value of R_1 selected, which cannot be larger than $R_{1\max}$ calculated above, taking into account the fact that a user node is equipped with an 8dBi omni-directional antenna. The 8dBi omni-directional antenna at user nodes is required in multi-hop cases for neighboring nodes to be within range of each other. However, in the single hop case, higher gain subscriber node antennas could be used as the node has to connect only to the AP.

Simulations and Results

The above analysis indicates that the 2-hop SMN model and the 1-hop, 2-hop and 3-hop SMNS models offer feasible solutions for the specific study area. In this section, following the same techniques discussed in Chapter 2, these models are used to evaluate the economic performance of broadband fixed wireless access systems based on multi-hop technology in the rural areas of Gallatin County, Montana.

802.11b vs. 802.11g

To compare results for this study with those of Case B in Chapter 2, first, the simulation for Case B in Chapter 2 is repeated by replacing 802.11b technology with 802.11g technology, while all the other parameters are kept unchanged. The cost assumptions for network equipment are basically the same as those that are summarized in Table 2.2. The additional cost assumption is for an 802.11g access point, which is

assumed to be \$1300. Furthermore, only the rural area is considered. Table 3.1 gives the financial results comparing the performance of Case B with 802.11b and 802.11g technologies applied. The results are expressed in terms of the NPV, the time to break even, and the capital cost per user. The cost/potential-user is the total capital cost in the ten-year study period divided by the number of potential users (namely the total population) at the end of ten-year study period. The number of real users is the number of potential users times the market share and market penetration.

Table 3.1. Simulation Result with 802.11b and g Technology Applied Using Conventional Single-Hop Point-to-Multipoint Topology

Items \ Models	Case B with 802.11b technology	Case B with 802.11g technology
NPV (\$k)	26176	-128,971
Years to break even	3	N/A
Capital cost/potential user (\$)	545.6	5019.7
Capital cost/real user (\$)	909.4	8366.2

The results in Table 3.1 confirm previous observations that 802.11b is a more cost-effective technology relative to 802.11g in low user density areas, since 802.11g as compared with b has increased capacity with shorter transmission range. As was pointed out in Chapter 2, the network for the rural area is always coverage-limited during the whole study period and the use of 802.11g in a point-to-multipoint single hop configuration doesn't improve the network economic performance. On the contrary, its performance is much inferior to that of 802.11b technology due to the costs of the additional APs and backhaul links needed to provide complete coverage. These results underscore the need to use multi-hop topology with 802.11g technologies for rural areas.

SMN2 vs. SMNS2

In these two multi-hop cases, appropriate antennas are selected to meet the coverage range requirement for the first hop. For SMN2, an omni-directional antenna with 12dBi gain is used at each AP. For SMNS2, sector antennas with 20dBi gain are used at each AP, and an omni-directional antenna with 8dBi gain is used at each user node in both cases. The costs for different antennas are taken into account in the study. The additional cost assumptions for the different antennas are given in Table 3.2, and the other costs assumptions are kept the same as indicated in Table 2.2. Table 3.3 gives the simulation results for these two models.

Table 3.2. Cost Assumptions for Antennas Applied

Component	Cost (\$)
8dBi gain omni-directional antenna	60
12dBi gain omni-directional antenna	250
20dBi gain sector antenna	575

Table 3.3. Simulation Results Comparison between SMN2 and SMNS2 Models

Items \ Models	SMN2	SMNS2
NPV (\$k)	27,642	35,987
Years to break even	4	2
Capital cost/potential user (\$)	529.9	290.8
Capital cost/real user (\$)	883.2	484.7

Comparing the third column of Table 3.1 (Case B in Chapter 2 with 802.11g technology) with Table 3.3, it is obvious that huge cost performance improvement is obtained by using multi-hop topology. Without the use of multi-hop, the network with 802.11g will not break even and it has a negative NPV value at the end of the ten-year study period. With multi-hop and 802.11g, the results for model SMN2 show that the

breakeven point occurs in 4 years and the NPV at the end of the ten-year study period is positive, far outperforming the 802.11g single hop baseline case. The results in Table 3.3 also show that SMNS2 has even better economic performance than SMN2. In the SMNS2 model, each coverage area is divided into three sectors and the use of higher gain sector antennas extends the first-hop range. In this coverage-limited network, further improvement in coverage has the additional advantage in that the increased capacity of 802.11g can be fully exploited. Furthermore, the APs on the same site share the same backhaul capacity, saving costs by reducing the total number of backhaul links.

SMNS1, SMNS2 and SMNS3 vs. Population Density ρ

The performance of multi-hop SMNS model is affected by several factors including the maximum transmission range of the first hop and the population density of the study areas. Figures 3.2 and 3.3 give the economic performance of SMNS models as a function of the population density ρ when the first-hop transmission range of SMNS1 is fixed at 2.2 miles and 1.1miles respectively. The population density ρ is in persons per sq mile, which is related to ρ_h by the equation $\rho = 3 \cdot \rho_h$, assuming 3 persons per house unit or node.

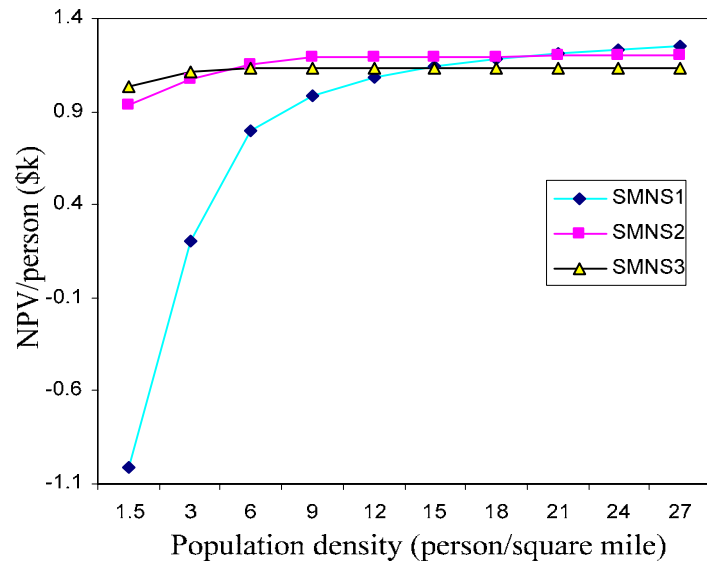


Figure 3.2. Performance of SMNS models vs. population density ($R_1 = 2.2$ miles).

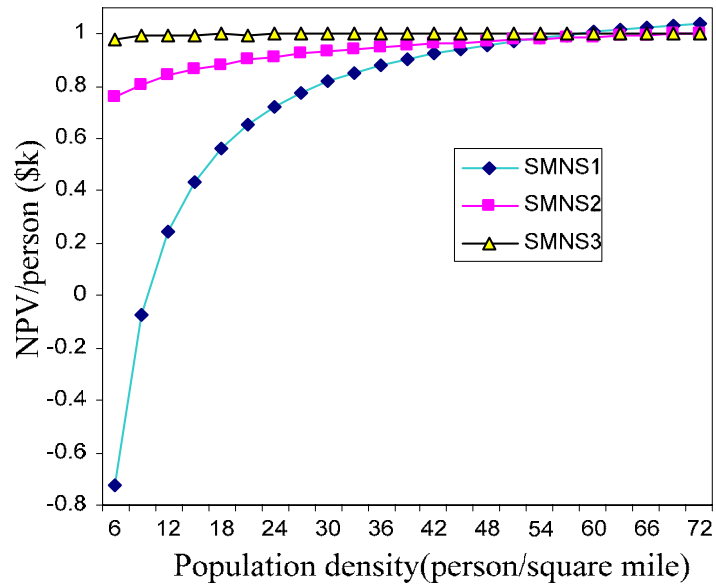


Figure 3.3. Performance of SMNS models vs. population density ρ ($R_1 = 1.1$ miles).

Several important points can be observed by comparing the results in Figure 3.2 with Figure 3.3. First, the results in both Figure 3.2 and Figure 3.3 indicate that when the population density ρ is low, SMNS3 has the best performance among the three and

SMNS1 has the worst performance. However, the domain in which the multi-hop SMNS models have better performance than the single hop models is related to the first-hop transmission range of SMNS1. In general, the shorter the first-hop transmission range, the bigger the population density domain where multi-hop models outperform single hop models.

Secondly, when the population density ρ increases to a critical value (about 21 persons/square mile in Figure 3.2 and 57 persons/square mile in Figure 3.3), thereafter the performance of the single and multi-hop models are comparable.

The third point is that as the maximum one hop transmission range decreases, as would be the case if terrain and other propagation factors were taken into account, the advantage of using multi-hop in low population density areas increases. For a realistic transmission range, the results in Figure 3.3 show that for a population density of 6 persons/square mile, the NPV/person is about -0.7 \$k, 0.75 \$k and 1 \$k for SMNS1, SMNS2 and SMNS3 respectively. This further demonstrates the necessity and importance of utilizing multi-hop topologies in rural areas with low population density. The advantage of multi-hop becomes greater as the capacity of the radio link increases and the transmission range decreases.

The fourth point is that it is cost effective to serve extremely low user density areas with multi-hop. When the population density is below about 2 persons/sq mile and about 10 persons/sq mile in Figure 3.2 and Figure 3.3 respectively, multi-hop cases have a positive NPV, but the single hop case has a negative NPV.

Sensitivity Analysis of the Critical Population Density

As observed from Figures 3.2 and Figure 3.3, when the population density increases to a certain value, the advantage of multi-hop SMNS model disappears, and the performance of multi-hop SMNS model becomes inferior to that of SMNS model with fewer hops.

The terms ρ_{c12} , ρ_{c13} and ρ_{c23} are defined as the critical population densities where the SMNS1 outperforms SMNS2, SMNS1 outperforms SMNS3 and SMNS2 outperforms SMNS3 respectively. For the case in Figure 3.2, the values for the three critical population densities are 21, 15 and 5 persons/square-mile respectively.

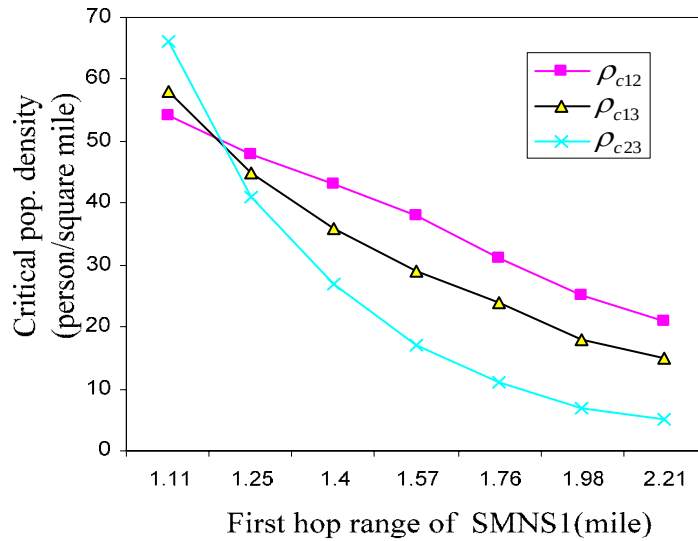


Figure 3.4. Critical population density vs. first hop transmission range of SMNS1.

Figure 3.4 shows that as the first hop transmission range of SMNS1 decreases, the critical population densities increase, indicating that the population density domain where the multi hop models have an advantage over the one hop model becomes larger. For example, if the transmission range is 1.1 miles, the critical population density $\rho_{c13} = 57$

persons/square mile, whereas for a range of 2.2 miles, $\rho_{c13} = 15$ persons/square mile. This behavior is due to the capacity limit of the hop nearest to the access point. As the range decreases, the density of users can increase, while the total number of users served is constant.

The results in Figure 3.4 indicate that the critical population density is related with the first hop transmission range of the SMNS1 model. Another factor, the average user traffic load, also has an effect on the critical population density. To investigate the effect of the average user traffic load on the critical population density, the first-hop transmission range of SMNS1 is fixed at 1.4 miles, while the average user traffic load is changed to test the sensitivity of ρ_{c12} to the value of average user load traffic.

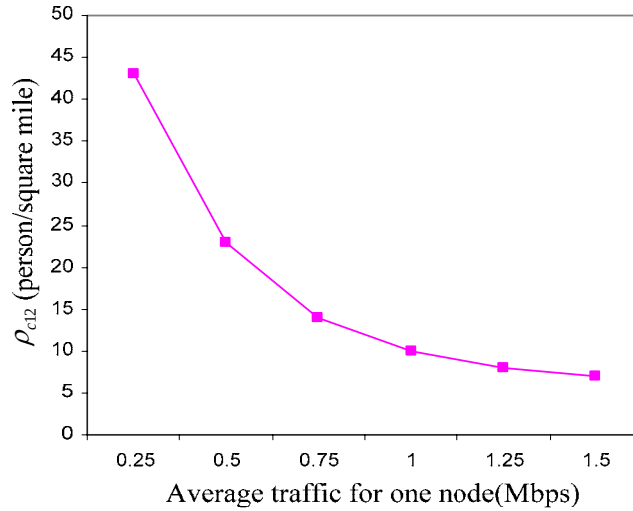


Figure 3.5. ρ_{c12} vs. average traffic load at a fixed first-hop radio range of SMNS1.

Figure 3.5 shows that as the average traffic load increases, the critical population density ρ_{c12} decreases. This further indicates that for the multi-hop SMNS models, as the average user traffic load increases, the maximum first hop range of the SMNS2 must

decrease to keep the network from becoming capacity-limited, the total coverage range of the SMNS2 model decreases.

Summary

Multi-hop models for WLANs have been proposed and their capacity limitations have been analyzed. For a given node density and average node traffic, the approach to selecting an appropriate model with a certain number of hops is described. Several multi-hop models based on 802.11g, are applied to low-density rural areas and their economic performances are evaluated and compared with that of similar scenarios in Chapter 2.

The simulation results show that the network economic performance with 802.11g technologies applied is inferior to the performance using 802.11b if the network is configured as a traditional single-hop, point-to-multipoint topology. However, it is cost effective to combine the use of sector antennas with multi-hop topology as demonstrated in the SMNS models for use in rural areas.

The sensitivity of critical population densities analysis also shows that there are domains of user node density and offered traffic load where multi-hop strategies provide a more cost effective approach to broadband wireless Internet access than conventional point-to-multipoint architectures. While the studies are based on ideal propagation models, the results indicate that decreasing the transmission range to account for terrain effects increases the advantages of multi-hop systems relative to single hop networks.

References

- [3.1] M. Frodigh, P. Johansson, P. Larsson, "Wireless ad hoc Networking – The art of networking without a network", *Ericsson Review*, no. 4, 2000, pp. 248-263.
- [3.2] W. Webb, "Mesh Technology Boosts Wireless Performance", *EDN*, pp. 50-63, October, 2003.
- [3.3] K. Mase, N.Karasawa, M. Kusumi, et al, "Wireless LAN with Wireless Multihop Backbone Network," *Proceeding of the International Conference on Wireless LANs and Home Networks*, Singapore, 5-7 December 2001.
- [3.4] Hung-Yu Wei, "Two-Hop-Relay Architecture for Next Generation WWAN/WLAN Integration," *IEEE Wireless Communication*, April 2004, pp.24-30.
- [3.5] Jahan Hassan, Sanjay Jha, "Cell hopping: a lightweight architecture for wireless communications," *IEEE Personal Communications*, Volume 10, Issue 5, Oct 2003 Page(s): 16 – 21.
- [3.6] "Tropos Networks Technology", *Tropos Networks whitepaper*, May, 2003.
- [3.7] Sanjit Biswas and Robert Morris, "Opportunistic Routing in Multi-Hop Wireless Networks", *Proceedings of the Second Workshop on Hot Topics in Networking (HotNets-II)*, Cambridge, Massachusetts, November 2003.
- [3.8] Jangeun Jun and M. L. Sichitiu, "The Nominal Capacity of Wireless Mesh Networks", *IEEE Wireless Communications*, October 2003, pp.8-14.
- [3.9] Romit Roy Choudhury, Xue Yang, Ram Ramanathan and Nitin H. Vaidya, "Using Directional Antennas for Medium Access Control in Ad Hoc Networks", *MOBICOM'02*, September 23-28,2002, Atlanta, Georgia, U.S.A, pp.59-69.
- [3.10] U.S. Census Bureau, Census 2000 Summary File 3(SF3)-Sample Data, compiled by Census & Economical Information Center, Montana Department of Commerce, 09/02.

CHAPTER 4

TOWARDS MOBILE WIRELESS COMMUNICAITON IN RURAL AREAS

Introduction

Chapters 2 and 3 have focused on fixed wireless communication in rural areas. Chapter 2 investigated and evaluated the feasibility and financial viability of providing fixed wireless access in rural and remote areas based on a baseline network structure. Nontraditional and innovative approaches were applied to improve the overall cost-benefits of the wireless network. In Chapter 3, one of the innovative technologies, multi-hop networking, was further investigated with the objective of finding domains where a multi-hop network structure has an advantage over single-hop network structure.

The study of multi-hop network technology in a fixed user scenario also brings forward the application research of Mobile Ad Hoc Network (MANET) technology in rural and remote areas. Multi-hop for fixed wireless applications shares some similarity with mobile ad hoc network technology, as both use multi-hop techniques to extend radio transmission range, which is critical for both fixed and mobile communications in rural and spares areas. However, there exist significant differences between a fixed multi-hop network and a mobile ad hoc network. A mobile ad hoc network in a rural area may be partially connected due to low node density and high node mobility. The provision of high-speed access to mobile users in rural areas is much more challenging.

In this chapter, first the similarities and differences between a fixed multi-hop network and a mobile ad hoc network will be addressed. Then the characteristics of a rural area mobile ad hoc network are analyzed and the challenges of applying MANET

technology for mobile wireless communications in rural and remote areas are pointed out.

Fixed Multi-hop Wireless Networks and Mobile Ad hoc Networks

An ad hoc network is a self-configuring network of wireless stations/nodes connected by wireless links. The term “ad hoc” is often used as slang to refer to an independent basic service set (IBSS) as defined in IEEE 802.11 standards [4.1]. In an ad hoc network, nodes communicate on a peer-to-peer basis without using an AP or any connection to a wired network. The wireless node can be mobile or static and the network is typically formed in a spontaneous manner with a dynamic topology. If the wireless nodes in the network are mobile, it is called a mobile ad hoc network.

As mentioned before, one similarity between a fixed multi-hop wireless network and a mobile ad hoc network is that they both use multi-hop techniques to extend the communication range. In each type of network, every node is essentially a router in that it not only receives/transmits information, but also relays/forwards traffic for other nodes. In both cases, information originating from a source node might go through several intermediate nodes before it finally reaches the destination. Nevertheless, there still exist considerable differences between a fixed wireless multi-hop network and a mobile ad hoc network.

A sample of a simplified fixed multi-hop wireless network is shown in Figure 4.1, in which the user nodes are represented by laptops. In a fixed multi-hop wireless network, the user nodes can be static or mobile. It is fixed in the sense that there always exists a central point, or access point (AP), which is generally fixed and that serves as a point of connection to the wired network infrastructure, as indicated in Figure 4.1. An access point

generally works as a gateway, providing access to the wired network, via the wireless medium for associated wireless nodes. When a user node needs to send data to another user node, which might be in the same local wireless network or outside, the data traffic will be sent to an AP first, either directly or through a multi-hop transmission. Then the AP will be responsible for sending the data to the desired destination, also through one hop or multi-hop transmission. In such a network, a user node first needs to register with an AP to make an association. Then the AP is responsible for its associated members for successful data transmission. The route to a user node receiving or sending data traffic to other nodes generally can be preconfigured through the AP.

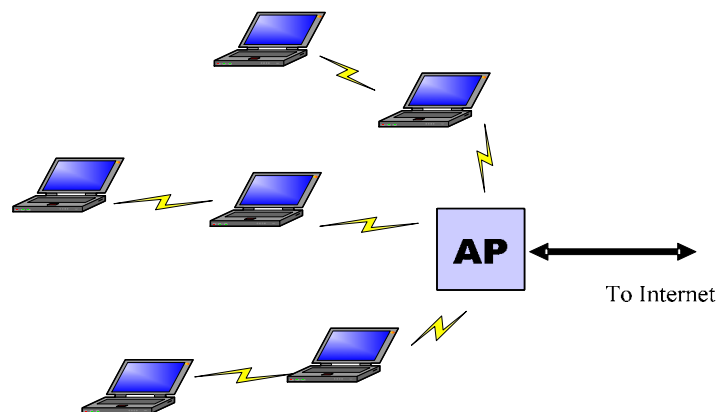


Figure 4.1. A simplified fixed multi-hop wireless network.

An example of a mobile ad hoc network is shown in Figure 4.2. The mobile nodes are also represented by laptops. A mobile ad hoc network is a self-configuring network of wireless stations connected by wireless links. Since the nodes may move randomly, the network topology is dynamic and arbitrary. In such a network, there is no central point, such as an AP in a fixed multi-hop wireless network case. User nodes communicate on a peer-to-peer basis without using an AP. From time to time, one or more nodes might

connect to a wired network and serve as gateway and provide a temporary link to a fixed network for other nodes. However, it is not necessary for such a gateway node to be fixed in the sense of location as well as function. Data transmission generally needs multi-hop routing without central coordination. When a source node wants to send data to a destination node, the data will be sent directly to the destination node if the destination node is in direct radio transmission range of the source node. If the destination node is out of direct radio range of the source node, then the source node first needs to find a route to the destination by using a specific routing protocol. After the route is found, the data will be sent out via the route by forwarding through the intermediate nodes. If any intermediate node on the route moves out of the range of its neighbors during this data transmission, then the data might be lost. Therefore, the design of an appropriate routing protocol that accounts for node mobility is crucial to the successful performance of a mobile ad hoc network.

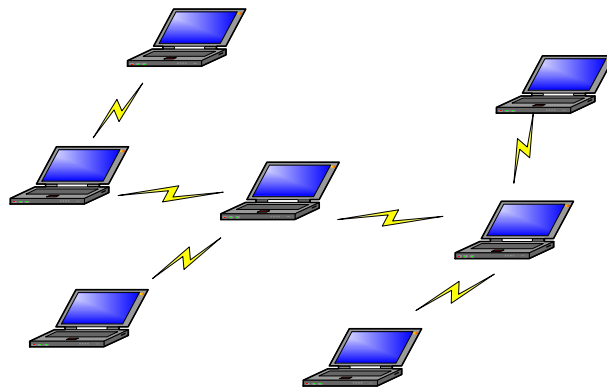


Figure 4.2. A sample of a simplified mobile ad hoc network.

Routing Challenges for Mobile Ad hoc Networks in Rural Areas

In a mobile ad hoc network, due to the mobility of user nodes, the connectivity and the topology of the network are always changing. The routing protocol applied to a mobile ad hoc network is critical to the realization of fast and effective information dissemination. In the following, the existing mobile ad hoc network routing protocols are briefly reviewed and the necessity to design a new routing protocol for mobile ad hoc networks for use in rural areas is addressed.

Overview of the Existing MANET Routing Protocols

Mobile ad hoc network (MANET) technology has shown great promise in urban and high user density areas. Much research work, mainly on efficient routing protocol design, has been done on mobile ad hoc networks. The existing routing schemes can be categorized into two big groups as topology-based and position-based [4.2]. Topology-based routing algorithms generally use available link states to make routing decisions for packet forwarding. Position-based routing algorithms use node position information to make routing decisions. Position-based routing algorithms are considered to have better scalability than topology-based routing algorithms since localized routing algorithms are used to achieve global delivery [4.3]. However, position-based routing algorithms generally require some kind of location information, which can be obtained through the use of the Global Position System (GPS) or other types of positioning service.

The topology based routing schemes can be further subdivided into proactive, reactive and hybrid approaches. Proactive routing schemes attempt to maintain consistent,

up-to-date routing information about the available paths even if these paths are not currently being used. Each node is required to maintain one or several tables to store routing information. The routing information is continuously and proactively updated when there are topology changes by propagating updates throughout the network. Sometimes proactive routing schemes are also called table-driven routing schemes. Proactive routing schemes are used in routing protocols such as Destination-Sequenced Distance-Vector Routing (DSDV) [4.4], Wireless Routing Protocol (WRP) [4.5], Optimized Link State Routing Protocol (OLSR) [4.5] and Topology Broadcast Based on Reverse Path Forwarding (TBRPF) [4.6].

Reactive routing schemes try to find or establish routes only when desired or on-demand, namely, when a node has data traffic to be sent. Generally, when a node has data to send, it will initiate a route discovery process in the network. After a route is discovered, data will be sent along the route. During data transmission, the route will be maintained by a routing maintenance procedure until the route is no longer desired or the destination node moves out of the existing route. Reactive routing schemes include Dynamic Source Routing (DSR) [4.7], Ad hoc On-demand Distance Vector routing (AODV) [4.8] and Temporally-Ordered Routing Algorithm (TORA) [4.9].

Hybrid routing schemes combine local proactive routing with global reactive routing to improve efficiency and scalability. An application of a hybrid routing scheme was proposed in Zone Routing Protocol (ZRP) [4.10]. An extensive review and comparison of different topology-based ad hoc routing protocols is given in [4.11] and [4.12].

Position-based routing schemes generally need a location service to determine the current position of a specific node. Different position-based routing schemes have been

proposed depending on the type of location service used, and also how the position information is exploited for packet forwarding. Location-Aided Routing (LAR) [4.13] uses node position information to enhance the route discovery process of reactive ad hoc routing algorithms. In Greedy Perimeter Stateless Routing (GPSR) [4.14], a sending node includes the approximate position of the destination node when sending out a data packet. An intermediate node forwards the packet to a neighbor lying in the general direction of the destination node. In Distance Routing Effect Algorithm for Mobility (DREAM) [4.15], each node maintains a database that stores the position information of all other nodes in the network. Each node updates its position information periodically through flooding packets in the network. One advantage of the position-based routing scheme is its scalability [4.3]. However, the price for this scalability is the additional location service required.

DSDV and WRP are two typical table-driven or proactive routing protocols. These routing protocols are proactive in that the routing tables maintaining all active links are continuously and proactively updated in all nodes. Having routing tables always being current and updated may be desirable in that there would be minimal delivery delay in sending the message. However, this also necessitates very high overhead traffic associated with updates of link status. The update traffic overhead increases when the network topology is rapidly changing. Proactive routing algorithms all have the same high control overhead problem and don't apply very well to network with highly dynamic topology changes or high node mobility.

The typical on-demand or reactive routing protocols are DSR and AODV. DSR is based on source routing. When a source node has data to send, it initiates a route

discovery process to find a complete path from source node to destination node. After a route is found, the source node sends out a data packet by specifying the complete path to the destination in the packet header and the intermediate nodes along this path simply forward the packet to the next hop indicated in the path. When there is no data traffic, there is no control overhead. However, since each data packet carries a source route, there is a large overhead incurred during packet delivery. AODV is based on DSDV but it improves on DSDV by creating routes on an on-demand basis. It also initiates route discovery only when a node has data to send. In contrast to the source route in DSR, the established route information is stored in the routing table of individual nodes on the path. When forwarding a packet, there is no control overhead associated with the delivery process. The route acquisition procedure is called a “pure on-demand route acquisition” as nodes that are not on a selected path do not need to maintain routing information and participate in routing table exchanges [4.8]. While there is no static overhead associated with these routing protocols, there is a slow start before transmitting data since a route discovery process needs to be accomplished beforehand.

Characteristics of MANETs in Rural Areas

In rural areas, a promising application of MANET technology is for vehicle-to-vehicle and vehicle-to-roadside communication. A network is formed by vehicles moving along a highway system, where each vehicle functions as a mobile node. Another name for this kind of mobile ad hoc network is Vehicular Ad hoc Network (VANET). Such a mobile ad hoc network in rural areas has its unique characteristics. First, in rural areas, generally there is very little or no existing communication

infrastructure. The VANET serves as a surrogate for fixed infrastructure. Secondly, in contrast with urban areas with high vehicle densities, the density of mobile vehicles is much lower in rural areas, and these mobile nodes might be sparsely distributed, resulting in high possibility of network partition. Thirdly, as a characteristic of VANET, mobile nodes in such a network can have high node mobility. Hence the network topology may be highly dynamic

Needs for New MANET Routing Protocols

The routing algorithms mentioned above have a common assumption: that the mobile networks are always connected and that a path can always be found from a source node to a destination node when an information message is being transmitted. The typical successful applications of these protocols are in cases where mobile nodes are closely spaced and interact with each other frequently. The characteristics of a mobile ad hoc network in rural domains pose new requirements on the routing protocol design, which are quite different from those conventional routing schemes directed to high node density domains. Further, nodes in rural areas may have high node mobility. High mobility further increases the unpredictability of the network topology. Direct application of those routing schemes reviewed in previous section may result in a very low routing efficiency or even a total failure.

Conventional topology-based routing schemes can't apply to mobile ad hoc networks in rural areas. As pointed out in [4.16], if a mobile ad hoc network is partially connected or fully disconnected, then the existing reactive routing schemes such as DSR and AODV will fail to discover a complete path. Failure to find a complete or connected path from

source node to destination node during the route discovery process will result in no data traffic delivery, and total failure of the routing process. Proactive routing protocols, such as DSDV and WRP, will fail to converge during routing information exchange stage due to network partitioning, resulting in a deluge of topology update messages and the failure of the routing process.

Position-based routing schemes generally require additional node physical position information during the routing decision process. A location service is needed as well to provide the position information of nodes. Generally, location service is provided based on position information derived using GPS or other positioning systems. Then this location service is leveraged to help the routing procedure, such as in GPRS and DREAM. In areas where GPS information is not available, such as mountainous terrain, or where GPS cannot be added to nodes, these position-based routing protocols cannot be used.

In summary, due to the unique characteristics of rural areas, existing routing protocols can't be directly applied. Instead, a new routing algorithm needs to be developed that can function under conditions of node sparsity and variable network topology.

Mobility Models for Mobile Ad hoc Networks in Rural Areas

As discussed in the previous section, an appropriate MANET routing protocol suitable for mobile ad hoc networks in rural areas is critical for successful application of mobile wireless communication. Once a routing protocol that takes into rural area network characteristics into account is designed, a method of accurately evaluating the performance of this specifically designed routing protocol is also needed. Current routing

protocols are generally evaluated based on random mobility models, which assume that mobile nodes in a network move randomly in an area. This is appropriate to simulate most of the common application scenarios. However, for mobile nodes moving on highway road systems or having other spatial constraints, routing protocol performance evaluation based on these random mobility models is not adequate. A mobility model that can closely reflect the real movement of mobile nodes is needed to accurately evaluate the performance of a routing protocol designed for such mobile ad hoc network and further improve the routing protocol design.

A mobility model defines the pattern of movement of a mobile node. It is a key component in characterizing node behavior and in evaluating routing protocol performance. The pattern of node movement influences the performance of routing protocols in ad hoc networks [4.17]. The two most commonly used mobility models for ad hoc network simulations are the Random Waypoint Mobility Model [4.7] and the Random Walk Mobility Model [4.18]. In the Random Waypoint Mobility Model, all nodes are initially randomly distributed within a simulation area. Then each node randomly chooses a destination (waypoint) and moves straight towards that destination with a randomly chosen speed. Once a node reaches the destination location, after pausing at its current location for certain time (pause time), it chooses another random destination and moves toward it with another randomly chosen speed. Each node continues this behavior, alternately pausing and moving to a new location during the whole simulation period. In the Random Walk Mobility Model, a node moves from one location to another location by randomly choosing a direction and a speed. Each movement occurs in either a constant time interval or a constant distance. At the end of

one movement, a new direction and speed are calculated and selected. During a movement, if a node reaches the boundary of a simulation area, it bounces off the boundary with an angle determined by the incoming direction and then moves on this new path. Besides these two most popular random mobility models, there are other mobility models either derived from these two mobility models or with some modifications added to improve the movement pattern. They are all based on the random movement assumption. These random mobility models are simple and easy to simulate, but sometimes not very realistic.

In a vehicular ad hoc network, the node movement pattern is not totally random. Their mobility is generally restricted to roadways, and constrained by speed limits and other factors. The random movement assumption in existing mobility models is not valid in this scenario. Hence, for a vehicular ad hoc network, a mobility model that can reflect the movement pattern of mobile nodes traveling on a highway system is needed. The design of an application oriented mobility model is very important for the accurate performance evaluation of the routing protocol designed for highway applications in rural areas. Accurate performance evaluation will help point out the right direction for modification and improvement during routing protocol design.

Along highway road systems in rural and remote areas, there is generally no fixed communication infrastructure available, except for occasional road sensors or other special purpose. Drivers on a road in a rural area may feel isolated from outside information world. Currently, the most common way to provide information to mobile users is through cellular service. However, in rural and remote areas, there exist huge spans with poor signal coverage or no coverage at all. Figure 4.3 shows a recent “drive

test” done by the Western Transportation Institute. The cellular service coverage along a major Montana highway connecting Bozeman to Yellowstone National Park was measured [4.18]. The figure indicates that those areas close to towns or cities are covered by cellular service. However, for those areas or areas far away from cities or towns, wide gaps exist with no signal coverage at all. This “drive test” result not only indicates the necessity of providing an alternative communication service to nodes in rural areas through the application of mobile ad hoc network technology but also explains why a new mobility model is needed to model the movement of a vehicular ad hoc network in rural areas.

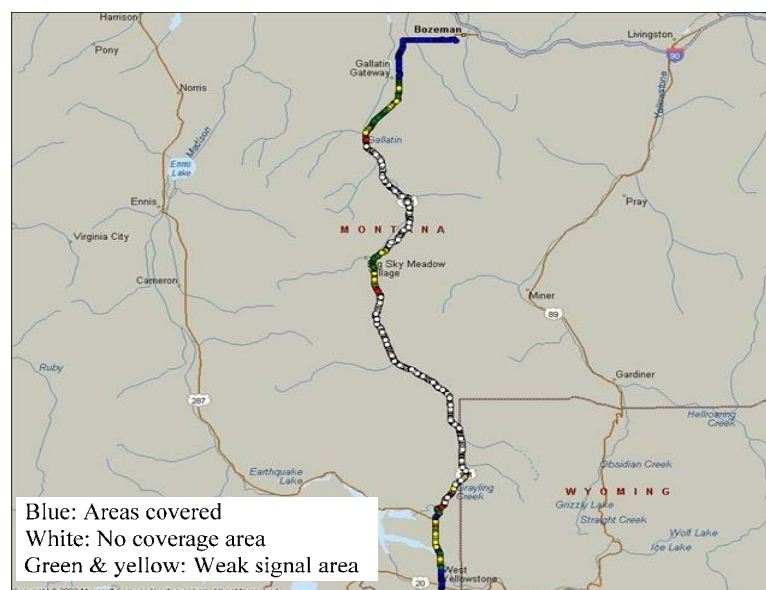


Figure 4.3. Cellular signal coverage test along a highway in Montana. Courtesy: Doug Galarus, Western Transportation Institute.

Summary

The similarities and differences between a fixed wireless multi-hop and a mobile ad hoc network have been discussed. In a mobile ad hoc network, an appropriate routing protocol is essential for successful data communication. Conventional MANET routing protocols are generally designed for fully connected networks. The unique characteristics of a mobile ad hoc network in rural and sparse areas require a new routing protocol. The importance of a mobility model to the accurate routing protocol performance evaluation has been discussed as well. In Chapter 6, a Border node Based Routing (BBR) routing protocol is proposed specifically for mobile ad hoc networks in rural areas. Before discussing the routing protocol design, in Chapter 5, a mobility model that reflects the node movement pattern in rural area, where the nodes are constrained to roadways, is developed first.

References

- [4.1] IEEE Std. 802.11, "Wireless LAN Media Access Control (MAC) and Physical Layer (PHY) Specifications," 1999.
- [4.2] M. Mauve, A. Widmer and H. Hartenstein, "A survey on position-based routing in mobile ad hoc networks," *Network, IEEE*, Volume 15, Issue 6, Nov.-Dec. 2001 Page(s): 30 – 39.
- [4.3] Ivan Stojmentovic, "Position-Based Routing in Ad Hoc Networks," *IEEE Communication Magazine*, Volume 40, Issue 7, July 2002, Page(s):128-134.
- [4.4] C. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers," *Computer Communication Review*, 24, October 1994.
- [4.5] P. Jacquet et al., "Optimized Link State Routing Protocol. *Internet draft, draft-ietf-manet-olsr-04.txt*, work in progress, Sept. 2001.
- [4.6] B. Bellur, R. Ogier and F. Templin, "Topology Broadcast Based on Reverse Path Forwarding (tbrpf)," *Internet draft, draft-ietf-manet-tbrpf-01.txt*, work in progress, Mar. 2001.
- [4.7] D. Johnson and D. Maltz, "Dynamic source routing in ad-hoc wireless networks," *ACM SIGCOMM*, August 1996.
- [4.8] C. Perkins and E. Royer, "Ad hoc on-demand distance vector routing," *2nd IEEE Workshop on Mobile Computing Systems and Applications*, February 1999.
- [4.9] V.D. Park, M.S. Corson, "A highly adaptive distributed routing algorithm for mobile wireless networks," *INFOCOM '97, Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies*. Proceedings IEEE Volume 3, 7-11 April 1997 Page(s): 1405 - 1413 vol.3
- [4.10] Z. Haas and M. Pearlman, "The performance of query control schemes for the zone routing protocol," *ACM SIGCOMM*, 1998.
- [4.11] E. Royer and C.-K. Toh, "A Review of Current Routing Protocols for Ad Hoc Wireless Networks," *IEEE Personal Communication*, Apr. 1999, pp. 46–55.
- [4.12] J. Broch et al., "A Performance Comparison of Multi-hop Wireless Ad Hoc Network Routing Protocols," *Proc. 4th ACM/IEEE Int'l. Conf. Mobile Computing and Networking MOBICOM '98*, Dallas, TX, USA, 1998, pp. 85–97.

- [4.13] Y. Ko and N. Vaidya, "Location-aided routing (LAR) in mobile ad hoc networks," *ACM MOBICOM*, November 1998.
- [4.14] B. Karp and H. Kung, "Greedy Perimeter Stateless Routing for Wireless Networks," *ACM MOBICOM 2000*, August 2000.
- [4.15] S. Basagni et al., "A Distance Routing Effect Algorithm for Mobility (Dream)," *Proc. 4th Annual ACM/IEEE Int. Conf. Mobile Computing and Networking, MOBICOM '98*, Dallas, TX, USA, 1998, pp. 76–84.
- [4.16] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Single-copy routing in intermittently connected mobile networks," *Sensor and Ad Hoc Communications and Networks, IEEE SECON 2004, First Annual IEEE Communications Society Conference*, on 4-7 Oct. 2004 Page(s): 235 – 244.
- [4.17] G.. Ravikiran, S. Singh, "Influence of mobility models on the performance of routing protocols in ad-hoc wireless networks," *Vehicular Technology Conference, VTC 2004-Spring, IEEE 59th*, Volume 4, 17-19 May 2004 Page(s): 2185 - 2189 Vol.4.
- [4.18] T Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *WCMC: Special issue on Mobile Ad Hoc Networking: Research, Trends and Applications*, 2(5): 483–502, 2002.

CHAPTER 5

GEOGRAPHIC AND TRAFFIC INFORMATION BASED MOBILITY MODEL
FOR MOBILE AD HOC NETWORKS IN SPARSE AND RURAL AREASIntroduction

The improvements in MANET technology and the ever-increasing safety requirements have turned vehicle safety communications into an important research topic. Vehicle safety communications is designed to provide drivers with real-time safety information through vehicle to vehicle or vehicle to infrastructure communications. Vehicle safety communication methods rely upon the creation of autonomous, self-organizing wireless communication networks, namely ad hoc networks to connect vehicles with fixed infrastructure and with each other. In rural and sparse areas, there is very little or no fixed infrastructure available. Also there is a lack of conventional cellular mobile communications infrastructure in rural and sparse areas as well. The “drive test” of cellular signal coverage along a highway in Montana, shown in Figure 4.3, illustrates that cellular service is very limited in rural and sparse areas. Figure 5.1 gives the Verizon wireless network coverage map for northwest part of U.S. Areas shown in yellow have cellular service. And the white color indicates where there is no coverage. From the figure one can see that there are large parts of the northwest where there is no wireless cellular service available. The coverage provided by wireless carriers is predominantly in urban areas and along major highways, not rural areas and minor roadways. Hence, vehicle ad hoc networks (VANETs) will play an important role in providing public safety

communications in rural areas.

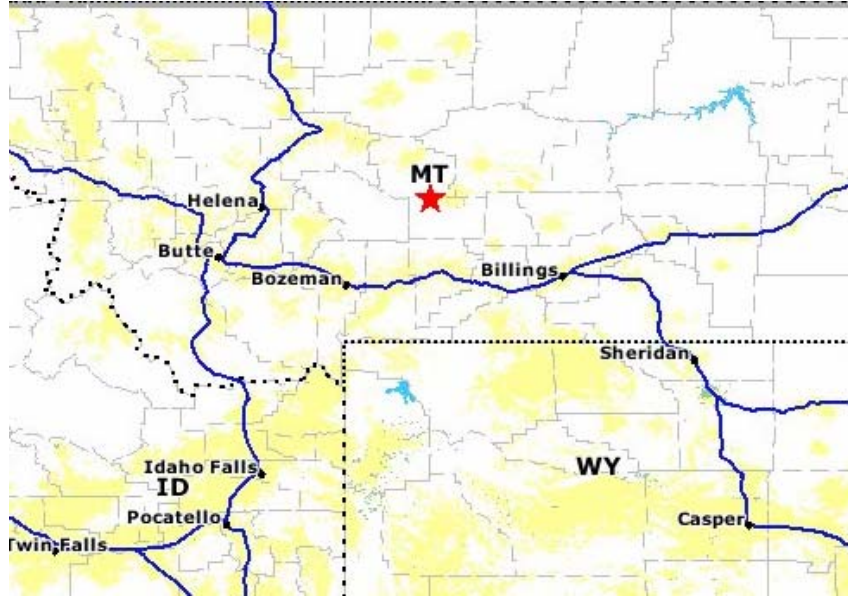


Figure 5.1. Verizon cellular coverage map in the northwest area of U.S. [5.1].

Ad hoc networks have been used as part of the communication infrastructure in vehicle communication systems such as in CarNet [5.2] and FleetNet [5.3]. For those vehicle systems, designed for high user density areas, both inter-vehicle communication and vehicle-infrastructure communication may occur. The routing protocols designed for these mobile ad hoc networks were proposed mainly for fully connected conditions. The performance of those routing protocols was typically evaluated with commonly used random walk or waypoint mobility models. However, ad hoc networks in rural areas are characterized as low node density and highly partitioned or partially connected, which means that generally there does not exist a complete path from a source node to a destination node at any specific time. The network might consist multiple node clusters, where the nodes inside each cluster may connect to each other through one or

multiple wireless hops. But the clusters are disconnected from each other. Routing algorithms appropriate for low node density regions and partitions have been less explored. There has been very limited investigation of routing protocol performance where node mobility may be constrained by roadways or other factors.

One difficulty in evaluating routing protocol performance on mobile ad hoc networks with low node density and partitions is the lack of suitable mobility models. In vehicular ad hoc networks, mobile nodes may move on predefined paths and the movement is not totally random and not accurately described by conventional mobility models. To explore ad hoc routing protocol performance in sparse areas, a Geographic and Traffic Information based Mobility Model (GTI mobility model) is proposed to model the movement under real world constraints. Based on this mobility model, performance of routing protocols can be better evaluated. The GTI mobility model is applicable under a range of node densities, but the effects of node mobility modeling are most pronounced in sparse networks, where disconnection and partitioning can adversely affect conventional routing protocols.

In the remainder of this chapter, first the related research work on mobility models is reviewed and the importance of mobility models on routing protocol performance evaluation is discussed. Then the GTI mobility model is proposed and described in detail. An application example is given by applying the GTI mobility model to a sample area in a rural region. The GTI mobility model is then used to evaluate the performance of an ideal routing protocol using simulation

Related Work on Mobility Models

A mobility model defines or reflects the movement of mobile nodes. Significant effects that a mobility model can have on the performance of an ad hoc routing protocol have been well studied in [5.4], where the performances of several popular routing protocols have been evaluated with several group mobility models including the Column Model, the Pursue Model and the Reference Point Group Mobility Model-Random Waypoint (RPGM-RW) Model.

A survey of current mobility models used for ad hoc routing protocol simulations is given in [5.5], where the mobility models are categorized into two major categories: trace-based models and synthetic models. Traces are the mobility patterns of mobile nodes observed in a real life system. Generally, traces can provide accurate movement information, especially when a large number of mobile nodes are involved and an appropriately long observation period is carried out. Synthetic models attempt to realistically model the movements of the mobile nodes without use of traces. Synthetic models are well developed and have been widely used for routing protocol performance evaluation. Compared with synthetic mobility models, relatively few trace-based mobility models have been developed.

Existing Synthetic Mobility Models

Commonly used synthetic mobility models can be further categorized into independent mobility models and group mobility models. The first group includes mobility models such as, the Random Walk Mobility Model which is well described in [5.5], the Random Waypoint Mobility Model [5.6], and Random Direction Mobility

Model [5.7]. The other group includes models such as the RPGM-RW [5.4], the Reference Point Group Mobility Model [5.8], and several other variations. Synthetic models are convenient to use, but sometimes a synthetic mobility model is not sufficient to characterize the movement of mobile nodes in real world scenarios.

Incorporation of obstacles in mobility modeling is examined in [5.9]. Since most of the commonly used mobility models are limited to random movements without consideration of obstacles, the mobility model proposed in [5.8] offers improvement over those random walk mobility models in that it takes into account the existence of obstacles and the restrictions that the obstacles impose on node movement, path selection and wireless transmission. The model can be used to model campus-like areas. The obstacle mobility model is also a synthetic mobility model, but it takes a big step towards realistic mobility modeling.

Compared with the commonly used random walk mobility models, the City Section Mobility Model reported in [5.5] and the Freeway and Manhattan Models proposed in [5.10] also add some constraints to the movement of the mobile nodes. These mobility models treat mobile nodes as vehicles traveling on a city road or a highway system, where the mobile nodes can only move on predefined paths and must also follow some traffic laws. These mobility models belong to the synthetic mobility model category as well. During the simulation, once the start point and the destination point are selected, nodes will move according to a path-finding algorithm, such as the shortest time traveled or the shortest distance traversed. Once the destination point is reached, the process will repeat.

Traces for Mobile Ad Hoc Networks in Rural Areas

In most cases, the commonly available synthetic mobility models are sufficient to test and evaluate the basic performances of routing protocols. However, a mobility model that closely reflects the real movements of mobile nodes will help to correctly evaluate a routing protocol and to optimize its performance. To study vehicular ad hoc networks, especially in sparse and rural areas, a GTI mobility model is proposed. With this model, it is easier to examine the connectivity characteristics of mobile ad hoc networks in domains where they are highly partitioned, partially connected or fully connected, by controlling some of the configurable parameters, such as, the average distance between vehicles and the average node density. Also, the GTI mobility model has potential for future extension with consideration of terrain effects, which will affect line of sight communications and enable inclusion of propagation factors. The GTI mobility model proposed is a trace mobility model that closely models the movements of mobile nodes in a real-world environment.

Highway planners use specialized tools, such as CORSIM [5.11], that characterize vehicular movement designed specifically for integration with transportation system modeling and analysis. While these tools contain accurate microscopic vehicle mobility models, they do not support studies of inter-vehicle communications and cannot be used in combination with a discrete time simulator to study ad hoc routing protocol performance.

Geographic and Traffic Information Based (GTI) Mobility Model

The GTI mobility model generates trajectories for each mobile node according to

geographic and vehicle traffic information by using a sampling method. Each trajectory is a trace file that records the path along which a mobile node will move for some period of time. For a specific simulation area, the basic geographic information about the area can be obtained from a digital map file, which includes position information (represented in altitude, latitude and longitude) for the start point and end point of each road, cross points, turn points and the physical length of each road. The necessary vehicle traffic information includes traffic load, velocity limits of mobile nodes on each road in that simulation area and other traffic rules that might affect node movements. The actual traffic information can be obtained through local transportation department offices if available, approximated using long term averages or treated as configurable parameters. Complete area information helps to develop accurate mobility models that closely match the actual application scenario.

Fundamentals of GTI Mobility Model

The proposed GTI mobility model is based on the topology information extracted from a digital road map and traffic information obtained from a local transportation department, both reflecting a specific simulation area. The topology information can be expressed as a graph that is composed of a set of vertices (V) and a collection of edges (E). Vertices are abstracted as the significant locations or positions from the road map, such as the start points, the end points of road sections as well as the cross points and the turn points along road sections. An edge is a link between two neighboring vertices and may have some additional information associated with it, such as velocity limits and average node densities. Edges are abstracted as the interconnections between locations.

An edge exists only when there is a real road connection between two neighboring positions in a road map. A path from one location to another is composed of one or more connected edges. Mobile nodes moving along a path can be considered as moving from one vertex to another along edges in the graph abstracted from a digital road map while taking into account the actual traffic information.

The Generation of Traces for Mobile Nodes

Based on the abstracted topology graph, trajectories for each mobile node moving in the simulation area for a certain period of time are generated under some simplifications and assumptions. First, roadways are all segmented and linearized based on the significant positions, such as turn points and intersection points, on the road map. If a higher granularity modeling is desired, then more positions along roads can be taken as additional anchor points. At the same time, the following assumptions are made:

1) Nodes move on each edge have the same average velocity v_{av} and velocity limit v_{limit} . A node moving on that edge will have a velocity v , where

$$v = (1 + k_v) \cdot v_{av} \quad (5.1)$$

Where k_v is a random number uniformly distributed between $(-0.1 \sim 0.1)$, and,

$$v \leq v_{limit} \quad (5.2)$$

2) Nodes are initially evenly distributed along each edge with some random deviation of initial positions, introduced by a random number k_p between $(-0.1 \sim 0.1)$. As indicated in Figure 5.1, if the coordinates of the two vertices V_1 and V_2 on an edge are (x_1, y_1) and (x_2, y_2) , and there are n nodes initially distributed on that edge, then the

initial position of node $N_i (0 \leq i \leq n)$ is expressed as (x_i, y_i) , where

$$x_i = \frac{x_2 - x_1}{n} \cdot (i + k_p) \quad (5.3)$$

$$y_i = \frac{y_2 - y_1}{n} \cdot (i + k_p) \quad (5.4)$$

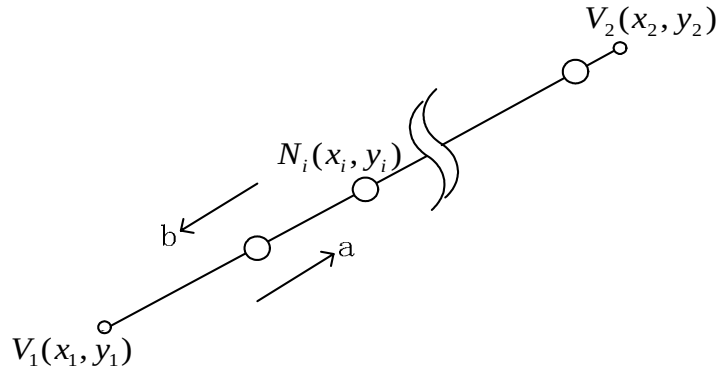


Figure 5.2. Initial node distribution on an edge.

3) At the beginning of the simulation, each node on a road section chooses its initial direction of motion randomly with equal probability, as indicated in Figure 5.2, with a probability given as

$$\begin{aligned} &P \{ \text{node chooses direction a} \} \\ &= P \{ \text{node chooses direction b} \} \\ &= 0.5 \end{aligned} \quad (5.5)$$

Once a node chooses an initial direction, it then continues in the same direction until a turn point or a cross point is reached.

4) When a node reaches a cross point and it has k directions to choose from to go forward, then the probability that the node choose any of the directions is the same, namely,

$$P \{ \text{a node chooses any of the } k \text{ directions} \} = 1/k \quad (5.6)$$

An example cross point with $k = 3$ is shown in Figure 5.3.

5) The total number of nodes in the simulation area during the simulation time is constant; In other words, nodes entering the simulation area are balanced by nodes leaving it.

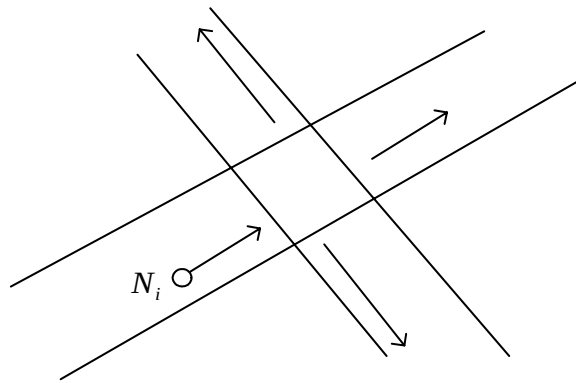


Figure 5.3. Direction choosing rule at a cross point with $k = 3$.

In the above assumptions, the parameters k_v and k_p are used to introduce some randomness to the velocities of different vehicles and their initial position distribution. If more randomness is desired, bigger value ranges can be chosen for k_v and k_p .

With all above assumptions, the steps that GTI mobility model uses to generate trajectories for mobile nodes are implemented using a C++ program. The flow diagram in Figure 5.4 summarizes the procedure in detail.

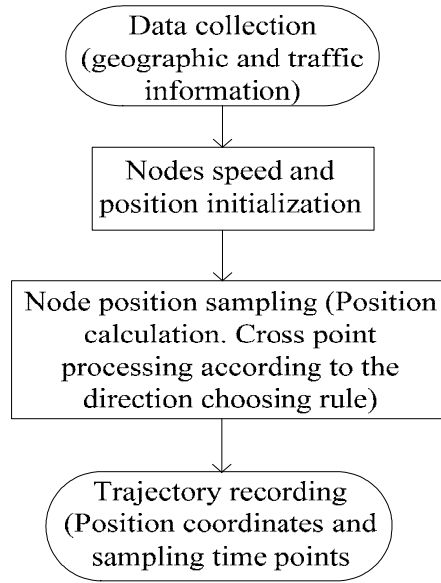


Figure 5.4. GTI mobility model trajectory generation flow diagram.

In the trajectory recording stage, position coordinates and accordingly the sampling time points are recorded to *.trj files. The header of each generated trajectory file has following format:

< # X position Y position Z position Traverse time>

Where # represents the index of the sampling point, X, Y and Z represent the latitude, longitude and altitude of the specific position respectively. Currently, the altitude of a specific position is not included in this mobility model. However, it can be included in the GTI mobility model in the future for modeling terrain effects on routing protocol design. The traverse time is the time interval between two neighboring sampling points.

Selection of the Sampling Time Interval in GTI Mobility Model

Trajectories generated by the GTI mobility model approximately represent the traces of mobile nodes moving in a specific area for a certain period of time. However, several

factors affect how closely these trajectories match the real movements of the mobile nodes in the real world. One factor is how much detail is retained when the graph is abstracted from a real digital road map. The more details that are included, the more closely will the trajectories match the real world scenario. Another important factor is the sampling time interval.

The shorter the sampling time interval is, the higher is the granularity of a trajectory. However, decreasing the sampling time interval will increase the size of the trajectory file. If there are thousands of nodes in the network, and total observation time for each node continues for several hours, then the total data increase will be significant. On the other hand, if the sampling time interval is too long, the resulting trajectory might not be accurate enough and some details of vehicle motion effects might be lost. The criteria used to select the sampling time interval is related to the node velocity and the radio transmission range of mobile nodes. Assuming that all nodes have the same radio transmission range, if two nodes are coming from opposite directions, the time that the two nodes are in the radio transmission range of each other can be expressed by:

$$T_R = \frac{2R}{v_1 + v_2} \quad (5.7)$$

In equation 5.7 R is the radio transmission range, v_1 and v_2 are the velocities of the two nodes coming from opposite directions respectively. The sampling time interval T_s should less than T_R , to guarantee that the two nodes will not pass without detecting the existence of each other. Assuming the maximum node velocity to be $v_{\max}$, then T_s should be:

$$T_s < \frac{R}{v_{\max}} \quad (5.8)$$

Generally, to achieve a higher granularity, a sampling time interval

$$T_s \ll \frac{R}{v_{\max}} \text{ should be selected.}$$

Application Example of the GTI Mobility Model

To demonstrate the use of the GTI mobility model, an application example is given by applying the GTI mobility model to the roadway grid of Yellowstone National Park (YNP), which is a large, sparse rural area with rough terrain and very little fixed communication infrastructure. The total Yellowstone page map is given in Figure 5.5 [5.12], which shows the basic geographic information of the park.

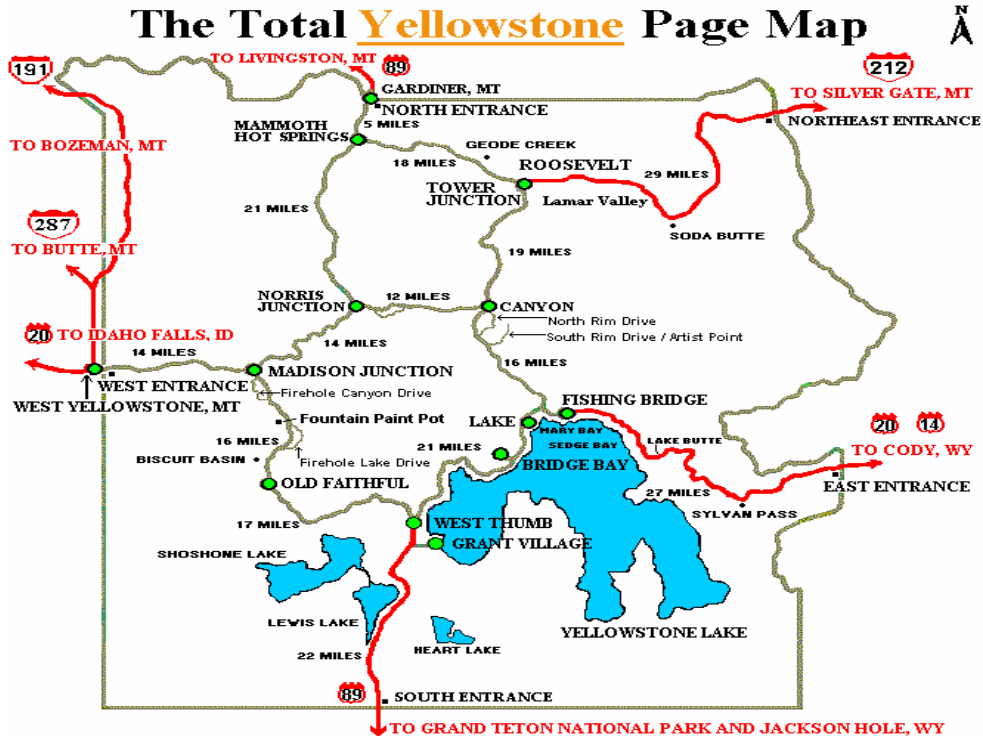


Figure 5.5. Geographic information of YNP [5.12].

For the purposes of demonstration, a mobile ad hoc network with only 100 mobile nodes is constructed based on the YNP roadway topology, though the actual traffic load might be much higher. Assuming the velocity limit to be $v_{\max} = 15m/s$, and a radio range of 1000 meters, characteristic of 802.11b wireless technology, a sampling interval of 6 seconds is chosen. This value is much less than $\frac{R}{v_{\max}} = 66.7$ seconds. Each trajectory simulated the movement of one node with 1-hour duration. In order to validate and visualize the trajectories generated, OPNETTM is chosen as the network simulation tool. The digital map of YNP and the generated trajectories are both input into OPNETTM. Figure 5.6 gives the initial node distribution as well as the trajectories in green lines for each mobile node. The trajectories are overlapping with each other at the beginning. Figure 5.6 shows that the road topology of YNP consists of two main loops and 5 branches that lead to each of the park entrances respectively. The roadways, shown in Figure 5.5, are linearized according to turn points and cross points, as shown in Figure 5.6. While some of the details of the area are lost in this approximation, the main characteristics of the highway system inside YNP remain.

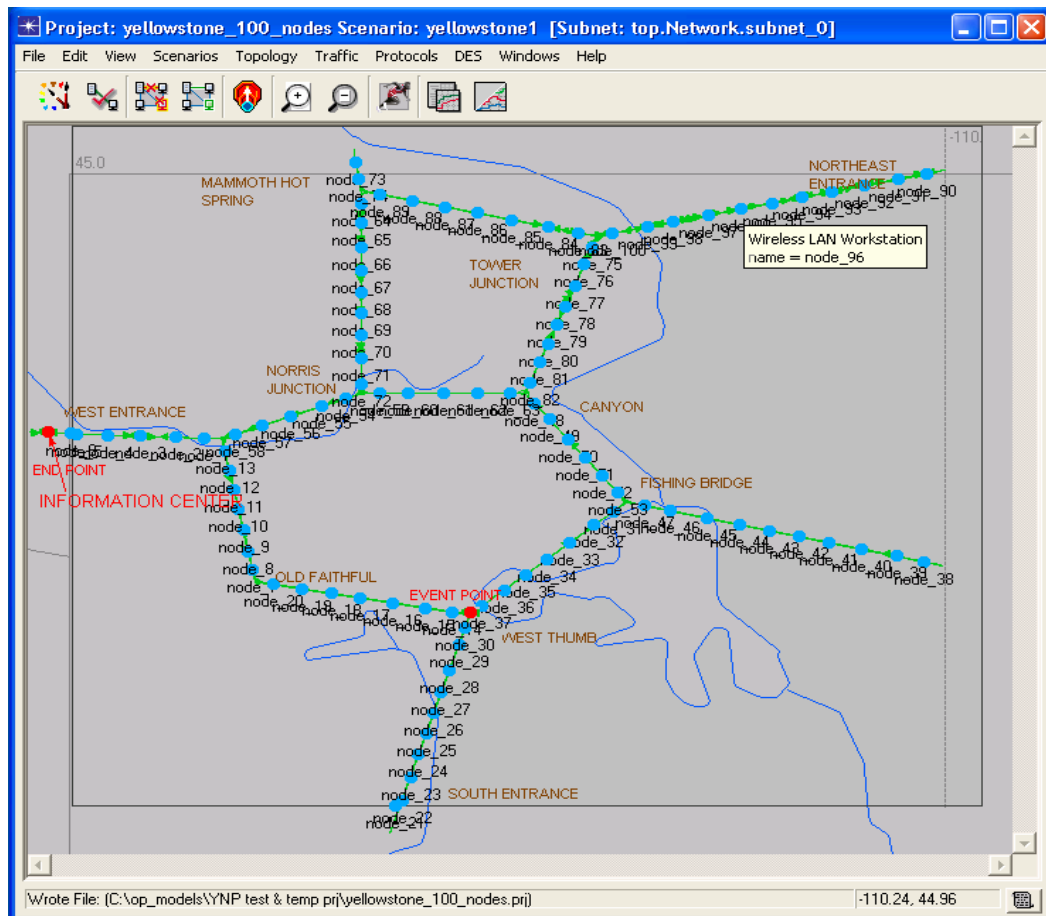


Figure 5.6. Mobile nodes and their trajectories in YNP.

Application of GTI for Evaluation of a Routing Protocol

Mobile ad hoc networks in sparse and rural areas can be characterized as partially connected with low node density and high mobility. With the motivation to design a routing protocol that is appropriate under these conditions, a simulation study is carried out to evaluate the performance of an ideal routing protocol. The basic rules for this ideal routing protocol are briefly described in the following paragraph. The ideal routing protocol is similar to an epidemic routing protocol, which was originally proposed in

[5.13] for partially connected ad hoc networks. Popular ad hoc routing protocols, such as DSR or AODV, are not chosen for the simulation because these routing protocols are mainly designed for a fully connected network. The focus here is partially connected or highly disconnected ad hoc networks.

We first use an ideal routing protocol for the network simulation rather than one of the currently available ad hoc protocols. First, the ideal routing protocol is simple and maintains the main rules of a practical routing protocol. The use of the ideal routing protocol helps to better investigate the connectivity characteristics of the underlying mobile ad hoc network, and thus enables validation of the GTI mobility model, without considering the many routing details of a practical routing protocol. Moreover, the simulation results based on the ideal routing protocol can also provide some insights into the design of a practical routing protocol that might be more effective for a partially connected ad hoc network.

The Ideal Routing Protocol

The ideal routing protocol is similar to a simplified version of the epidemic routing protocol proposed in [5.13]. For purposes of simplification, the ideal routing protocol uses ideal message exchange rules and also makes the following assumptions for information delivery.

Ideal message exchange rules:

- 1) Message hand offs occur when moving nodes are within radio range of each other.
- 2) Information exchange is instantaneous when two nodes are within radio range of

each other.

Assumptions:

- 1) The message processing time at each individual node is zero.
- 2) Nodes keep the messages when they move on.
- 3) The number of nodes in the simulation period is constant.
- 4) Nodes move in accordance with the trajectories files.
- 5) The simulation ends once the message reaches the destination.

Simulation Environment and Parameters

The simulation scenario is designed as follows: A source node or *Event node*, which represents a node that has an accident or has some local traffic related information, is located at the West Thumb cross point of YNP (see Figure 5.5). This node generates data traffic and sends this data to the destination node or *End node*, representing the Information Center, located at the west entrance of YNP. Both the *Event node* and the *End node* are shown by red dots in Figure 5.6. The ideal routing protocol is used for the information delivery. The explored questions are: Can the event information be transmitted from the *Event node* to the *End node* through the mobile ad hoc network? If the message can be successfully delivered, what is the transit time (T_{trans}) that it takes to transmit a message from the *Event node* to the *End node*?

The radio transmission range of each node is noted as R . The approximate total physical road length after linearization inside YNP is 194.36 miles. Assuming that there are a total of N nodes inside the park, then the average node separation (L_{av}) is

approximated by $\frac{194.36 * 1.609 * 1000}{N}$ meters. A parameter α is defined as the ratio of

the radio transmission range R and L_{av} , namely $\alpha = \frac{R}{L_{av}}$.

Based on the basic geographic information and traffic data obtained from the park administration office, simulations were carried out using a range of initial node velocities and velocity dispersions, and position randomness. Table 5.1 summarizes some general simulation parameters.

Table 5.1. Simulation Parameters

Parameters	Value
Simulation Area	YNP highway system
Total simulation time	2 hours
Approximate total physical road length after linearization	194.36 miles
Total number of nodes	1400
Average distance between neighboring vehicles (L_{av})	223.4 meters
α	0.45~2.0

Simulation Results

Since some randomness exists in the GTI mobility model due to the addition of randomness to the initial node distribution, node velocity and direction chosen, trajectories generated with each use of the model are different even with the same initial configuration parameters. With the parameters indicated in Table 5.1, for each test case, trajectories for all mobiles nodes are generated for 15 trials and the transit times are derived by simulation.

The first set of studies explores network connectivity characteristics as a function of the ratio α . Table 5.2 and Table 5.3 summarize the dependence of transit time with α when the average node movement velocity is between 12.1~14.7 (m/s) and 24.2~29.4 (m/s) respectively.

Table 5.2. Transit Time vs Ratio α (Node Velocity: 12.1~14.7 m/s)

α	Transit time (T_{trans})		
	Avg (s)	Max (s)	Standard deviation (s)
0.45	5463.0	7077.3	1103.7
0.90	4968.2	6692.0	847.3
1.0	4376.1	5813.6	532.0
1.35	0	0	0
>1.35	0	0	0

Table 5.3. Transit Time vs Ratio α (Node Velocity: 24.2~29.4 m/s)

α	Transit time (T_{trans})		
	Avg (s)	Max (s)	Standard deviation (s)
0.45	2620.6	3355.9	500.4
0.9	2551.4	3235.6	467.5
1.0	2516.8	3233.9	470.1
1.35	0	0	0
>1.35	0	0	0

The simulation results indicate that the connectivity characteristic of the network is strongly dependent on α , while the movement velocity has very little effect on network connectivity. Both simulations show that during average traffic load hours, if α is less

than 1.0, then the network is partially connected. When α is 1.0 or less, the radio transmission range is not more than the average distance between neighboring vehicles, and the message delivery is mainly dependent upon the movement of the mobile nodes themselves, instead of due to wireless forwarding by the intermediate nodes hop by hop. The average transit time is about 5000 seconds when average node velocity is 13.4 m/s and goes to about 2500s when the average node velocity doubles. The transit time obtained in both cases is close to the time for a vehicle to move from the position of the *Event node* to the position of the *End node* at the corresponding average vehicle velocity. Note that when the radio range is greater than the average vehicle separation, the transit time goes to zero, characteristic of an ideal routing protocol, as the message processing time, transmission time and propagation time have been set to zero.

To explore whether the ad hoc network connectivity characteristics discovered by using the idea routing protocol simulation is correct, we also investigated the probability distribution of the average neighbor node degree (number of neighbors within radio range) when α changes from 0.45 to 2.25. The probability density functions of the average neighbor node degree for all nodes during the 2-hour simulation time when node movement velocity is averaged at 13 m/s are shown in Figure 5.7.

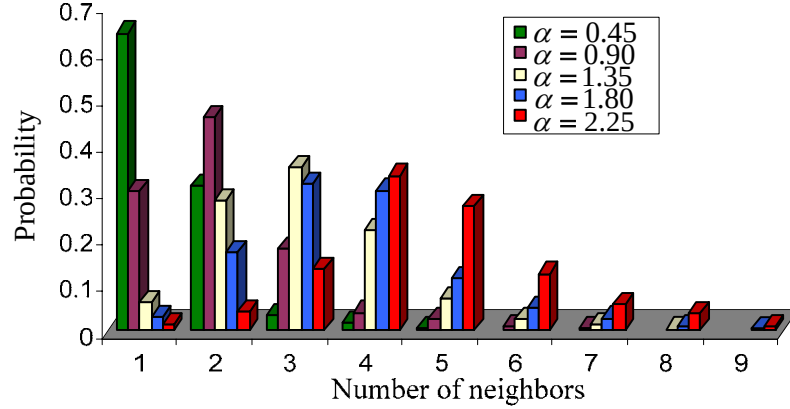


Figure 5.7. Probability density function of average neighbor node degree.

Figure 5.7 indicates that when α is 1.0 or less, the probability that a node has only one neighbor is high (about 65% when $\alpha = 0.45$ and 33% when $\alpha = 0.90$). In a network, a group of nodes that connect to each other through either one hop or multi-hop wireless communication are called a node cluster. If each node has a high probability of having only one neighbor, then there are many node clusters in the network. Each node cluster has no more than two nodes, and node clusters are disconnected from each other, then the network is highly partitioned. To make the network connected, the probability that nodes have at least two neighbors should be high. Hence both the simulation results using the ideal routing protocol and the neighbor node degree investigation indicate that the ad hoc network is highly partitioned when α is less than 1.0.

When the radio range ratio α is equal to or larger than 1.35, the simulation results indicate that the network is fully connected. Figure 5.7 indicates that when $\alpha = 1.35$, the probability that a node has more than two neighbors is very high, increasing to more than 90%. When radio range further increases, the probability that nodes have more than two

neighbors is even higher. The delivery of the message at these radio ranges is totally dependent upon the immediate forwarding of intermediate nodes, and the transit time is essentially zero for this study, since node-forwarding latency is assumed to be 0 in the ideal routing protocol.

Studies were also carried out to examine the effects of the initial velocity randomness k_v on GTI mobility model behavior. The GTI mobility model assumes that each node has a velocity $v = (1 + k_v) \cdot v_{av}$, where k_v is a random number uniformly distributed between $(-0.1 \sim 0.1)$ and v_{av} is the average node velocity, which is assumed to be same for all nodes. Simulations were performed with different percentages of velocity randomness as initial conditions. For these simulations, α is fixed at 1.0 and the average velocity v_{av} is equal to 30mi/hr.

Table 5.4. Transit Time vs Initial Velocity Randomness

k_v	Transit time (T_{trans})		
	Avg (s)	Max (s)	Standard deviation (s)
Uniform(-0.1,+0.1)	4376.1	5813.6	532.0
Uniform(-0.25,+0.25)	4561.4	6400.0	693.9
Uniform(-0.5,+0.5)	5122.0	6389.8	638.0

The simulation results in Table 5.4 show that different amount of initial velocity randomness applied in GTI mobility model does not have much effect on transit time performance. Though the standard deviation for transit time increases slightly, the average transit time is always near 5000 seconds. The average transit time is mainly determined by the average node velocity, while the amount of initial velocity randomness

has only minor effect.

The effects of the initial position randomness k_p on GTI mobility model behavior were also studied. The GTI model assumes that nodes are initially evenly distributed along each edge with some random deviation of initial positions, introduced by a random number k_p between $(-0.1 \sim 0.1)$. To study the effects of initial position randomness, simulations were performed with different percent of position randomness. In these simulations, α is fixed at 1.0 and the average velocity v_{av} is equal to 30mi/hr, with k_v uniformly distributed between $(-0.1 \sim 0.1)$. The simulation results are given in Table 5.5.

Table 5.5. Transit Time vs Initial Position Randomness

k_p	Transit time (T_{trans})		
	Avg (s)	Max (s)	Standard deviation (s)
Uniform(-0.1,+0.1)	4376.1	5813.6	532.0
Uniform(-0.25,+0.25)	4540.3	6464.4	733.2
Uniform(-0.5,+0.5)	4637.3	6230.5	617.7

The simulation results in Table 5.5 show that the transit time performance are quite similar for different amounts of initial node position randomness. This further indicates that initial node position randomness plays a less important role than the average distance among neighboring nodes.

Summary

In this chapter, a GTI mobility model has been proposed for modeling the movements of mobile nodes based on geographic and traffic information of the simulation area. This mobility model gives a closer match of the mobile node movements with real world scenarios. The outputs of the mobility model are trace-like trajectories for each mobile node in a specific simulation area that can be incorporated into OPNETTM or other network simulation tools for routing protocol performance evaluations.

In the simulation, at different values of α and different average node velocity, the transit time has been calculated by applying an ideal routing protocol to the mobile ad hoc network, and the neighbor node degree has been investigated as well. Increasing α from much less than 1 to bigger than 1 makes the network change from highly partitioned to fully connected. Both the simulation results and the neighbor node degree investigation indicate the same connectivity characteristic of the network. The behavior of the GTI mobility model has been further studied by varying the node initial velocity and position randomness and the consistent simulation results validate the GTI mobility model.

These studies based on the ideal routing protocol serve as foundation for the development of a novel practical routing protocol that can tolerate network partitioning and high node mobility that will be carried out in Chapter 6. The new protocol is tested by simulation and the results are given in Chapter 7, where the new routing protocol is evaluated based on not only the GTI mobility model but also on a random walk mobility model. The performance of the protocol is evaluated in detail and the effects of choice of the mobility model on the performance evaluation are further discussed in that chapter.

References

- [5.1] Online, <http://www.verizonwireless.com>.
- [5.2] R. Morris, J. Jannotti, F. Kaashoek, J. Li, and D. S. J. De Couto, "CarNet: A scalable ad hoc wireless network system," *9th ACM SIGOPS European Workshop: Beyond the PC: New Challenges for the Operating System*, Kolding, Denmark, Sep. 2000.
- [5.3] A. Ebner and H. Rohling, "A self-organized radio network for automotive applications," *Conf. Proc. ITS 2001, 8th World Congress on Intelligent Transportation Systems*, Sydney, Australia, Oct. 2001.
- [5.4] G. Ravikiran, S. Singh, "Influence of mobility models on the performance of routing protocols in ad-hoc wireless networks," *Vehicular Technology Conference, VTC 2004-Spring, IEEE 59th*, Volume 4, 17-19 May 2004, pp. 2185 - 2189 Vol.4
- [5.5] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *WCMC: Special issue on Mobile Ad Hoc Networking: Research, Trends and Applications*, 2(5): 483–502, 2002.
- [5.6] David B. Johnson and David A. Maltz, Dynamic source routing in ad hoc wireless networks. *Mobile Computing*, edited by Tomasz Imielinski and Hank Korth, chapter 5, pages 153–181. Kluwer Academic Publishers, 1996.
- [5.7] http://www.cs.ucsb.edu/~eroyer/txt/icc01_optden.ps
- [5.8] X. Hong, M. Gerla, G. Pei, and C. Chiang. "A group mobility model for ad hoc wireless networks," *In Proceedings of the ACM International Workshop on Modeling and Simulation of Wireless and Mobile Systems (MSWiM)*, August 1999.
- [5.9] Amit Jardosh, Elizabeth M. Belding-Royer, Kevin C. Almeroth, and Subhash Suri, "Towards realistic mobility models for mobile ad-hoc networking," *In Proc. of ACM Mobicom 2003*, pp.217–229, San Diego, CA, 2003.
- [5.10] F. Bai, N. Sadagopan, and A. Helmy, "IMPORTANT: A framework to systematically analyze the Impact of Mobility on Performance of Routing protocols for Ad hoc Networks", *Proceedings of IEEE INFOCOM*, 2003.
- [5.11] <http://mctrans.ce.ufl.edu/featured/TSIS/Version5/corsim.htm>
- [5.12] <http://www.yellowstone-natl-park.com/map.htm>
- [5.13] A. Vahdat and D. Becker, "Epidemic routing for partially connected ad hoc networks," Duke University, *Tech. Rep. CS-200006*, Apr. 2000.

CHAPTER 6
BORDER NODE BASED ROUTING PROTOCOL
FOR MANETS IN RURAL AREAS

Introduction

As discussed in Chapter 4, mobile ad hoc networks in rural areas have unique characteristics including low node density, high node mobility and high probability of network partition. Generally there is no fixed communication infrastructure that can be readily used to interconnect with the ad hoc network. The design of a routing protocol for such a mobile ad hoc network is challenging due to the unpredictable and highly dynamic topology. Under these circumstances, it is difficult or may not be possible to form a complete end-to-end multi-hop route which can last for the period of time needed to complete a data transmission.

The general approach to information transmission for partially connected ad hoc networks is to relay messages hop by hop, not necessary continuously, but at discrete time intervals as links become available. The data can be stored in intermediate nodes for some time before it can be further forwarded to another intermediate node or to the destination node. This is sometimes called the message relay approach. Generally a message may consist of one or several packets. With the message relay approach, data delivery may incur a long delay before it reaches the destination node. As a comparison, in a fully-connected mobile ad hoc network, if a route or path for a data packet is found, the delivery of the data packet will occur in a relatively short time, determined by a

combination of the propagation, processing and transmission delays. An ad hoc network that uses the generalized message relay idea is also called a Delay Tolerant Mobile Network (DTMN) [6.1].

Data delivery in a sparse mobile ad hoc network is generally based on the store-and-forward message relay approach [6.2]-[6.5]. The message ferrying approach was presented in [6.3], where a set of special mobile nodes called message ferries move around the deployment area according to known routes while other nodes transmit data to distant nodes out of range by using ferries as relays. A similar method is proposed in [6.4], where some nodes called data mules are used to collect data in a sparse sensor network. In that sensor network, nodes are generally static, not mobile, except for the data mules. Another approach using message relay was proposed in [6.5], in which mobile hosts actively modify their trajectories to transmit messages to minimize the transmission delay. All approaches mentioned above are mobility-assisted and proactive in nature since nodes modify their trajectories proactively to assist communication. However, it is not always the case that non-randomness in the movement of nodes can be exploited to help data delivery. For example, for vehicular ad hoc networks in sparse and rural areas, there are no vehicles that can serve as “message ferries”, and there is generally no repetition in the individual node’s trajectory.

Epidemic routing was first introduced for partially connected ad hoc networks in [6.6]. In the epidemic routing algorithm, random pair-wise exchanges of messages occur among mobile nodes whenever they meet each other. The movement inherent in the mobile nodes themselves is exploited to help deliver the data to the destination when the network is partially connected. The epidemic algorithm is flooding based, and it trades

system bandwidth and node buffer space for the eventual delivery of a message.

To control flooding or save system bandwidth and node buffer space, different flooding control schemes have been proposed [6.7]-[6.10]. A comparison of broadcast techniques in mobile ad hoc network is given in [6.11]. However, control schemes proposed in [6.7]-[6.10] all assume that nodes have some knowledge or history about other nodes in advance. Opportunistic forwarding strategies using the position information obtained through a GPS system are explored in [6.7]. Probabilistic metric “delivery predictability” is explored in [6.8] to select the better next step candidates. The “delivery predictability” function is based on the history of encounters, assuming nodes know how many times they encounter other nodes. Similarly, a forwarding decision based on the “utility function” is proposed in [6.9], in which more information about other nodes, including the nodes recently noticed and the most frequently noticed, the power level, the rediscover interval etc., are used to calculate the utility function. An opportunistic exchange algorithm using a spatio-temporal relevance function to manage the nodes buffer space was proposed in [6.10], focusing on resource discovery in urban areas, such as finding an empty park lot or notification of a traffic jam.

Flooding control schemes based on knowledge or history information about other nodes, as discussed above, are not appropriate to application scenarios in rural areas. The low node density, combined with the difficulty of obtaining the information used in the routing determinations limits the effectiveness of these schemes. Furthermore, the assumption that nodes will have GPS-based location information is an additional constraint and there may not be repetition in node trajectories as needed in some of the approaches.

In this chapter, a Border node Based Routing (BBR) protocol is proposed for mobile ad hoc network in sparse and rural areas. The BBR routing protocol is mainly based on broadcast and it applies the store-and-forward approach used in epidemic routing. However, instead of simply flooding the network, a flooding control scheme will be explored that uses one-hop neighbor information. The epidemic routing algorithm proposed in [6.6] also did not take into account real application oriented characteristics such as high node mobility in a vehicular ad hoc network. The effects of high mobility on neighbor discovery and further data delivery will also be explored in BBR routing.

In the remainder of this chapter, the general concept of the BBR routing protocol will be introduced and then the routing algorithm will be presented in detail. The conceptual data structures used in BBR routing will then be discussed. Simulation results and comparisons of performance with other routing protocols are presented in chapter 7.

BBR Routing Protocol Overview

The BBR routing protocol can be used for sending messages from any node to any other node (unicast) or from one node to all other nodes (broadcast) in a mobile ad hoc network. It is based mainly on broadcast, but unicast is also used in certain scenarios to save bandwidth. The delivery of unicast messages is realized by combining the use of broadcast and unicast.

The general design goals of the BBR routing protocol are to optimize the broadcast behavior for ad hoc networks with low node density and high mobility and to deliver messages with high reliability while minimizing delivery delay.

The BBR routing protocol has two basic functional units. One is the neighbor

discovery algorithm, and the other is the border node selection algorithm. The neighbor discovery process is responsible for collection of current one hop neighbor information for each node at any time. The border node selection process is responsible for selection of the right candidate/candidates for further forwarding of a particular packet. The decision process is based mainly on the one-hop neighbor information collected in the neighbor discovery process.

In the following section, the general assumptions that the BBR routing protocol is based on will first be briefly discussed. Then the neighbor discovery algorithm and border node selection algorithm will be described in detail.

Assumptions

The protocol design is based on the following assumptions that include consideration of the characteristics of mobile ad hoc networks in rural areas. First, no node location information is available. Second, the only communications paths available are via the ad-hoc network itself. There is no other communication infrastructure. Third, node power is not the limiting factor for the design. Fourth, communications are message oriented. Real time communication traffic is not supported. Fifth, each node in the ad-hoc network has the same radio transmission range.

Although the original intent is to design a routing protocol that can best fit partially connected mobile ad hoc networks with mobility patterns similar to vehicular traffic constrained to roadways, the BBR routing protocol can be used in any ad hoc network with any node mobility pattern

Neighbor Discovery Algorithm

Neighbor discovery is the process whereby a node discovers its current neighbors. For a particular mobile node, any other node that is within radio transmission range is called a neighbor. All the neighbors of a particular mobile node consist of a neighbor set. Here, only one-hop neighbors are considered. Since all nodes may be moving, the neighbors for a particular mobile node are always changing. The neighbor set is dynamic and needs to be updated frequently.

Generally, neighbor discovery is realized by using periodic Hello messages for neighbor node detection. Each node informs other nodes of its existence by sending out Hello messages periodically. A node updates its neighbor node set after receiving Hello messages from other nodes.

The neighbor discovery algorithm used in BBR routing is similar to the neighbor discovery protocol (NDP) proposed in the Zone Routing Protocol (ZRP) [6.12]. The NDP in ZRP is a MAC-level based neighbor discovery protocol where a periodic Hello beacon is sent out by the node MAC layer to advertise its existence. The neighbor discovery algorithm in BBR routing is a network layer-level based NDP. The Hello message is sent out by network layer. The advantage of using network layer based NDP is that all routing functions are accomplished in the network layer, without consideration of the specific MAC layer technology used.

Neighbor Discovery Process. Every node broadcasts a Hello message periodically and maintains a Neighbor Table to keep its current one-hop neighbor information. The time interval between two successive Hello messages is called the Hello interval. This

time interval also determines how frequently a node updates its Neighbor Table. To avoid Hello message collisions, nodes act independently and send out Hello messages asynchronously.

A Hello message is a very simple packet with the node ID as its only content, as shown in Figure 6.1. The header of the Neighbor Table is indicated in Figure 6.2. If a Hello message is received, the Boolean value of “Arrival” will be set to be true for that node entry. If a Hello message is received for the first time from a new neighbor, the value of “Last Recorded” will be set to be -1 to indicate that this is a new table entry. If a Hello message from a neighbor is received, and the neighbor is already in the Neighbor Table, then the value of “Last Recorded” will be left unchanged during the Hello message recording stage. The flow diagram for the message recording procedure is indicated in Figure 6.3.

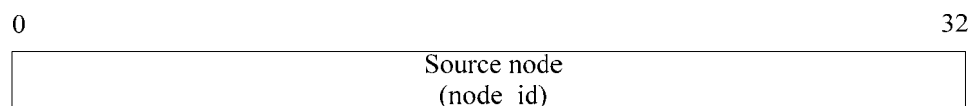


Figure 6.1. Packet format of Hello message.

Neighbor node ID (node_id)	Arrival (Boolean)	Last recorded (int)
-------------------------------	----------------------	------------------------

Figure 6.2. Table header for Neighbor Table.

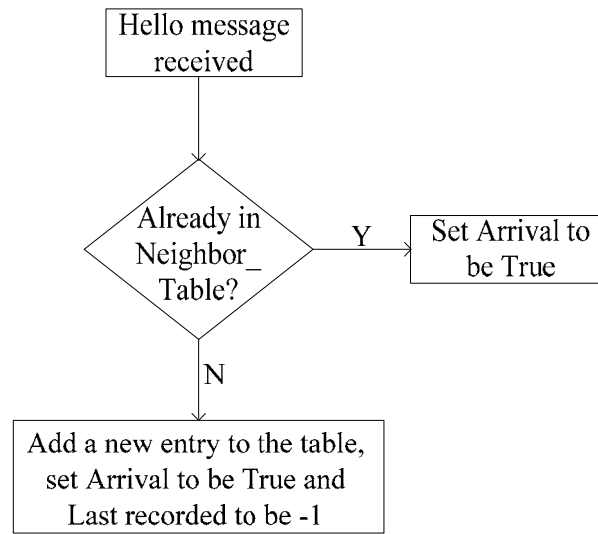


Figure 6.3. Hello message information recording procedure.

At time points separated by one Hello interval, a node updates its Neighbor Table according to all the Hello messages received during the previous Hello interval. A node records received Hello message information to its Neighbor Table at any time, but it updates its current neighbor information only at fixed time points separated by the Hello interval. If a node doesn't receive a Hello message from a neighbor node for a MaxHelloLoss number of times, the node will consider that the neighbor is lost and delete that node from its Neighbor Table. The value of MaxHelloLoss can be configured to account for different amounts of node mobility. Figure 6.4 gives the pseudo code used to update a Neighbor Table, which is mainly based on that of NDP in ZRP [6.12], with some modifications.

```

for each entry in the Neighbor Table
{
  if (Arrival == False)
  {
    if (Last recorded >= MaxHelloLoss)
    {
      remove that neighbor entry;
    }
    else
    {
      Last recorded ++;
    }
  }
  else
  {
    if (Last recorded == -1)
    {
      new neighbor found++;
    }
    Arrival = FALSE;
    Last recorded = 0;
  }
}

table_update_timer += Hello_interval;

If (new neighbor found > 0)
{
  set interrupt("new neighbor/neighbors found");
}

```

Figure 6.4. Pseudo code for updating a Neighbor Table.

In the neighbor discovery process, one very important parameter is the Hello interval. This parameter determines how frequently Hello messages are sent out and also how frequently a Neighbor Table will be updated and subsequently how accurately the Neighbor Table information reflects the real neighbor situation of a particular node. The effect of the Hello interval setting on the packet delivery ratio will be discussed in detail in Chapter 7 with the simulation results.

Distributed Border Node Selection Algorithm

In BBR routing, border nodes are selected per broadcast event. A border node has the responsibility of saving received broadcast packet/packets and forwarding the packet/packets when appropriate. Also, a border node is generally located at the edge of the radio transmission range for a specific broadcast from a source node. For a group of nodes that receive the same broadcast message, only those nodes selected to be border nodes will keep the received data information and rebroadcast it later when those nodes meet new neighbors. The overall objective is to the broadcast message as far as possible and as soon as possible while minimizing the total number of nodes involved in rebroadcast. BBR uses a distributed algorithm, the decision whether a node is a border node or not for a particular broadcast event is made independently by an individual node based on its one-hop neighbor information and the received broadcast information.

Broadcast Based Routing Algorithm. Broadcast is very commonly used in wired IP networks. In one definition of broadcast by flooding or pure flooding, a node or a host, on receiving a broadcast message for the first time, has the obligation to rebroadcast the message. In a mobile ad hoc network, one main disadvantage of the pure flooding scheme is that it can lead to the broadcast storm problem [6.13]. This is partially due to the nature of the wireless medium and the use of the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) mechanism by the 802.11 Medium Access Control (MAC) layer standards. Other disadvantages include waste of bandwidth and the possible loss of the broadcast message due to collisions as well as network partitioning.

To improve broadcast efficiency in a mobile ad hoc network, a cluster-based

broadcast scheme is proposed in [6.13], where a mobile network is separated into clusters, and there are gateway nodes in each cluster. Gateway nodes forward the broadcast information. The cluster-based broadcast scheme has high efficiency for fully connected ad hoc networks, but for a partially connected ad hoc network, the broadcast information might easily get lost because when two clusters are disconnected, the broadcast information can't propagate to the other cluster. Another broadcast scheme proposed for mobile ad hoc networks is multipoint relaying [6.14], in which a smallest set of nodes in the network called multipoint relays are selected. The rebroadcast by these multi-point relays will make the whole network be covered for a broadcast event. Again, this broadcast scheme only works for fully connected ad hoc networks. In BBR, the broadcast scheme is based on border node selection. Part of the proposed border node selection algorithm is similar to the self-pruning scheme proposed in [6.15]. However, the self-pruning scheme is not specifically designed for mobile ad hoc networks and it doesn't provide a solution to the possible broadcast storm problem. A brief comparison of the cluster-based broadcast scheme, the multipoint relay scheme and the proposed BBR routing algorithm is given in Table 6.1.

Table 6.1. Comparison of Multipoint Relay Scheme, Cluster-Based Broadcast Scheme & BBR Routing Scheme

Schemes Items	Multipoint relaying	Cluster-Based Broadcast	BBR Routing Scheme
Neighbor info needed	One hop and two hop neighbor info	Neighbor table, 2-hop-topology link state table	One hop info
Processing	Multipoint relays selection	Gateways, cluster head nodes selection	Border node selection
Reliability	Guaranteed delivery to all the two hop neighbors	Low reliability when the network is highly dynamic and partition	High reliability for network with dynamic and partition
Forwarding decision	Multipoint relay nodes forward only	Gateway nodes forward only	Border nodes forward only
Forwarding time	Right after the receiving of the broadcast message	Right after the receiving of the broadcast message	Store and forward
Data packets	Data info	Source router header + Data info	Data info + list of the one hop neighbor
Control packet	Hello message with list of its one hop neighbor	Hello message + neighbor list + adjacent cluster ID	Hello message

Heuristic for the Selection of Border Nodes. Based on intuition, for a specific broadcast, an ideal candidate to forward a packet would be a node that is located at the edge of the radio transmission range of the source node. To select out the border node based only on one-hop neighbor information, two approaches can be explored: one is based on the minimum common neighbor concept and the other is based on the maximum uncommon neighbor concept. Before further exploring these two approaches, some terms used frequently are first introduced.

Covered: A node j is covered by a node k when the node j is within the direct radio transmission range of the node k and can receive packets from node k when node k sends out packet. When node j receives a packet from node k , node j

is covered by the broadcast of node k . As radio transmission is reciprocal, if a node j is covered by node k , then node k is covered by node j .

Neighbor set: A one-hop neighbor set of node i , noted as N_i , consists of those nodes that are covered by node i .

Common neighbor set: An arbitrary node i has a one-hop neighbor set N_i .

Another node s has a one-hop neighbor set N_s . A third node j is called a common neighbor of node i and node s , if and only if $j \in N_i$ and $j \in N_s$.

The common neighbor set of node i and s , noted as N_{is} consists of all common neighbors of node i and node s .

Uncommon neighbor set: For two nodes i and s , the uncommon neighbor set U_{is} is defined as $N_i - N_{is}$. The uncommon neighbor set U_{si} is defined as $N_s - N_{is}$.

1) *Border node selection based on minimum common neighbors:* In this approach, a border node is selected based upon the intuitive notion that nodes at the edge of radio transmission range should have fewer common neighbor nodes with the broadcast source node, as compared to those nodes that are closer to the source node. Nodes with minimum common neighbors with the broadcast source node will be chosen as border nodes. As indicated in Figure 6.5, a circle delineates the direct radio transmission range of the node located at the center of the circle. R is the radio transmission range. The red circle indicates the radio transmission range of node s . Similarly, the blue circle, brown circle and yellow circle indicates the radio transmission ranges of node a , node b and

node h respectively. For example, suppose node s is a broadcast source. Nodes at the edge of the radio transmission range, such as node b and node h , as compared to nodes closer to the broadcast source node, such as node a , have fewer common neighbors with node s . Selection of nodes such as b or h or both as the border nodes for further rebroadcast is appropriate in that this selection results in maximum range and rapid information dissemination while saving bandwidth by minimizing unnecessary rebroadcasts.

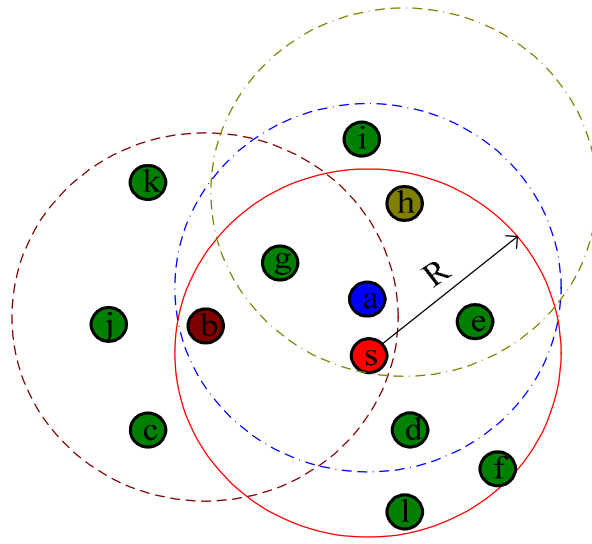


Figure 6.5. A typical broadcast and node distribution in a connected network

2) *Border node selection based on maximum common neighbors*: In this approach, a border node is selected based upon the intuitive notion that nodes at the edge of radio transmission range might have more uncommon neighbor nodes with the broadcast source node, as compared to those nodes that are closer to the source node. This observation is true when the network is fully connected. Under this network condition, border node selection based either on maximum common neighbors or on minimum uncommon neighbors is basically equivalent. For example, as indicated in Figure 6.5,

using either approach, node b will be selected as a border node. However, for a partially connected network, clusters of nodes are always disconnected with each other. A simple example is given in Figure 6.6, where the nodes cluster composed of nodes c, f, j and k is disconnected with the nodes the cluster centered on node s . Under this network condition, using the maximum uncommon neighbor approach, no border node can be selected as all neighbors of node s have zero uncommon neighbors with node s . Using minimum common neighbor approach, a node at the edge of the radio transmission range of node s , such as node b , can still be selected as a border node. In other words, the maximum uncommon neighbor approach does not apply well to a partially-connected ad hoc network for border node selection.

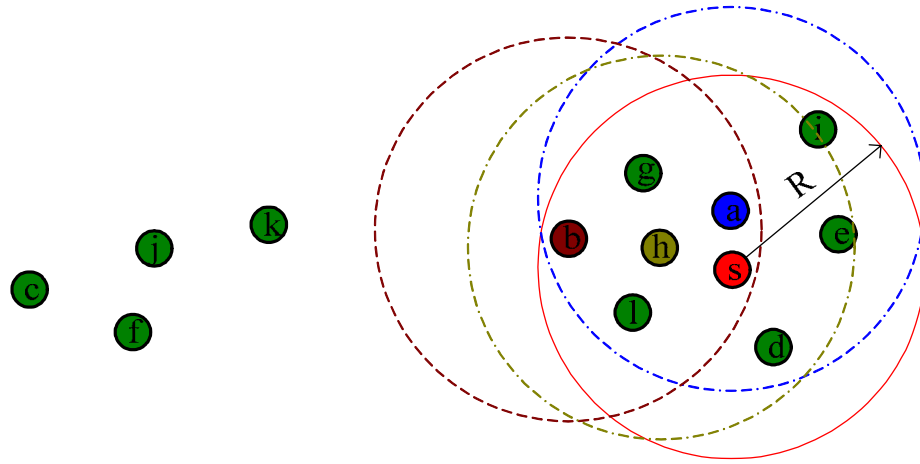


Figure 6.6. A typical broadcast and node distribution in a partially connected network.

As the routing algorithm should work well not only for a fully connected ad hoc network, but more importantly, it should have good performance for a partially-connected ad hoc network, we conclude that the border node selection algorithm based on the minimum common neighbor approach is preferable

Border Node Selection Scheme and Rules. When a node has a packet to forward and there is no available neighbor, it keeps the packet in the forward buffer and broadcasts it later when there are neighbors in range. A source node that generates a data packet is by default chosen as a border node. A node that broadcasts or rebroadcasts a data packet will use a packet format as indicated in Table 6.2, which has a list of its current neighbors attached. The “Common neighbor #” field is set to be the number of the common neighbors between the current node and the previous node that broadcasts the data packet. If a node is originating a data packet, then the “Common neighbor #” field is set to be zero. Each packet has its unique packet ID, generated by the originating node. The packet ID remains unchanged as the packet moves from source to destination.

Table 6.2. Broadcast Data Packet Format

Source node ID	Destination node ID	Common neighbor #	Reserved	Packet id	Packet content	Neighbor list
----------------	---------------------	-------------------	----------	-----------	----------------	---------------

With the BBR routing algorithm, every node has three tables/buffers: the Neighbor Table, the Border Node Selection Table and the Forward Table. The Neighbor Table is used to save the information about current one hop neighbors. The Border Node Selection Table is used to select the border node. The Forward Table is used to buffer data packets for future forwarding. When a node receives a data packet, it first searches its Forward Table to see whether there is already a packet entry with the same packet ID. If there is a data packet with the same packet ID already in the Forward Table, then the data packet will be ignored. Otherwise, the node checks the neighbor list of the received data packet and follows the following procedures based on the following cases.

Case 1: Single neighbor on the neighbor list of the broadcast message. The neighbor node is the only node on the neighbor list. Then no border node selection will be made and it is a border node by default. The node will check its one hop current neighbor list. If there are no additional nodes on this list, then it will store the data packet in its Forward Table. It will carry this data packet and rebroadcast for a total of p times at different time points in the future when there are new neighbors within its transmission range. The rebroadcast parameter p is configurable and indicates the willingness of intermediate nodes to forward a data packet. If this node has additional neighbor nodes within range, then it will rebroadcast immediately and rebroadcast $p - 1$ times in the future when new neighbors coming into range.

Case 2: Multiple neighbors on the neighbor list of the received broadcast data packet. There are multiple neighbors on the neighbor list. Those nodes receiving the data packet for the first time will initiate two timers, an access delay timer T_{ad} and a maximum delay timer T_{max} . The access delay timer T_{ad} is used to decide when a node needs to rebroadcast if it has to do so. The maximum delay timer T_{max} is used to decide when a node should initiate the border node selection process. The values of these two timers will be discussed in the section where the detailed operation of the protocol is described. The value of T_{ad} is node dependent, while the value of T_{max} is the same for the group of nodes receiving the same broadcast data packet. During the T_{max} time interval, each node in the group decides to rebroadcast or not when its T_{ad} timer expires. The decision is made depending on whether all its current one hop neighbors are covered or not, namely, whether they have received the broadcast packet information or not. If a

node needs to rebroadcast when its T_{ad} timer expires, the rebroadcast packet will have a format as shown in Table 6.3, which looks almost the same as that of the data packet format indicated in Table 6.2. However, the values of the fields “Node ID”, “Common Neighbor #” and “Neighbor list” are all different from that of the received broadcast packet. During the whole T_{max} interval, each node will listen continuously. Rebroadcast packets from its neighbors will also be recorded and saved temporally for the use by the border node selection procedure. At the end of the T_{max} time interval, a node will check whether it is the node with the least common neighbor number with the previous broadcast source based on all packets received and recorded in its Border Node Table, if it is, it will select itself as a border node. Otherwise it is not a border node. As a node can only receive packets from the source node and from its common neighbors with the source node, the least common neighbor comparison is carried out between itself and its common neighbors with source node.

Table 6.3. Rebroadcast Packet Format

Node ID	Destination node ID	Common neighbor #	Reserved	Packet ID	Packet content	Neighbor list
---------	---------------------	-------------------	----------	-----------	----------------	---------------

To illustrate the border node selection process in Case 2, an example network is given in Figure 6.7 with the details of the border node selection process shown in Table 6.4.

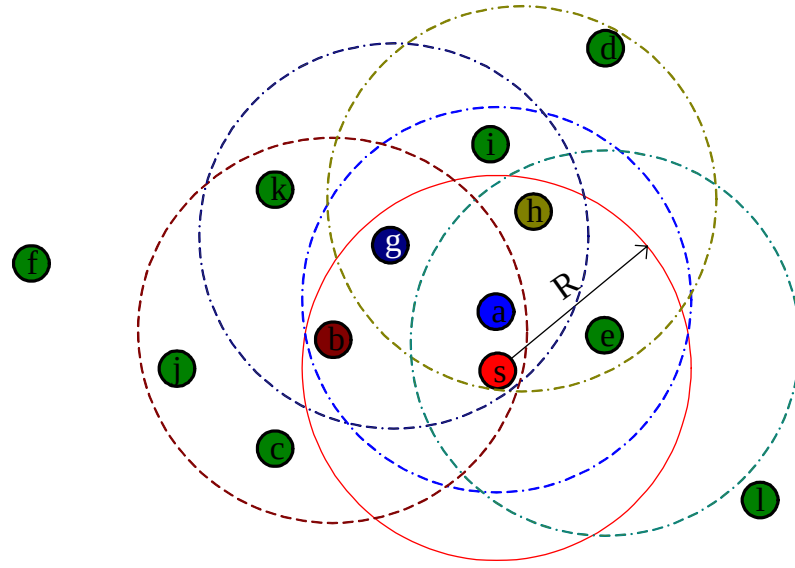
Figure 6.7. A typical network with a broadcast source node s .

Table 6.4. Border Node Selection Process Illustration

Node	Neighbor set	Common neighbor set	Common neighbor number	Node action when T_{ad} expires	Border node decision when T_{max} expires
s	$\{h, e, g, b, a\}$				
h	$\{s, a, g, e, i, d\}$	$\{a, g, e\}$	3	Rebroadcast	yes
e	$\{s, a, h\}$	$\{a, h\}$	2	Keep silent	yes
g	$\{s, a, b, h, i, k\}$	$\{a, b, h\}$	3	Rebroadcast	no
b	$\{s, a, g, j, c, k\}$	$\{a, g\}$	2	Rebroadcast	yes
a	$\{s, b, g, e, h, i\}$	$\{b, g, e, h\}$	4	Keep silent	no

Border node function. Every node selected to be a border node will save the received packet in its Forward Table. When a new neighbor come into range, the border node will rebroadcast the packet for a total of $p-1$ ($p \geq 1$) or p times at different time points according to whether it has rebroadcast or not when its T_{ad} timer expires. The exact value for p reflects the willingness of intermediate nodes to forward a specific data packet.

Conceptual Data Structures

This section describes several data structures and provides an overview of their use in the BBR routing protocol. Each node implementing BBR must maintain a Neighbor Table, a Border Node Selection Table and a Forward Table. The Neighbor Table has already been described in detail in the section that discusses the Neighbor Discovery Algorithm. Here the Border Node Selection Table and the Forward Table data structures are introduced.

Border Node Selection Table

The Border Node Selection Table is used when a node needs to decide whether it is a border node. In order to decide whether a node is a border node or not for a specific broadcast event, during the $T_{\max}$ time interval, a node will temporally save all packets received with the same packet ID in its Border Node Selection Table. In other words, the border node is selected per packet ID. The header of the Border Node Selection Table is indicated in Table 6.5. At one time, there might be several border node selection processes going on, each corresponds to a different packet ID. Each row under the table header in the Border Node Selection Table is called a table entry.

Table 6.5. Border Node Selection Table

Packet ID	Forward Counter	Common Neighbor #	Neighbor List	Rebroadcast	Evhandle_1	Evhandle_2	Packet Content
...	...	...	...	...	...	...	...

The Border Node Selection Table is realized by using a hash table. The final destination address of the data packet is chosen as the key to the hash table. The Forward

Counter is records how many hops that a specific packet has been forwarded. Counters Evhandle_1 and Evhandle_2 are used to realize the two timers T_{ad} and T_{max} respectively. An incoming packet is put into the Border Node Selection Table and saved as a “source broadcast” entry if there is no packet with the same packet ID in the current table. The “Rebroadcast” field indicates whether the node has rebroadcast or not when T_{ad} expires. If a node has rebroadcast when T_{ad} expires, this field will be set to be TRUE, otherwise it will set to be FALSE. This packet entry will initiate two timers and also cause the packet content to be saved. A packet arriving later with the same packet ID as the “source broadcast” packet will be stored in the Border Node Selection Table as a “rebroadcast” entry without setting the two timers and saving the packet content.

A node makes a decision whether it should rebroadcast or not when the first timer T_{ad} of the “source broadcast” entry expires. When the second timer T_{max} expires, a node will decide whether it is a border node or not. After the border node selection process is finished, all packets with that specific packet ID are removed from the Border Node Selection Table.

Forward Table

The Forward Table is used to save those packets that need to be rebroadcast at least once by a node. Table 6.6 gives the header of the Forward Table.

Table 6.6. Forward Table Header

Packet ID	Forward Counter	Rebroadcast Counter	Packet Content	Insert Time
...	...	...	...	...
...	...	...	...	...

Every time a node meets a new neighbor, it will first check whether the new neighbor is the destination node of some packets saved in its Forward Table. If so, the packets are unicast to the destination node and removed from the Forward Table. Otherwise the node will pick a sequence of data packets that are destined to the same destination node from its Forward Table to rebroadcast. The destination node can be selected randomly or by using a First In First Out algorithm. The Forward Counter records how many times the packet has been forwarded by intermediate nodes. Every time a packet is rebroadcast, the corresponding Rebroadcast Counter is decreased by 1. If the Rebroadcast Counter decreases to zero, the packet will be removed from the Forward Table. The value of the Rebroadcast Counter is initialized to the value of $p - 1$ or p when a packet is put into the Forward Table depending on whether the “Rebroadcast” field of the border node select entry is TRUE or FALSE respectively. The “Insert Time” field records the time when the packet is saved in the Forward Table, which can be used for Forward Table resource management when the buffer space is limited. It will also be used to prevent the repetition transmission problem incurred by two nodes that may discover each other at different future time points. The repetition transmission problem will be further discussed at the end of this chapter.

BBR Options Header Format

Nodes using the BBR routing protocol employ a specific BBR data packet or control packet format. These packets have an additional option header. A BBR packet is encapsulated into a regular IP packet before it is sent out to the network. This section describes the BBR Options Header format and four types of BBR options are defined for

use within the BBR Options Header.

Fixed Portion of the BBR Options Header

Similar to the DSR routing protocol, the fixed portion of the BBR Options Header is used to carry information that must be present in every BBR Options Header. Figure 6.8 shows the format of the fixed portion of the BBR Options Header. This BBR Options Header follows the IPv4 header for every BBR packet.

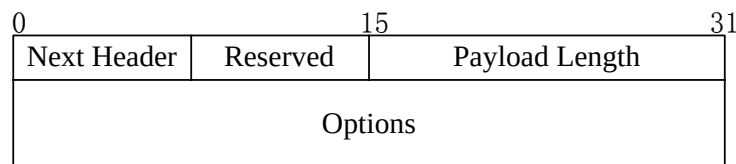


Figure 6.8. The fixed portion of the BBR Options Header format.

Next Header: The Next Header is an 8-bit selector field that identifies the type of header immediately following the BBR Options Header. It uses the same value as in the “Protocol” field of IPv4 header, which is indicated in Figure 6.9.

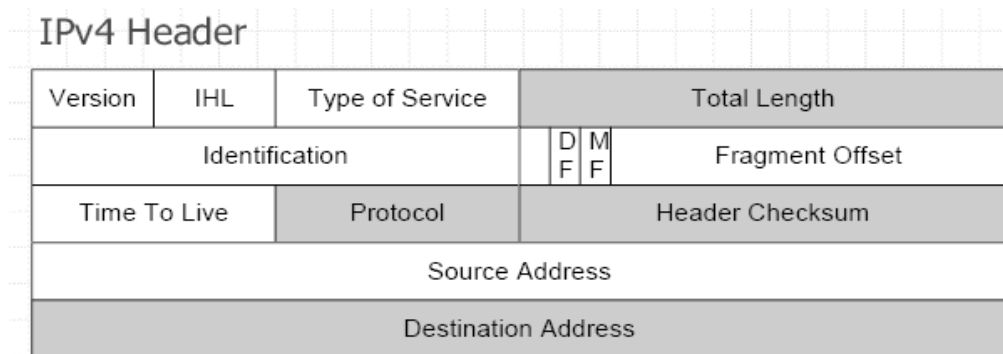


Figure 6.9. The IPv4 header.

Reserved: This field is reserved for future use and must be sent as 0 and is ignored on reception.

Payload Length: This field is the length of the BBR Options Header, excluding the

4-octet fixed portion. In DSR the value of the Payload Length field defines the total length of all options carried in the DSR Options Header, since in DSR one packet can carry several options at the same time. For BBR this field is not needed as only one option per packet is considered. Hence this value will be the same as the value of the “Option Data Length” field in the “Options” field. It is included here for convenience and for later possible expansion.

Options: This is a variable-length field. The length of the Options field is specified by the Payload Length field in this BBR Options Header. It contains one type of BBR option, encoded in type-length-value (TLV) format.

The following types of BBR options are defined for use within a BBR Options Header:

Neighbor Discovery Option (Option 1).

Data and Neighbor Destination Option (Option 2).

Data and None Neighbor Destination Option (Option 3).

Data Acknowledgement option (Option 4).

Neighbor Discovery Option (Option 1)

The Neighbor Discovery Option in a BBR Options Header is encoded as indicated in Figure 6.10. Those fields of the Neighbor Discovery Option and some fields of the IPv4 header are set with the values specified in the following discussion

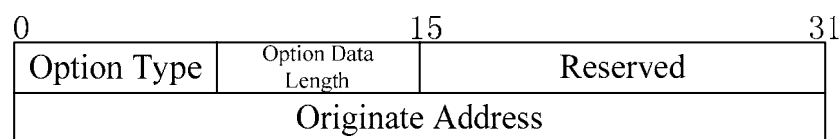


Figure 6.10. The Neighbor Discovery Option header.

Fields of the Neighbor Discovery Option:

- 1) Option Type: 1;
- 2) Option Data Length: This is an 8 bit unsigned integer. It is the length of the option, in octets, excluding the Option Type and Option Data Length fields. For the current definition of the Neighbor Discovery Option, this field must be set to 6 octets.
- 3) Originate Address: This is the IP address of the sending node.

IP fields:

- 1) Source Address: This must be set to the IP address of the node sending this packet.
- 2) Destination Address: This must be set to the IP limited broadcast address (255.255.255.255).
- 3) Hop Limit (TTL): This must be set to 1.

Data and Neighbor Destination Option (Option 2)

When a node originates or receives a data packet and finds that the destination of the data packet is one of its current neighbors, it will unicast this data packet directly to that destination node instead of using broadcast. The Data and Neighbor Destination Option in a BBR Options Header is encoded as indicated in Figure 6.11.

0		15		31	
Option Type	Option Data Length	Reserved		Forward Count	
Identification (Packet ID)					
Final Destination Address					

Figure 6.11. Data and Neighbor Destination Option header.

Fields of the Data and Neighbor Destination Option:

- 1) Option Type: 2;
- 2) Option Data Length: This is an 8-bit unsigned integer. It is the length of the option, in octets, excluding the Option Type and Opt Data Len fields. For the current definition of the Data and Neighbor Destination Option, this field must be set to 14 octets.
- 3) Forward Count: This is an 8-bit unsigned integer. It records how many times the packet has been forwarded before reaching the final destination node.
- 4) Identification (Packet ID): This is a unique value generated by the initiator (original sender) of the data packet. The node initiating a new data packet generates a unique identification value for each data packet. This identification number can be a concatenation of the node's ID (the least significant 16 bits of the IP address for the node) and a locally generated packet ID (16 bits). This value allows a forwarding node to determine whether it has recently seen a copy of this data packet. If this Packet ID is found by a forwarding node in its Forward Table, this forwarding node will decrease the Rebroadcast Counter value of the packet entry by 1. If at this time, the Rebroadcast Counter is zero, it then removes the packet from the Forward Table.
- 5) Final Destination Address: This is the final destination IP address of the data packet.

IP fields:

- 1) Source Address: This must be set to the IP address of the node sending this packet. It can be the IP address of the node that originates the data packet or the

IP address of an intermediate node that forwards the data packet.

- 2) Destination Address: This is the same address as in the Data and Neighbor Destination Option's "Final Destination Address".

Data and None Neighbor Destination Option (Option 3)

When a node finds that the destination node of a data packet is not in its current neighbor list and needs to broadcast or rebroadcast the data packet, the node will broadcast the data packet with its current one-hop neighbor list attached by padding the Data and None Neighbor Destination Option. The Data and None Neighbor Destination Option in a BBR Options Header is encoded as in Figure 6.12.

0	15	31
Option Type	Option Data Length	Comm. Neigh. # Forward Count
Identification (Packet ID)		
Final Destination Address		
Neighbor Address 1		
Neighbor Address 2		
:		
Neighbor Address n		
:		

Figure 6.12. Data and None Neighbor Destination Option header.

Fields of Data and None Neighbor Destination Option:

- 1) Option Type: 3.
- 2) Option Data Length: This is an 8-bit unsigned integer. It is the length of the option, in octets, excluding the Option Type and Option Data Length fields. For the current definition of the Data and None Neighbor Destination option, this field must be set to $(n + 3) \times 4 + 2$ octets.

- 3) Common Neighbor #: This is an 8-bit unsigned integer. It records how many common neighbors there are between this node and the node previously broadcasting the packet. If this node originates a data packet or rebroadcasts a data packet from its Forward Table, this value will be set to 0.
 - 4) Forward Count: This is an 8-bit unsigned integer that records how many times the packet has been forwarded before it reaches the destination.
 - 5) Identification: Same as in Option 2.
 - 6) Originate Address: This is the IP address of the source node that first generates the data packet.
 - 7) Final Destination Address: This is the destination IP address of the data packet
 - 8) Neighbor Addresses [1... n]: These are IP addresses on this node's neighbor list
- IP fields:*

- 1) Source Address: This must be set to the IP address of the node forwarding this packet.
- 2) Destination Address: This must be set to the IP limited broadcast address (255.255.255.255).

Data Acknowledgement Option (Option 4)

This option is optional and is not needed for the BBR routing protocol to work. A data acknowledge packet with the Data Acknowledgement Option is a control packet. In order to reduce the complexity of the BBR routing protocol it is not desirable to process a control packet in the same way that a data packet is processed for border node selection. However, if conventional broadcast is used to deliver control packets, substantial

overhead traffic will result and consume considerable bandwidth when the network becomes fully connected. This is contrary to the original design goal of BBR routing, which in part is to optimize the broadcast behavior by selecting only border nodes for packet forwarding. On the other hand, the use of a data acknowledgement packet will make the hidden node problem severe when a node tries to send multiple packets to the same destination node at a time.

The purpose of using a data acknowledgement packet is to remove those copies of a received data packet that might still exist in the Forward Tables of those intermediate nodes when a data packet actually reaches its final destination node. As discussed in the assumptions for BBR routing, if the buffer space is assumed big enough, then a data acknowledgement packet does not need to be used as the packet will ultimately be discarded from the Forward Table after being forwarded for a maximum of p times. p is a configurable parameter that can be adjusted to optimize the network performance under different forwarding conditions. However, this option is still considered and also implemented here for future possible improvement or modification to adjust to specific application scenarios. This option can be switched off if necessary during simulation to achieve a better packet delivery rate. A data acknowledgment packet, if used, will have a Data Acknowledgement Option inserted into the BBR Options Header. The Data Acknowledgement Option is encoded as indicated in Figure 6.13:

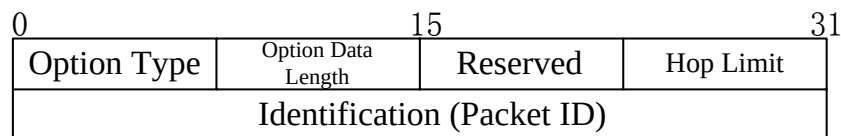


Figure 6.13. Data Acknowledgement Option header.

Fields of the Data Acknowledgement Option:

- 1) Option Type: 4.
- 2) Option Data Length: This is an 8-bit unsigned integer. Length of the option, in octets, excluding the Option Type and Option Data Length fields.
- 3) Hop Limit: The Hop Limit is the number of times the Data Acknowledgement packet can be rebroadcast before it should be discarded. This value should copy the Forward Count value of the acknowledged data packet.
- 4) Identification (Packet ID): This value is copied from the Identification (Packet ID) field of the packet being acknowledged.

IP fields:

- 1) Source Address: This must be set to the IP address of the node forwarding this packet.
- 2) Destination Address: This must be set to the IP limited broadcast address (255.255.255.255).

Detailed Operation of the BBR Routing Protocol

In this section, the general packet processing for the BBR routing protocol is described in detail. General packet processing describes generation of control and data packets, how a BBR packet is encapsulated into a regular IP packet, and packet handling procedures associated with the different BBR options when a BBR packet is received.

General Packet Processing

Originating a Packet. When originating a Hello packet, a node using BBR routing

will add a BBR Neighbor Discovery Option to the packet and then use one hop limited broadcast. When originating a Data Acknowledge packet, a node using BBR routing will add a BBR Data Acknowledgement Option to the packet and then use one hop limited broadcast. When originating a data packet, a node using BBR routing must perform the following sequence of steps:

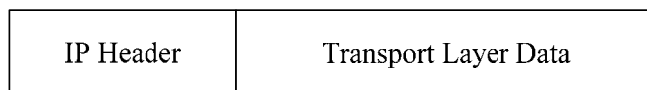
- 1) Search the node's one hop neighbor list for the address given in the IP Destination Address field in the packet's IP header.
- 2) If the IP Destination Address is in its neighbor list, which means that the destination node is only one hop away, then the node unicasts this packet to that neighbor directly by adding the Data and Neighbor Destination Option.
- 3) If the IP Destination Address is not in its neighbor list, then the node saves one copy of the packet in its Forwarding Table and broadcasts the packet by adding the Data and None Neighbor Destination Option. When broadcasting the packet, the node replaces the IP Destination Address field with the IP "limited broadcast" address (255.255.255.255) and copies the original IP Destination Address to the Final Destination Address field of the new Data and None Neighbor Destination Option added to the packet.
- 4) The node then broadcasts the packet.

Add a BBR Options Header to a Packet. Similar to the DSR protocol, a node originating a packet adds a BBR Options Header to the packet to carry information needed by the routing protocol. A packet must not contain more than one BBR Options Header. A BBR Options Header is added to a packet by performing the following sequence of steps (these steps assume that the packet contains no other headers that must

be located in the packet before the BBR Options Header):

- 1) The node inserts a BBR Options Header after the IP header but before any other header that may be present, as indicated in the Figure 6.14.

Standard IP packet



Standard IP packet with BBR Header

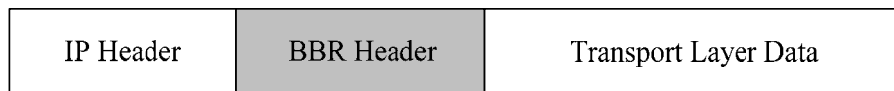


Figure 6.14. BBR Options Header position in a regular IP packet.

- 2) The node sets the Next Header field of the BBR Options Header to the Protocol field of the packet's IP header.
- 3) The node sets the Protocol field of the packet's IP header to the protocol number assigned for BBR (In our implementation, integer 202 is used.).

Originating a Hello Packet. A Hello packet is a control packet generated by BBR routing protocol. Hello packets are generated periodically by each node to discover its current one hop neighbors. A node originating a Hello packet adds a BBR Neighbor Discovery Option to the packet according to the following sequence of steps:

- 1) The node creates a Neighbor Discovery Option, and appends it to the BBR Options Header in the packet.
- 2) The IP Source Address field of the Hello packet is set to the Originate Address of the Neighbor Discovery Option.
- 3) The IP Destination Address is set to the IP limited broadcast address

(255.255.255.255).

- 4) The node broadcasts the Hello packet with the Time To Live field of the IP header set to 1.

Originating a Data Packet. When originating a data packet, the node generates a unique identification number (packet ID), which is a concatenation of the node's ID address (the least significant 16 bits of the IP address for the node) and a locally generated packet ID (16 bits). Then depending on whether the final destination address is in the node's neighbor list or not, the packet will be processed according to the following sequences of steps.

If the final destination address is in the current neighbor list:

- 1) Generate a Data and Neighbor Destination Option. The Option Type and Option Data Length are set according to Option 2. The Reserved field is set to zero. The Forward Count is also initialized to 0. The Originate Address is set to the IP Source Address. The Final Destination Address is set to the IP Destination Address.
- 2) The packet is sent to its destination node by unicast.

If the final destination address is in not the current neighbor list:

- 1) Save a copy of the packet to the node's Forward table (The originating node is selected as the border node by default). Set the Rebroadcast Counter to $(p - 1)$.
- 2) Generate a Data and None Neighbor Destination Option. The Option Type and Option Data Length are set according to Option 3. The Common Neighbor # is set to zero. The Forward Count is also initialized to 0. The Originate Address is set to the IP Source Address. The Final Destination Address is set to the IP

Destination Address. Neighbor Addresses [1... n] are the IP addresses of all the node's current one hop neighbors.

- 3) Set the IP Destination Address in the IP header to the IP limited broadcast address (255.255.255.255).
- 4) The node transmits the packet to all its neighbors by broadcast.

Originating a Data Acknowledgement Packet. A Data acknowledgement packet is a control packet generated by the BBR routing protocol, and it is used to remove the copies of the received packet that might exist in the Forward Tables of intermediate nodes once the data packet has reached its final destination node. Data acknowledgement packets are generated only by destination nodes. A node originating a data acknowledgement packet adds a BBR Data Acknowledgement Option to the packet according to the following of steps:

- 1) The node creates a Data Acknowledgement Option, as described in Option 4, and appends it to the BBR Options Header in the packet.
- 2) The IP Source Address field of the data acknowledgement packet is set to the IP address of the sending node.
- 3) The IP Destination Address is set to the IP limited broadcast address (255.255.255.255).
- 4) The node broadcasts the data acknowledgement with the Time To Live field of the IP header set to 1.

Handling an Unknown BBR Option. Nodes implementing the BBR routing protocol must handle all options specified previously. If a node finds an unknown option, it discards the whole packet.

Processing a Received Packet

When a node receives any packet (whether for forwarding, or as the final destination of the packet), if that packet contains a BBR Options Header, then that node must process the packet according to the option contained in the BBR Options Header.

If the BBR Options Header contains a Neighbor Discovery Option, then it must be processed according to section “*Processing a Received Neighbor Discovery Option*”.

If the BBR Options Header contains a Data and Neighbor Destination Option, then it must be processed according to section “*Processing a Received Data and Neighbor Destination Option*”.

If the BBR Options Header contains a Data and None Neighbor Destination Option, then it must be processed according to section “*Processing a Received Data and None Neighbor Destination Option*”.

Processing a Received Neighbor Discovery Option. When a node receives a Hello packet, it extracts the Source Address of the packet and updates its Neighbor Table by adding a new entry if this is a new neighbor or updates the “Last recorded” field of the entry specific to a neighbor if the neighbor is already in the Neighbor Table. When a node updates its Neighbor Table and finds a new neighbor or neighbors, it initiates a process as described in the Forward Table section. If a node repeatedly fails to receive the Hello packet from a specific neighbor for a number of MaxHelloLoss times, it removes the neighbor from its current neighbor list.

Processing a Received Data and Neighbor Destination Option. When a node receives a data packet contains a Data and Neighbor Destination Option, it examines the packet to

determine if it is the final destination node of that packet. If it is not, it then discards the packet. If it is the final destination node of the received data packet, it removes the BBR Options Header and the included BBR option, and passes the rest of the data packet to the upper layer.

Processing a Received Data and None Neighbor Destination Option. When a node receives a data packet contains a Data and None Neighbor Destination Option, it processes the packet according to the following steps:

- 1) The node checks the Final Destination Address of the Data and None Neighbor Destination Option header, If the final destination node is one of its current neighbors, then it unicasts this packet to that neighbor by modifying the option header to be a Data and Neighbor Destination Option. At the same time it puts a copy of the packet into its Border Node Selection Table, initiating the timer $T_{\max}$ with $T_{\max} = a \times (n \times \Delta t)$, where n is the total number of neighbors on the neighbor list of the received packet. The parameter Δt is the estimated transmission delay for sending one packet, which can be approximated, by (packet length / data speed). a ($a \geq 1$) This parameter is used to increase the value of the timer to make sure that the node receives all the rebroadcast packets that might be coming from the neighbors of the previous forwarding node. When the $T_{\max}$ timer expires, any packets that arrived during the $T_{\max}$ time interval with the same packet ID are removed from the Border Node Selection Table.
- 2) Otherwise, the node checks if this packet is already in its Forward Table, if YES,

it decreases the Rebroadcast Counter of the message with the same packet ID by 1 and then discards the packet. If the Rebroadcast Counter reaches 0, the node removes this entry from the Forward Table.

- 3) If this packet is not in its Forward Table, the node checks the option header. If it is the only neighbor on the neighbor list of the Data and None Neighbor Destination Option, and this node has its own neighbors except the previous forwarding node, then it rebroadcasts the packet with a modified Data and None Neighbor Destination Option attached, saves a copy to the Forward Table and sets the Rebroadcast Counter to be $(p-1)$. If it is the only neighbor and it has no other neighbors except the previously forward node, then it saves a copy to the Forward Table and sets the Rebroadcast Counter to be p .
- 4) If this packet is not in its Forward Table, and there are multiple neighbors in the neighbor list of the Data and None Neighbor Destination Option, then the node saves a copy of the packet and records the packet header information to the Border Node Selection Table, calculates the number of common neighbors between this node and the node that previously sent the packet and marks it as n_c .
- 5) The node initiates another timer T_{ad} to account for access delay. The value of T_{ad} is set to $T_{ad} = (i-1)*\Delta t$, where i is the position of the node address on the neighbor list of the received packet.
- 6) During the T_{ad} period, the node records the header information of each incoming packet with the same packet ID in the Border Node Selection Table

and checks which of its current neighbors are covered by this incoming broadcast, here “covered” means a neighbor node is on the neighbor list of the incoming broadcast packet, which also means the neighbor node has received the broadcast packet.

- 7) When the T_{ad} Timer expires, the node checks whether all of its current neighbors have been covered. If all neighbors have been covered then the node remains silent. If some of its neighbors are not covered, then the node rebroadcasts the packet and modifies the Data and None Neighbor Destination Option by first calculating its Common Neighbor number, then increasing the Forward Counter by 1, and also attaching its list of current neighbors. The node then sets the packet’s IP Source Address to the node’s IP address and the IP Destination Address to the limited broadcast address 255.255.255.255. It keeps recording the packets received during the T_{max} interval.
- 8) When T_{max} expires, the node checks the value of Common Neighbors # fields of all the received packets with the same packet ID, determines the minimum and marks it as n_{min} . The node then determines whether it is a border node in accordance with the following selection rules:
 - a) If $n_c = n_{min}$, then it is a border node according to the border node selection rule. The node saves a copy of the packet in its Forward Table, setting the Rebroadcast Counter to p if no broadcast occurs during the T_{ad} time and to be $(p - 1)$ if it has broadcast once during the T_{ad} period.

- b) If $n_c > n_{\min}$, it is not a border node and discards the packet.
- c) The node then removes all packets with the same packet ID that have been saved in Border Node Selection Table for this border node selection process.

Processing a Received Data Acknowledgement Option. When a node receives a packet contains a Data Acknowledgement option, it processes the packet according to the following steps:

- 1) The node removes the packet whose Identification number (packet ID) is the same as the identification number in the Data Acknowledgement Option of its Forward Table.
- 2) The node decreases the value of the Hop Limit in the Data Acknowledgement Option by 1. If the Hop Limit is 0, then it discards this data acknowledgement packet.
- 3) The node then rebroadcasts the data acknowledgement packet after a random back off time.

Flowchart for Received Packet Processing. The flowchart for processing a received packet is shown in Figure 6.15.

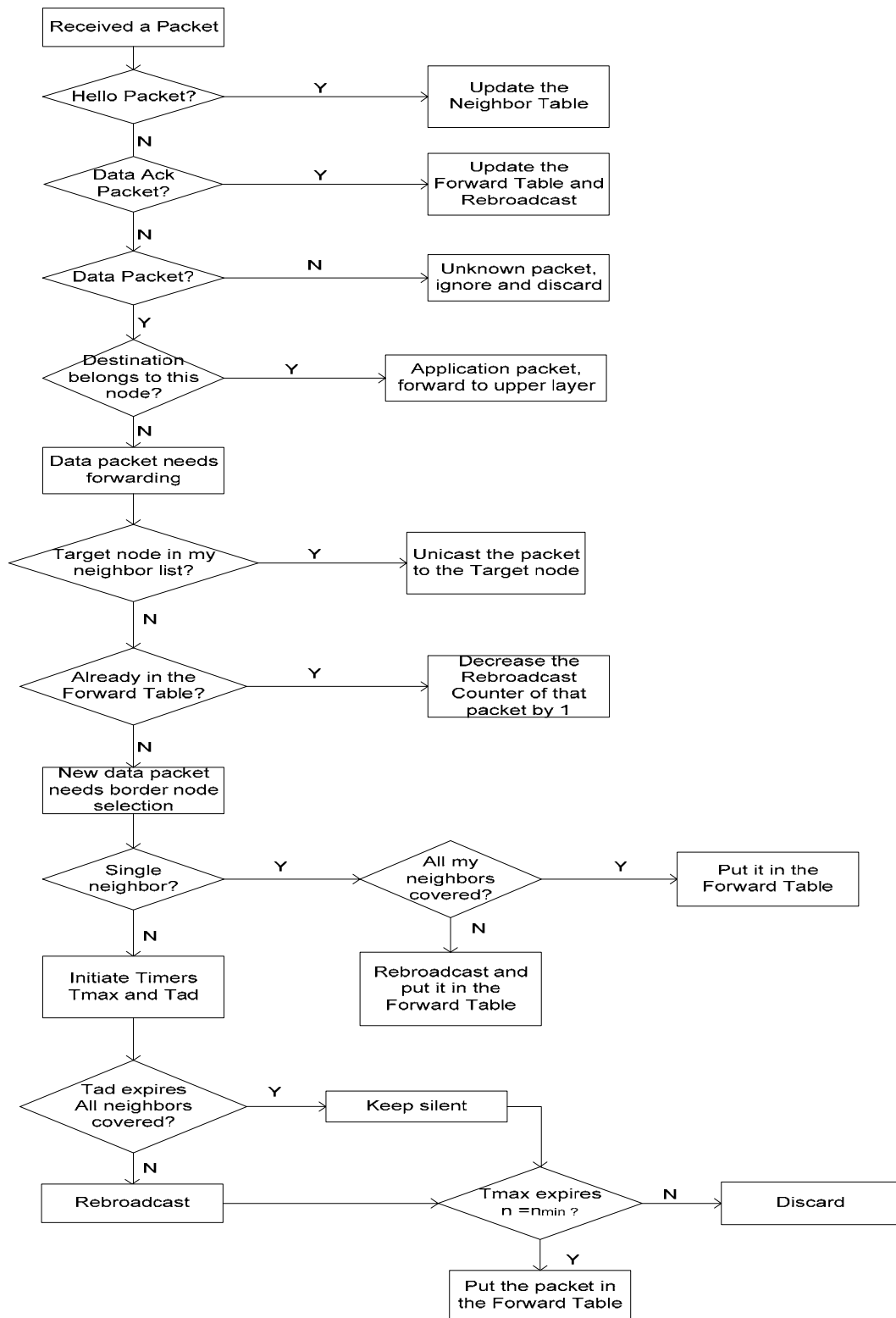


Figure 6.15. Flowchart for received packet processing.

Protocol Constants and Configuration

The BBR implementation supports the following configuration variables and provides a mechanism enabling the value of these variables to be modified by system management. For each configuration variable below, a default value is specified to simplify configuration, as indicated in Table 6.7.

Table 6.7. Some Configurable Parameters and Default Values

HelloInterval	2 seconds
MaxHelloLoss	2 times
MaxRebroadcast(the value of p)	3 attempts
TranDelaySlot	3 milliseconds

Some Features of BBR Routing Protocol

In this section, some features of BBR routing protocol are discussed. These features are added or might be added to the routing protocol to improve the routing performance of the BBR routing algorithm.

Scheduled Forwarding for Mitigating the Hidden Node Problem.

The hidden node problem is inherent to the wireless medium. In a wireless network, a hidden node is defined as a node that is within the interference range of an intended destination but out of the sensing range of a sender. Figure 6.16 gives a node distribution with three nodes A , B and C . The radio transmission range of each node is indicated by a circle. Node A is sending data to node B . Node C , which is out of the radio transmission range of node A senses a “free” medium, and has no idea of any data traffic going on between node A and B . If at this time, node C sends data to node B , there will be collision

at node *B*. In this scenario, node *C* is “hidden” from node *A* and it is called a hidden node of node *A*. Vice versa, node *A* is also a hidden node to node *C*.

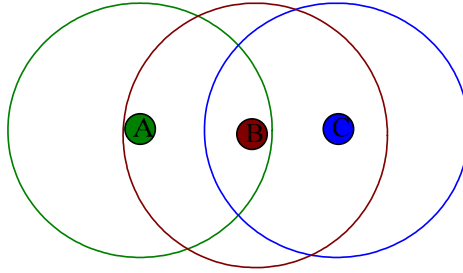


Figure 6.16. Node distribution with possible hidden nodes problem.

In BBR routing, one common scenario where a hidden node problem exists is indicated in Figure 6.17 and illustrated by the following example. Assume that node 1 moves in the upward direction and meets two new neighbors: node 2 and node 5. Assume further that node 1 has multiple packets to send to a specific node. Node 1 broadcasts these packets sequentially and they are received by node 2. Note that nodes 1 and 3 are hidden from each other. Node 3 after it receives a packet from node 2 forwards this packet by either using unicast or broadcast. Since node 1 and 3 are hidden nodes to each other, if the forwarding at node 3 occurs at the same time when node 1 sends a packet to node 2, a collision will occur at node 2. In such case, the probability of having a collision at node 2 is high and some packets from node 1 might be lost.

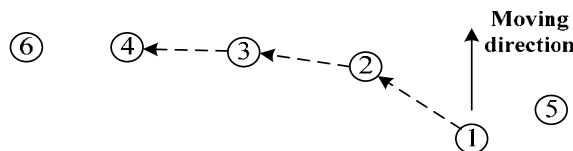


Figure 6.17. Hidden node problem in BBR routing.

In the 802.11x wireless MAC protocol, the hidden node problem is solved by using Request to Send/ Clear to Send (RTS/CTS) control packets to reserve the wireless medium. Regular RTS/CTS frame exchange is used as the protection mechanism for unicast transmissions. However, the RTS/CTS approach can't be used with broadcast frames. So this approach can't be used in the BBR routing algorithm as it relies on broadcast. Note that in the special case where node 1 has only 1 data packet to be forwarded, the hidden node problem will have no effect on that one packet transmission. However, the BBR routing algorithm design is message oriented. One message from a source node to a destination node can consist of several packets. The existence of the hidden node problem in this case results in packet loss during forwarding if no measures are taken.

The hidden node problem can't be solved completely, but one way to mitigate the problem is to use scheduled forwarding. When a sending node wants to forward multiple packets, it sets the packet inter-arrival time to be nonzero. In other words, a time interval is added between two consecutive packet transmissions. The general rule for selecting the time interval is to allow a hidden node sufficient time to forward a packet before the next consecutive packet is transmitted by the source. For example, in the above scenario, the source node (node 1) schedules the times to transmit its packets to allow the hidden node (node 3) to forward packets without incurring collision at the intermediate node (node 2). The necessary time interval is roughly the time from node 1 sending out a packet to node 3 finishing forwarding the same packet and can be approximated by using equation 6.1. In the equation, the packet transmission time is considered and the propagation time is neglected.

$$t_{sh} \approx 3 \cdot \frac{Packet_Length}{Transmission_Rate} \quad (6.1)$$

The time interval t_{sh} is related to data transmission time, which is in turn related to average data packet length and transmission rate of a network. For a typical data packet length of 1024 bits, and with a transmission rate of 2Mbps (In 802.11b, this transmission rate is used for broadcast frames), the value of t_{sh} is about 1.5ms (This value will be even smaller with a higher transmission rate). One minor disadvantage of using scheduled forwarding is that it increases the probability that an intermediate node (node 2 in above scenario) moves out of the radio range of a source node before the source node can transmit all its packets due to the extra added time intervals between packets. This is an issue for high mobility networks. However, assuming a node moves with an average speed of 20m/s, with a radio range of 100 meters, the time that two nodes are within radio range of each other is about 5 s. The added time interval is in the order of milliseconds, and is a negligible side effect of scheduled forwarding and can be ignored.

Redundant Transmission Prevention.

As described previously in this chapter, the neighbor discovery process requires that each node sends out a Hello packet periodically. To make the collision probability of hello packets as low as possible, each node sends out Hello packet independently and asynchronously. Also each node updates its Neighbor Table periodically and asynchronously.

Neighbors discover each other at different time points when they meet each other, which might result in redundant data packet transmission in some cases. Figure 6.18

gives a scenario where two nodes *A* and *B* are moving towards each other. Assume that node *B* has data packets buffered in its Forward Table. When these two nodes move into the radio range of each other, one of the two nodes, such as node *B*, might find node *A* as its new neighbor before node *A* finds node *B* as its new neighbor. When node *B* discovers new neighbor node *A*, it sends out a sequence of data packets destined to a specific destination node. Node *A* receives these packets and saves them in its Forward Table. If there is no other data packets in node *A*'s Forward Table except those packets just received from node *B*, then when node *A* discovers node *B* as a new neighbor, it rebroadcasts those packets just received from node *B* back to node *B*. Node *B*, when receiving these data packets, ignores them if it still has copies of these data packets in its Forward Table or saves these data packets back into its Forward Table again. The retransmission from node *A* is redundant and should be avoided.

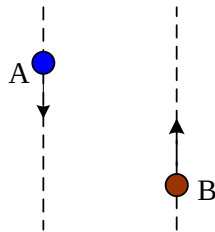


Figure 6.18. Redundant transmission when two nodes meet each other.

One way to solve the redundant transmission problem is to use a Hello message combined with checking the data packet storage time. In the above scenario, when node *B* meets a new neighbor node *A* and wants to forward a sequence of data packets, node *B* sends out an extra Hello message first before it sends out the data packets. Node *A* receives the Hello message from node *B* and records that message to its Neighbor Table.

As above, node *A* later updates its Neighbor Table and finds a new neighbor, node *B*. Before node *A* sends out a sequence of data packets, it checks the storage time of those packets, if it finds that the storage time of those packets is less than one Hello Interval (the Hello_Interval is described below), then it does not forward those packets and hold those packets for a longer time until another new neighbor appears. In this way, the redundant transmission of node *A* can be avoided and those data packets from *B* can be prevented from being unnecessarily transmitted back to node *B*.

Effects of the Hello Interval Length on the Delivery Ratio.

The Hello_Interval is a configurable parameter in the neighbor discovery process and is the length of time between consecutive Hello messages and is also the time interval between two updates of the Neighbor Table. The shorter the Hello_Interval is, the higher the frequency that a node sends out Hello messages and updates its Neighbor Table. A short Hello_Interval helps to maintain accurate neighbor information with the cost of higher control overhead since Hello messages are control messages sent out periodically regardless of whether there is any data traffic.

Increasing the Hello_Interval time can reduce network control overhead. However, a long Hello_Interval time might result in obsolete neighbor information, especially for networks with high node mobility where neighbor information changes rapidly and continuously. Hence choosing an appropriate value for the Hello_Interval is important to maintaining up to date neighbor information while at the same time minimizing the control overhead.

The value of the Hello_Interval affects the accuracy of neighbor information, which in turn has an effect on the packet delivery ratio. Figure 6.19 gives a scenario where two nodes are moving towards each other. The time that the two nodes within the radio range of each other can be expressed as $t_R = \frac{2R}{v_1 + v_2}$, where R is the radio range of each node, v_1 and v_2 are the speeds of node 1 and 2 respectively. The length of the Hello_Interval is noted as t_h . Each node sends out Hello messages and updates its Neighbor Table at fixed time points, indicating by the short arrows along time axis. The time interval between two neighboring update time points is t_h . As Hello messages are sent out by each node asynchronously, the time points at which node 1 and node 2 send out Hello messages and update their Neighbor Tables are not in alignment with each other.

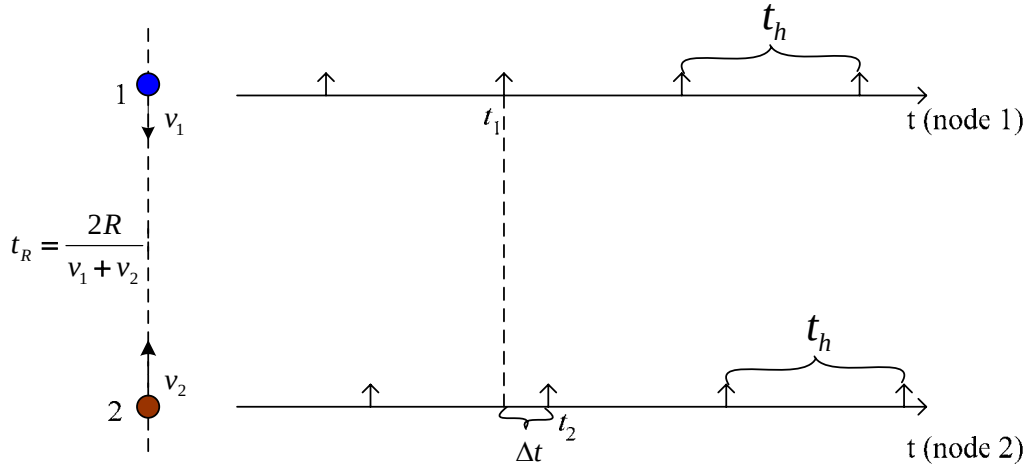


Figure 6.19. Illustration of short Hello_Interval problem.

If a node meets a new neighbor and finds that there are some packets in its Forward Table destined to that neighbor node, it forwards all those packets by using unicast. In Figure 6.19, for an example, node 2 has some packets in its Forward Table that are destined to node 1. When they come close to each other, at time t_1 , node 2 receives a Hello message from node 1 and records that information to its Neighbor Table. At time t_2 , node 2 updates its Neighbor Table and finds a new neighbor node 1. Node 2 checks its Forward Table and finds there are data packets destined to node 1, and then node 2 begins to unicast those data packets to node 1 assuming that node 1 is still its current neighbor. However, if $t_R < \Delta t$, then when node 2 begins to unicast data packets to node 1 at time t_2 , node 1 has already moved out the radio range of node 2, but node 2 is not aware of that. After finishing forwarding, node 2 discards all copies of data packets it has forwarded to node 1, thinking that all packets forwarded to node 1 have reached their destination. However, in fact, node 1 doesn't receive all or even part of the packets forwarded by node 2. The result is that packets are lost during forwarding and those lost packets might never get delivered.

From the Figure it is clear that the maximum value of Δt is t_h . A small value t_R corresponds to a short radio transmission range with high node mobility. The possibility that data packets got lost during forwarding under the above scenario is high when t_R is short and t_h is long. Hence for a sparse network with high node mobility, to improve data delivery ratio, one criterion is to choose a value for t_h that is no larger than t_R . Namely, choose t_h that satisfies $t_h < t_R$.

Selection of t_h that is less than t_R reduces the packet loss without guarantee that packets will be finally delivered. If a long packet sequence needs to be unicast, some packets at the end of sequence may still be lost when nodes move out radio range of each other. If the memory space of individual nodes is big enough then a Data Acknowledgement packet method can be used to assure delivery. In this approach, when a node forwards a unicast packet to its destination, the packet will not be removed from the forwarding node table until an acknowledgement is received for that specific packet. This may result in higher repeat delivery ratio and more Forward Table space will be needed. As discussed before, using a Data Acknowledgement packet in the BBR routing algorithm has some disadvantages, but if high delivery ratio is the most important goal for a network, then it still can be used to improve the data delivery ratio.

Border Node Selection Optimization for a Fully Connected Network.

Any node participating in a specific border node selection procedure has two steps to follow: first, when the access delay timer T_{ad} expires, a node needs to decide whether it should rebroadcast or not, depending on whether all its current neighbors are in range of the broadcast of a source node and/or possible rebroadcasts of neighbors of the source node. Second, when the maximum delay timer T_{max} expires, a node needs to decide independently whether it is a border node or not. These two steps actually accomplish two different tasks. The first step assures dissemination of packets in a locally connected node cluster. The second step is responsible for data packet delivery to other node clusters that are not connected to the node cluster where the source node is located.

The value of $T_{\max}$ is directly related to the total neighbors that a source node has when it broadcasts a data packet. The more neighbors a source node has, the longer $T_{\max}$ is. When a network becomes fully connected, the probability that a source node has a large number of neighbors is high. If a source node has large number of neighbor nodes, then $T_{\max}$ is quite long, which means each node taking part in that specific border node selection procedure needs to wait for a long time before it can make its decision whether to be border node or not.

When $T_{\max}$ is long, before a node can make a decision to be a border node, it may have already moved out the node cluster where it received a broadcast, and then fail to forward data packets to new neighbors. This slows the data packet delivery speed and also causes data packet copies to stay in Border Node Table unnecessarily. When a network becomes fully connected, the task accomplished by the second step is not that important as the whole network now is a single cluster of nodes connected together and delivery of a data packet in such case can be accomplished through step 1.

To improve data packet delivery efficiency in a fully connected ad hoc network, the approach used is to keep the first step as usual while to limit the number of nodes taking part in a border node selection procedure in the second step. By limiting the total number of nodes participating in a specific border node selection procedure, the time for a node to make a decision to be a border node or not is shortened without affecting the total data delivery ratio in a fully connected ad hoc network.

As discussed in Chapter 5, the average neighbor node degree N (neighbors within radio range of each node) can reflect the connectivity characteristic of an ad hoc network.

A threshold value for N that delineates a partially connected network from fully connected network is noted as N_t . It is observed that on average, if N is less than N_t , then statistically the network is partially connected. On the contrary, if N is larger than N_t , the network is statistically fully connected.

In the BBR routing algorithm, a value of N equal or larger than N_t can be used to limit the total number of nodes participating in a border node selection procedure, noted as N_b . When a source node broadcasts a data packet, it will check its neighbor list. When it has a neighbor number larger than N_b , it will randomly select a number of N_b neighbor nodes and put these nodes at the beginning of the neighbor list attached. If it has a neighbor number less than N_b , it will attach the neighbor list in the order the neighbors appear in the Neighbor Table. A node decides whether it should proceed with a border node selection procedure by simply checking its position in the attached neighbor list of the source node. If its position is not larger than N_b , then it should proceed with border node selection. Otherwise, it doesn't need to execute a border node selection for that specific broadcast.

Summary

In this chapter, a BBR routing algorithm has been proposed for sparse MANETs as would be typical in rural areas. The combination of the store-and-forward concept with the border node selection rules makes the BBR routing protocol specifically suitable for a partially connected mobile ad hoc network. At the same time, it is also applicable to fully connected mobile ad hoc networks. The border node selection algorithm has no restrictions regarding the node mobility pattern. It can be used not only on mobile nodes with random movement patterns, but also on mobile nodes constrained to highway movement.

The BBR routing protocol is based mainly on the distributed border node selection algorithm, together with the neighbor discovery algorithm. This chapter gives a detailed description of the data structures used in the BBR routing algorithm, the BBR Options Header format and the detailed operations of the BBR routing protocol. Then some problems inherent in the BBR routing algorithm that may have effects on the routing protocol performance have been discussed and approaches used to solve or alleviate these problems have been proposed. Details of the implementation of the BBR routing protocol using OPNETTM Modeler and the performance evaluation of the BBR routing protocol in partially connected ad hoc networks as well as in fully connected mobile ad hoc networks with different mobility patterns are given in Chapter 7.

References

- [6.1] K. Harras, K. Almeroth and E. Belding-Royer, "Delay Tolerant Mobile Networks (DTMNs): Controlled Flooding Schemes in Sparse Mobile Networks ", *IFIP Networking 2005*
- [6.2] Juang, P., Oki, H., Wang, Y., Martonosi, M., Peh, L., Rubenstein, D., "Energy-Efficient Computing for Wildlife Tracking: Design Tradeoffs and Early Experiences With ZebraNet". *In ASPLOS (2002)*
- [6.3] Wenrui Zhao, M.H. Ammar, Ellen Zegura, "A Message Ferrying Approach for Data Delivery in Sparse Mobile Ad Hoc Networks," *MobiHoc'04*, May 24-26, 2004, Roppongi, Japan
- [6.4] Shah, R., Roy, S., Jain, S., Brunette, W.: Data MULEs: Modeling a Three-Tier Architecture for Sparse Sensor Networks. *In IEEE SNPA Workshop (2003)*
- [6.5] Q. Li and D. Rus, "Communication in disconnected ad hoc networks using message relay," *J. Parallel Distrib. Comput.*, vol. 63, no. 1, pp.75–86, 2003
- [6.6] A. Vahdat and D. Becker, "Epidemic routing for partially connected ad hoc networks," *Duke University, Tech. Rep. CS-200006*, Apr. 2000
- [6.7] Jason LeBrun, Chen-Nee Chuah, Dipak Ghosal, Michael Zhang, "Knowledge Based Opportunistic Forwarding in Vehicular Wireless Ad Hoc Networks," *IEEE VTC Spring 2005*
- [6.8] A. Lindgren, A. Doria, and O. Schelen, "Probabilistic routing in intermittently connected networks," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 7, no. 3, pp. 19–20, 2003
- [6.9] X. Chen and A. L. Murphy, "Enabling disconnected transitive communication in mobile ad hoc network," in *Proc. of Workshop on Principles of Mobile Computing*, collated with PODC'01, Aug. 2001
- [6.10] B. Xu, A. Ouksel, and O. Wolfson, "Opportunistic resource exchange in inter-vehicle ad-hoc networks," presented at the *IEEE Int. Conf. Mobile Data Management*, Berkeley, CA, Jan. 2004
- [6.11] Brad Williams, Tracy Camp, "Comparison of Broadcasting Techniques for Mobile Ad Hoc Networks", *MOBIHOC'02*, June 9-11, 2002, EPFL, Lausanne, Switzerland
- [6.12] Charles E. Perkins, *Ad Hoc Networking*, Addison-Wesley, 2001

- [6.13] S. Ni, Y. Tseng, Y. Chen, and J. Sheu, “The broadcast storm problem in a mobile ad hoc network”, *Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking*, Seattle, Washington, United States
- [6.14] Amir Qayyum, Laurent Viennot, Anis Laouiti, “Multipoint Relaying for Flooding Broadcast Messages in Mobile Wireless Networks”, *Proceedings of the 35th Hawaii International Conference on System Sciences – 2002*
- [6.15] H. Lim and C. Kim. Multicast tree construction and flooding in wireless ad hoc networks. In *Proceedings of the ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWIM)*, 2000

CHAPTER 7

BBR ROUTING PROTOCOL IMPLEMENTATION AND PERFORMANCE EVALUATION

Introduction

The BBR routing algorithm was developed in Chapter 6 and some conceptual data structures were proposed. The detailed operations for a node to run the BBR routing protocol were also specified. This chapter introduces the implementation of the BBR routing algorithm in OPNETTM modeler. Then, the performance of BBR routing protocol under different mobility patterns is comprehensively evaluated and its performance is compared with the DSR routing protocol and also with a Random Node Selection (RNS) algorithm.

BBR Routing Protocol Process Model Development in OPNETTM Modeler

In this section, first MANET protocol modeling in OPNETTM is briefly introduced. Then the BBR routing protocol process model development in OPNETTM is described in detail.

MANET Protocol Modeling in OPNET

OPNETTM is a popular network modeling and discrete time simulation tool. Its wireless module provides a platform to model wireless LAN networks, including mobile ad hoc networks. Its MANET library has a variety of node models to model fixed and mobile nodes, including wireless LAN routers, servers and work stations, and mobile ad

hoc network (MANET) stations. These node models use the 802.11x wireless MAC layer protocol. The framework of these node models provides an open source for adding new MANET routing protocols and customizing existing routing protocols.

Node Model for a Mobile MANET Station. In a pure ad hoc network, there is no central control node and also there are no servers or separate routers. Each MANET node also serves as a router. The node model applied for this study is a mobile MANET station model: `manet_station_adv`. The following introduction about the MANET routing protocol modeling will be based on the `manet_station_adv` node model, which was modeled in OPNETTM with a layered architecture following the OSI model. Figure 7.1 gives the node model architecture of a `manet_station_adv` [7.1]. In the node model, the module `traf_src` serves as an application layer source generating raw packets as well as a sink to accept received packets. The `ip_encap` module encapsulates packets coming from the higher layer into IP packets. It also de-encapsulates IP packets arriving from the lower layer to forward data information to the higher layer. This module serves as an interface between the higher layers and the `ip` module. The `ip` module is responsible for IP routing functions, packet fragmentation and reassembly. The `arp` module implements the Address Resolution Protocol (ARP) which maps IP addresses to physical network addresses to interface to certain Data Link Layer modules. The `wireless_lan_mac` module implements the 802.11x MAC layer protocol. The `wlan_port_rx_0_0` module models a wireless LAN radio receiver, and `wlan_port_tx_0_0` module models a wireless LAN radio transmitter. These two together can model the radio characteristics of a wireless channel.

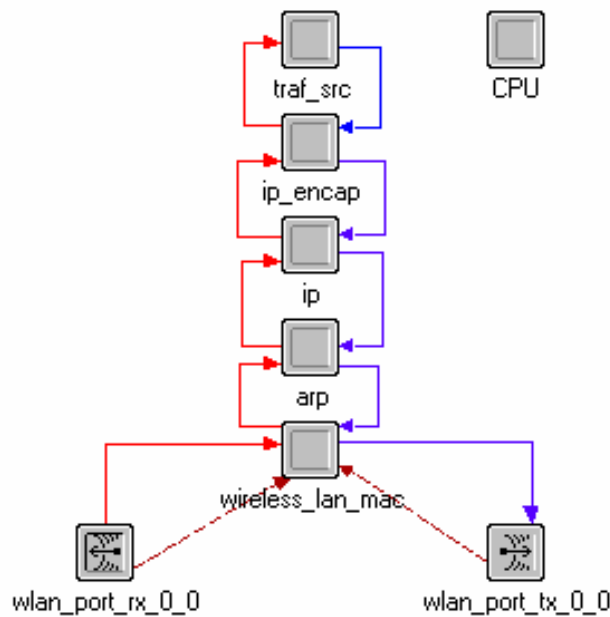


Figure 7.1. Node model architecture for a mobile MANET station node [7.1].

Manet_mgr Architecture. OPNETTM has implemented several popular ad hoc routing protocols, such as DSR, AODV, OLSR and TORA. The DSR, AODV and TORA protocols are implemented over IP using a MANET manager process: manet_mgr. The manet_mgr process provides a common interface to multiple MANET routing protocols and is extensible for custom protocols. The relationship between the manet_mgr process model with the ip module and the MANET routing protocol process models is indicated in Figure 7.2 [7.2]. In the figure, ip_dispatch is the process model for the ip module. Manet_mgr is one of the child processes of the ip_dispatch process model. These process models are the child processes of the manet_mgr process model. The manet_mgr performs the following functions: registers itself with the ip module, determines the MANET protocol configured, creates child processes accordingly and then waits for invocation. The manet_mgr process model can be invoked by routing processes to send

packets to the ip module and also can be invoked by the ip module to send packets to its MANET routing child processes.

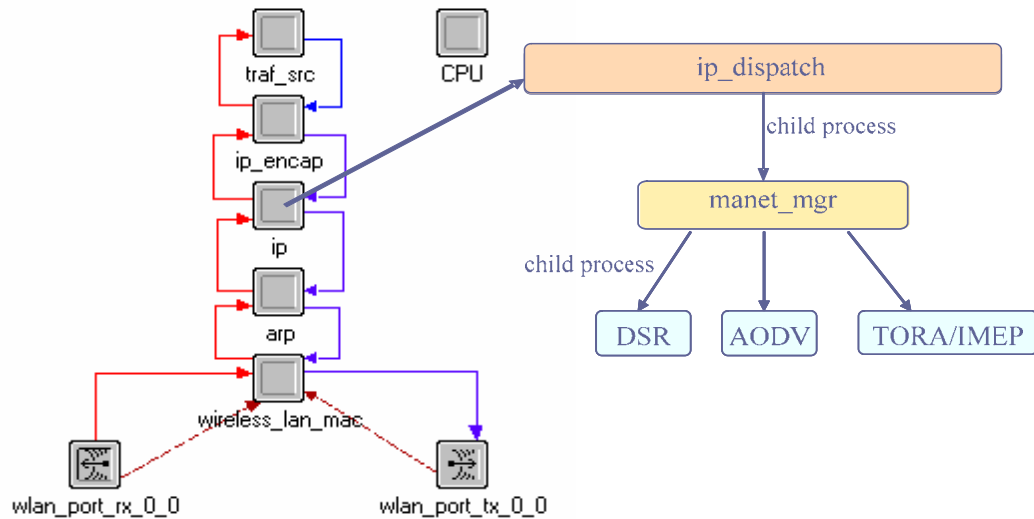


Figure 7.2. Manet_mgr architecture in a MANET station node model [7.2].

Adding a Custom MANET Protocol. A custom MANET protocol can readily be added to OPNET™ if the standard MANET protocol is implemented over IP. The DSR routing protocol implementation in OPNET™ modeler serves as a good reference. A custom MANET protocol can be instantiated as a child process of the manet_mgr process model, as indicated in Figure 7.3. The sequence to add a MANET routing protocol is summarized as follows:

1. Build a process model to implement the protocol core of a custom protocol, writing code for the custom process model including necessary external C files and header files.
2. Register the custom protocol with IP, declaring the protocol name in the IP

program so that IP can recognize the new protocol. The protocol name used in an IP packet is stored as an integer. Currently, in IP, the Internet Protocol, numbers from 138~252 are still available for use, here number 202 is chosen as the protocol name for the BBR routing protocol.

3. Attach the custom process model to the manet_mgr, declare the custom protocol as a child process of the manet_mgr, at the same time declare user configurable protocol attributes in the manet_mgr process interface.
4. Parse the IP interface configuration, adding code into the ip_dispatch process model to parse the custom protocol configuration during IP initialization.
5. Spawn the custom routing protocol child process. Add code to the manet_mgr to spawn the custom process when notified by IP during initialization.
6. Start communication by using the OPNET™ Kernel procedures, Interface Control Information(ICI)s, interrupts for sending and receiving packets

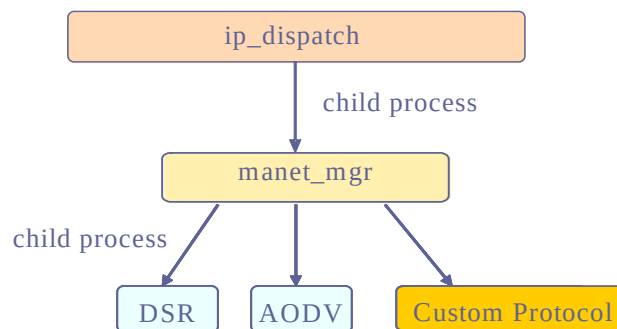


Figure 7.3. Adding a custom routing protocol [7.3].

BBR Routing Protocol Development in OPNET™

BBR Protocol Process Model. Implemented over IP, the BBR routing protocol is claimed as a child process of manet_mgr, as indicated in Figure 7.3 with the Custom

Protocol replaced by BBR Protocol. A process model in OPNET™ is represented by using a finite state machine (FSM), which defines the states of the module and the criteria for changing the states. Figure 7.4 gives the FSM representation of the BBR protocol process model. The “init” state takes care of some of the initialization procedures, such as initializing state variables, registering the statistics, initializing various buffers and tables used by the protocol, and registering the BBR routing protocol as a higher layer in IP and assigning the protocol number to the BBR routing protocol. After initialization, the process immediately makes a transition to the “wait” state. While a process remains at this state, it will periodically send out Hello messages and update the Neighbor Table when a hello timer expires. At the same time it waits for a packet to arrive and processes the received packets according to the rules specified in the BBR routing protocol.

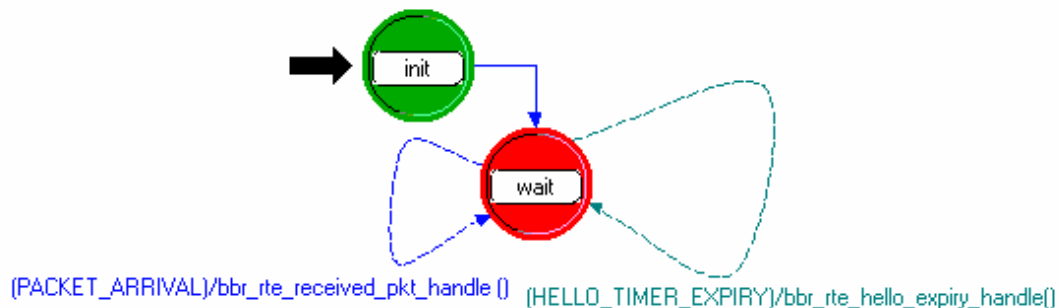


Figure 7.4. Finite state machine representation of the BBR process model.

The main functions of the BBR protocol are coded inside the bbr_rte process model (bbr_rte.pr.m). Some external files (ex.c) and header files (.h) are also coded to support the main functions of the BBR protocol, and include:

- bbr_border_node_selection.ex.c: deals with BBR Border Node Selection Table and handles the border node selection process.

- `bbr_forward_buffer.ex.c`: handles the BBR Forward Table.
- `bbr_neighbor_discovery.ex.c`: handles the BBR Neighbor Table and neighbor discovery process.
- `bbr_notif_log_support.ex.c`: handles simulation log messages.
- `bbr_pkt_support.ex.c`: deals with creation and deletion of BBR packets.
- `bbr_support.ex.c`: supports generic support functions
- `bbr_receive_buffer.ex.c`: handles received packets. The receive buffer is not a necessary component for the BBR routing protocol. It is implemented here only for the purpose of collecting statistics
- `bbr.h`: defines constants, timers and data structures including local and global statistic handles, data tables and buffers that are used by the BBR routing protocol.
- `bbr_pkt_support.h`: defines data structures for the Options packet field in the BBR header, to support normal packet operation.
- `bbr_ptype.h`: includes all prototypes used by the BBR routing protocol.

BBR Protocol Packet Format. The BBR packet format is shown in Figure 7.5. In the BBR packet format, length values of the “Options” field and the “data” field are variables and are unknown until a packet is generated. After a packet is generated, the “Options” field contains one of the BBR options and the length value of the “Options” field is set according to the BBR option specified. The “data” field generally contains a higher layer data packet, and the length of the “data” field is set to the real length of the packet.

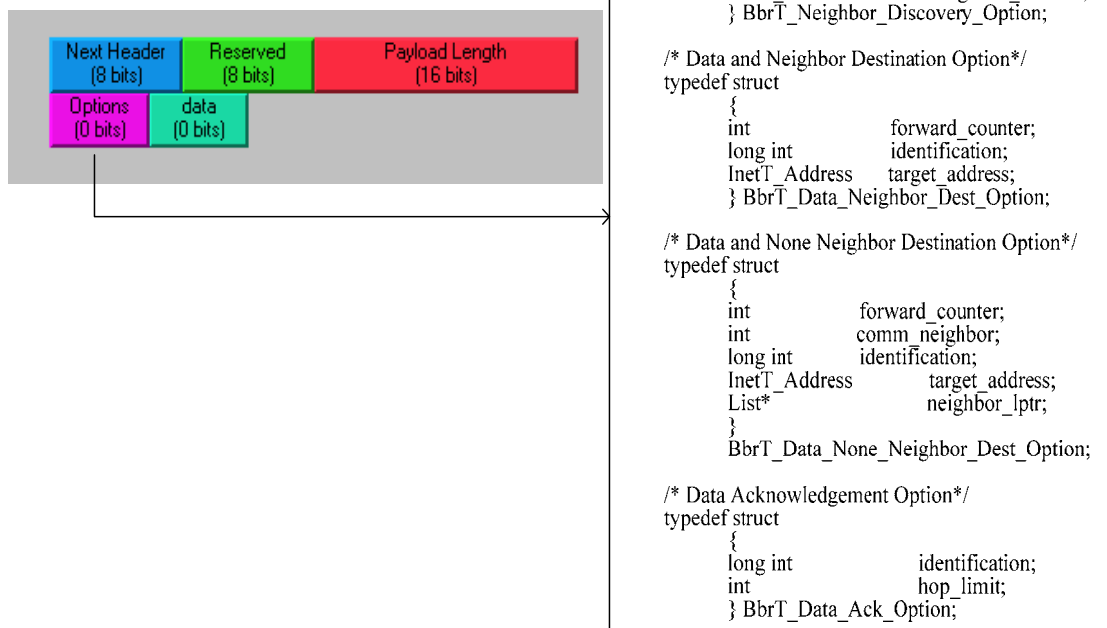


Figure 7.5. BBR packet format.

BBR Protocol Packet Flow. The BBR protocol uses its own routing algorithm and it does not use the forwarding table of IP. There are two types of packets: data packets and control packets. Specifically, the BBR control packet includes Hello packets and data acknowledgement packets, others are data packets. Each packet (data/control) is appended with a BBR header and goes through the BBR protocol module. For IP, a data packet appears the same as a control packet since there is no routing function needed from IP. However, a packet goes through different paths in the IP module depending on where a packet comes from and whether it is a data packet or a control packet.

1. Control traffic packet flow. Control traffic packet flow in BBR is exactly the same as that of DSR routing protocol [7.4]. Control packets in BBR use one hop limited broadcast. Figure 7.6 gives the control traffic packet flow when a control packet is generated by the MANET routing protocol. Figure 7.7 shows the control traffic flow when a control packet is received from a lower layer. These two figures are redrawn here for completeness.

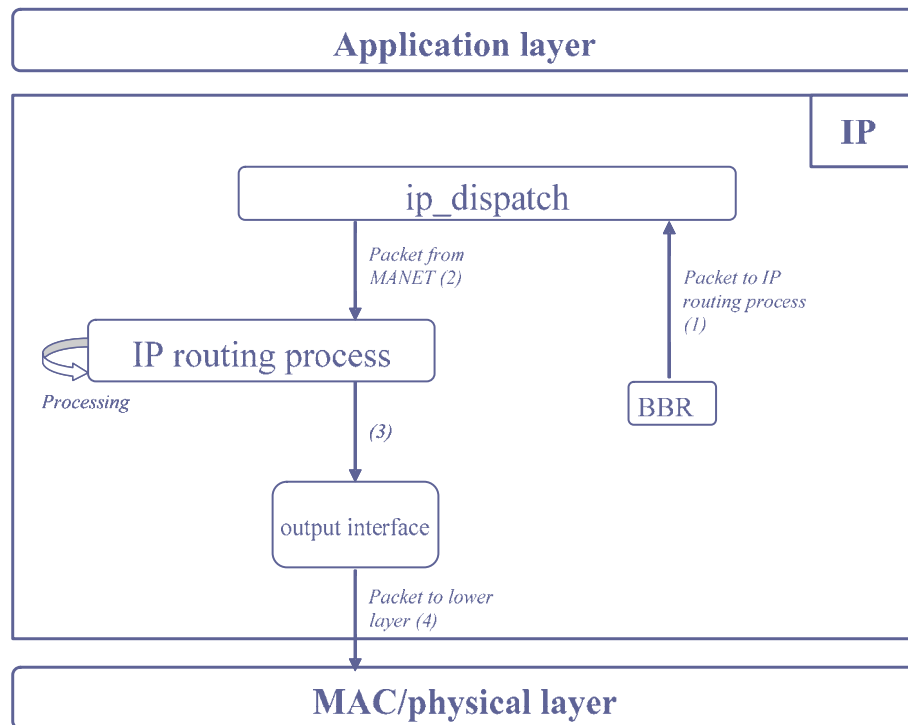


Figure 7.6. Control packet from routing protocol [7.4].

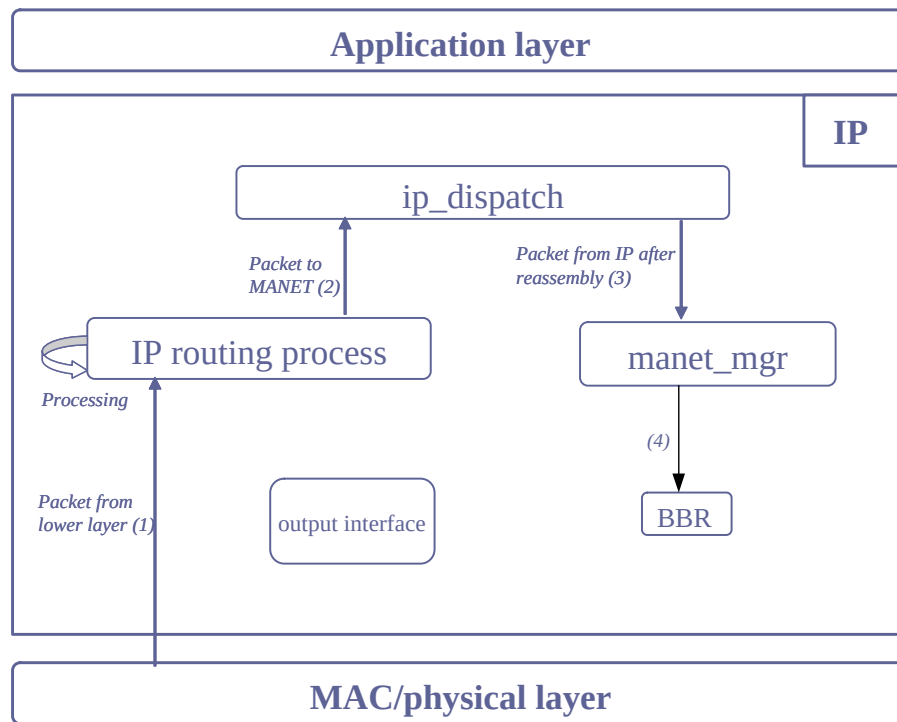
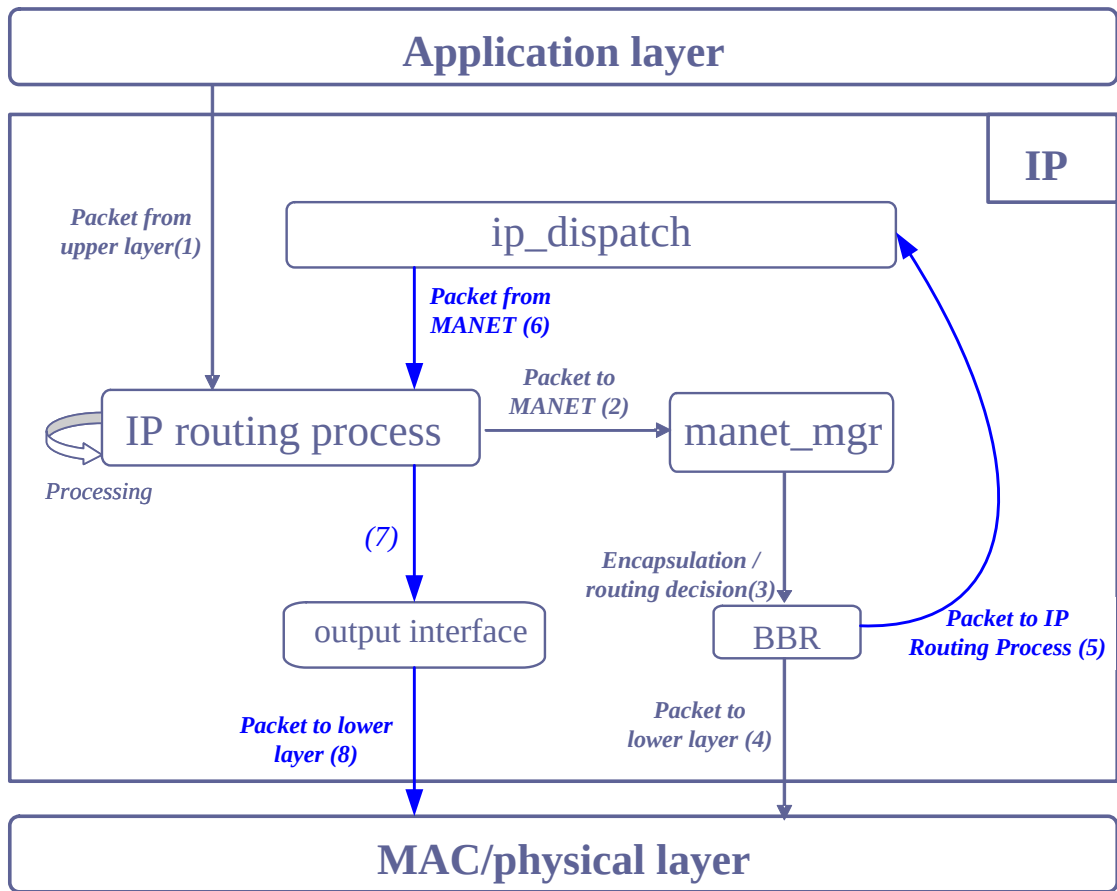


Figure 7.7. Control packet from lower layer [7.4].

2. Data packet from the application layer. When a data packet comes from the application layer, BBR will process it differently depending on whether the final destination node is on a node's current neighbor list or not. Figure 7.8 gives the packet flow when a data packet comes from the application layer.



Legends

Route 1-4: packet from upper layer, and neighbor is the destination node of the packet.

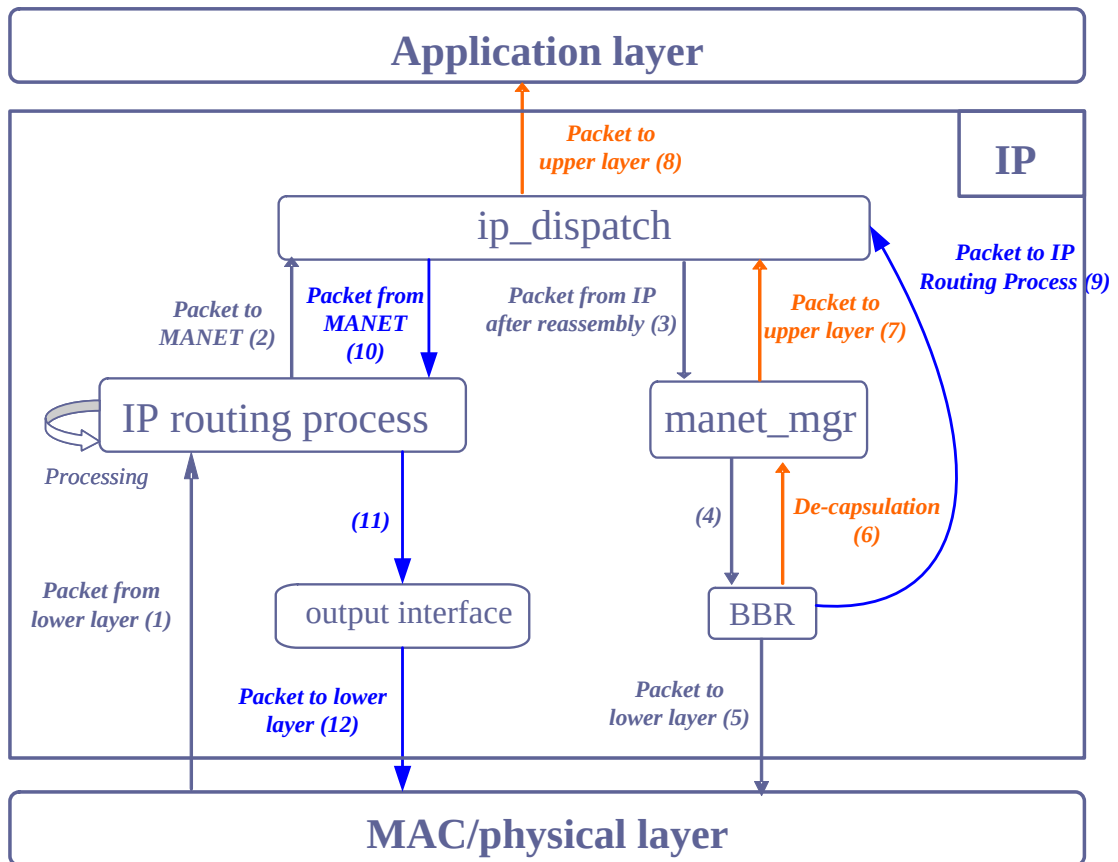
Route 1-8: packet from upper layer, and neighbor is not the destination node of the packet.

Figure 7.8. Data packet flow when a data packet is from application layer.

When a data packet comes from the application layer, if the IP routing process finds that the node is configured with the MANET routing protocol, then the IP routing process will send the packet to the manet_mgr process model without looking up its forwarding table. The manet_mgr process will send the packet to one of its child process models according to the MANET protocol configured. When the packet arrives at the BBR protocol model, the BBR protocol model will encapsulate the packet and make decisions about routing and forwarding. According to BBR routing algorithm, if the final

destination node is on the neighbor list of the current node, then the node will forward the packet to the destination by using unicast, which is indicated in the figure as Route 1-4. If the destination node is not in the neighbor list of the current node, then the node will forward the packet by using one hop limited broadcast after processing the packet according to BBR routing algorithm, as indicated by Route 1-8. For this route, a packet goes through steps 5 to 8 the same way as a control packet when a control packet is generated by the routing protocol.

3. Data packet from a lower layer. When a data packet comes from a lower layer, the packet will be processed differently, depending on whether the receiving node is the final destination node or not. The packet flow will also be different. Figure 7.9 gives the packet flow diagram when a packet is received from a lower layer. When a data packet from a lower layer reaches the BBR process model, the BBR process model will forward the packet to the upper application layer if this node is the final destination node of the packet, as indicated by Route 1~4-6~8. Otherwise, this node will check whether the destination node is on the current neighbor list, and if it is, then the packet will be sent to the lower layer directly using unicast to forward the packet to a neighbor destination node as indicated by Route 1~4-5. If the neighbor node is not on the current neighbor list, then the data packet will be sent to the ip_dispatch process model after the BBR header is modified according to specified BBR forwarding procedures. Then the ip_dispatch process model will send the packet to the lower layer by using one hop limited broadcast.



Legends

Route 1~4-5: packet to lower layer, and neighbor is the destination node of the packet.

Route 1~4-6~8: packet destined for this node.

Route 1~4-9~12: packet to lower layer, and neighbor is not a destination node of the packet.

Figure 7.9. Data packet flow when a data packet is from a lower layer.

Performance Evaluation

The OPNETTM process model implements the basic BBR routing protocol process along with added features described in the last part of Chapter 6, including scheduled forwarding, redundant transmission prevention and border node selection optimization. The process model is integrated with the existing mobile MANET station node model to simulate a mobile node in an ad hoc network. To evaluate the performance of BBR

routing protocol comprehensively, the performance is observed and measured under a variety of conditions. To make comparisons, the performance of the DSR routing protocol is also evaluated under the same network configurations. The reason DSR is chosen as the basis for comparison is that DSR is a very popular ad hoc routing protocol for mobile ad hoc networks and it has better and more consistent performance at all mobility rates than other standard routing protocols [7.5], such as AODV, TORA and DSDV. In the following, the general behavior of the BBR routing protocol is first evaluated with a random node mobility model applied. Then the performance of BBR is evaluated with the GTI Mobility Model described in Chapter 5. The performance of BBR and DSR are compared with each other as each mobility model is applied. Finally, the performance of BBR is also compared with a Random Node Selection algorithm.

General Simulation Scenario

This scenario is designed to evaluate some general behaviors of the BBR routing protocol. Simulations are designed to enable sensitivity analysis of some of the configurable parameters of BBR and to evaluate the routing protocol performance under different network connectivity characteristics and mobility.

The simulation area is a two-dimensional space of 1000×1000 square meters. The number of nodes in the network is assumed to be 50. These nodes are initially uniformly placed within the simulation area. For the initial stationary distribution, the average distance among neighboring nodes which is noted as L_{av} , can be approximated by assuming these nodes are completely evenly distributed in the simulation area. Then

$$L_{av} = \frac{(1 + \sqrt{2})}{2} \cdot \sqrt{\frac{(1000 \times 1000)}{50}} = 171.4 \text{ meters.}$$

A parameter α is defined as the ratio between the radio transmission range R and L_{av} , namely $\alpha = \frac{R}{L_{av}}$. This parameter characterizes the degree of network connectivity. When $\alpha < 1$, the nodes are on average separated by more than the radio range, and the network is disconnected. For $\alpha > 1$, the average node separation is less than the radio range, and the network becomes gradually connected.

The Random Waypoint Mobility Model [7.6] with pause time equal to zero is used to model the movement of each mobile node in the network.

The MAC protocol used is IEEE 802.11b with a 2 Mbps data rate. The data traffic is defined as follows: Shortly after the simulation begins, each node randomly chooses a destination node other than itself and sends out data packets for 20 seconds. The data packet inter-arrival time is uniformly distributed with an average of 2 seconds. The packet size has an exponential distribution with an average of 1024 bits.

For the BBR protocol configurable parameters, values indicated in Chapter 6 are generally used. In some simulations, some of the parameters are varied analyze the performance sensitivity of parameter value selection. Table 7.1 summarizes the basic parameter values used in the simulations.

Table 7.1. Simulation Parameters

Network simulator	OPNET™ Modeler	
Simulation area	1000m × 1000m	
Number of nodes	50	
Mobility model	Random waypoint	
Node speed	Uniform(0, 20) m/s	
MAC protocol	IEEE 802.11b	
Data rate	2Mbps	
Data traffic	Packet inter-arrival time	Uniform(1,3) s
	Packet size	Exponential Average:1024bits
BBR configurable parameters	HelloInterval	2 s
	MaxHelloLoss	2 times
	MaxRebroadcast	3 attempts
	TranDelaySlot	3 ms

For all simulations, the results are averaged over 15 runs with different random number seeds. The results are illustrated in the figures by the mean and error bars indicating the standard deviation. Several metrics measured to evaluate the protocol performance are defined as follows:

- Packet delivery ratio: the ratio between the number of non-repeat packets delivered to the destination and the number of packets sent by the source.
- Delay: the average end-to-end transmission delay by taking into account only the packets non-repeatedly received.

Packet Delivery Ratio vs. Simulation Time. This simulation tests the BBR routing protocol performance as a function of simulation time. In this simulation, the radio

transmission range R is specifically chosen to be 8 meters to make the network very sparsely connected. With this radio range, α is 0.047. The percentage of the study area covered by a single node is 0.02%. The simulation time is varied from 1000 seconds to 190,000 seconds. The performance metric measured is the packet delivery ratio. Figure 7.10 plots the packet delivery ratio as a function of the simulation time and the vertical bars on the data points are the root mean square values for the fifteen runs. The figure indicates that when the simulation time is short, the packet delivery ratio is low. The packet delivery ratio increases rapidly with the increase of the simulation time until it reaches about 75% with the simulation time to be about 10,000 seconds. After that, further increase the simulation time from 10,000 seconds to 19,000 seconds, the packet delivery ratio remains at a relatively constant level. This is because in a sparse network, packet delivery mainly depends upon the movement of mobile nodes to carry forward data packets. In short time period, the possibility that a forwarding node or a source node meet a destination node is relatively low, resulting in the low delivery ratio. Increasing the simulation time improves the possibility that a forwarding node or a source node will meet a destination node, which results in the increased packet delivery ratio. However, once the simulation time increases to the point that is long enough for all packets in the network to be delivered, further increase in the simulation time does not help to further improve the delivery ratio.

From the figure, it is also noted that the final delivery ratio does not reach 100%. This is because packets get lost during forwarding due either to the hidden node problem or because the neighbor information is not being updated frequently enough. The hidden node problem was discussed in detail at the end of Chapter 6. The loss due to the update

frequency is related to the length of the HelloInterval, and the effects of the HelloInterval on the packet delivery ratio will be studied in simulations later.

Note that in these simulations, the nodes initiate packet transmissions for only the first 20 seconds. If the nodes sent packets for the entire simulation time rather than just for the first 20 seconds, then the additional traffic would result in a lower delivery ratio as those packets generated at the end of simulation time might not have enough time to be delivered. This is especially true when the network is highly partitioned and the delivery of packets is mainly dependent upon the movement of the mobile nodes themselves. We discuss the delivery delay in a following section.

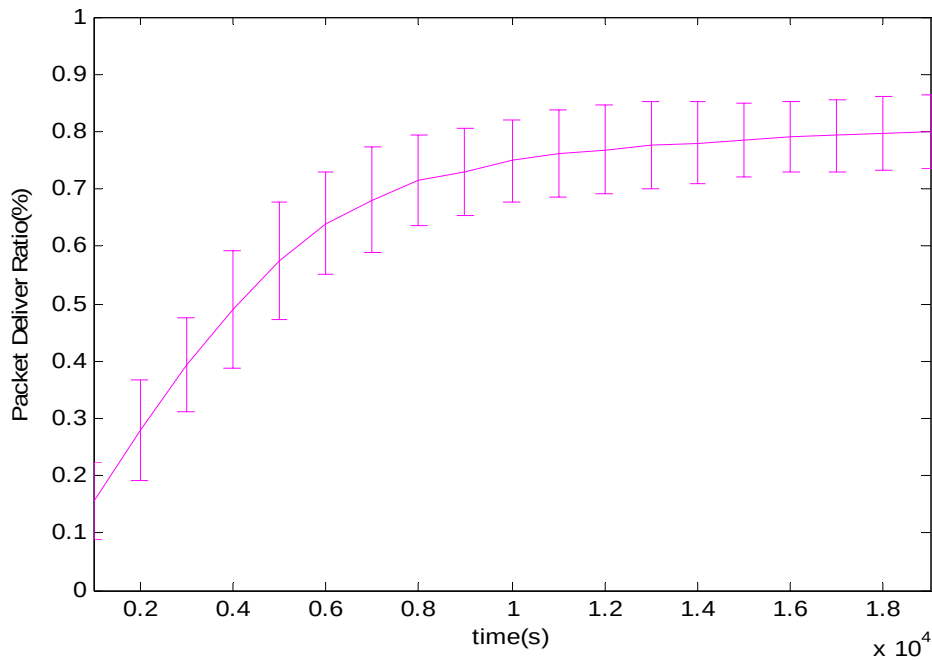


Figure 7.10. Packet delivery ratio vs. simulation time.

Packet Delivery Ratio vs. MaxRebroadcast Number. The MaxRebroadcast number is one of the BBR configurable parameters, and it determines the maximum number of

times a border node will rebroadcast a data packet. To test the sensitivity of the routing protocol performance to this parameter, two sets of simulations were done. The radio transmission range R in one simulation was set is 8 meters, which is corresponds to $\alpha = 0.047$. In the other simulation set, R was chosen be 80 meters corresponding to $\alpha = 0.47$. The simulation time for each run was 20000 seconds. Figure 7.11 shows the relation between the packet delivery ratio and the MaxRebroadcast number at different values of α .

These results indicate that the packet delivery ratio is not sensitive to this parameter when $\alpha = 0.047$. The packet delivery ratio remains constant when the MaxRebroadcast number changes from 1 to 8. However, when $\alpha = 0.47$, the packet delivery ratio increases significantly when the MaxRebroadcast number increases from 1 to 2. Then it remains constant as the MaxRebroadcast number increases from 2 to 8.

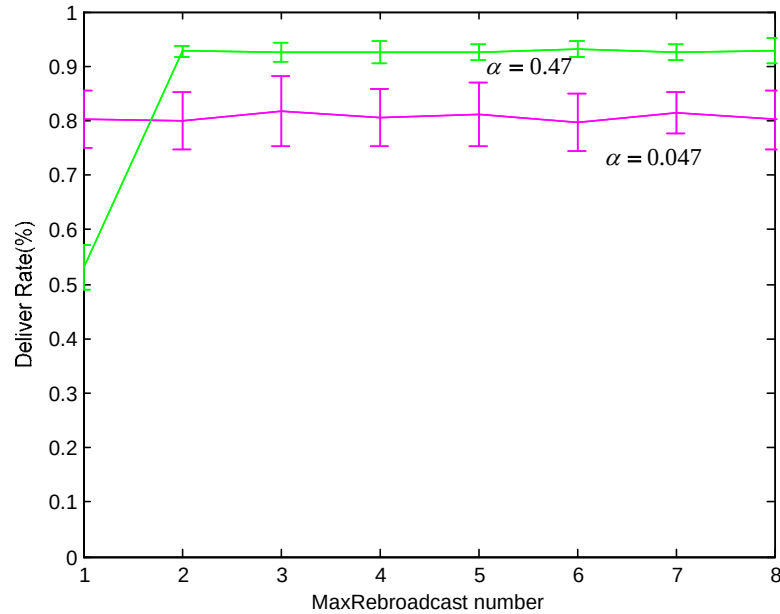


Figure 7.11. Packet delivery ratio vs. the MaxRebroadcast number.

The sensitivity of the packet delivery ratio to the MaxRebroadcast parameter at different values of α can be explained as follows. When $\alpha = 0.047$, the network is highly partitioned. During most of the simulation time, a node typically has only one neighbor. Under these conditions, the border node selection is simplified. The only neighbor will always be selected as a border node and carry the packet forward to other nodes by movement. However, when $\alpha = 0.47$, during most of simulation time, a node will have more than one neighbor. For each broadcast, a border node process will be carried out as described in Chapter 6. As defined in the BBR protocol, if a node rebroadcasts once when the access delay timer T_{ad} expires, then when the maximum delay timer T_{max} expires and the node makes a decision to be a border node, it will save the packet in its Forward Table and rebroadcast $p - 1$ (p is the MaxRebroadcast number) times later. Under these conditions, if the MaxRebroadcast number p is equal to 1, then those border nodes who have broadcast once when T_{ad} expires actually will not carry data forward for further delivery, resulting in some packets being lost during the border node selection process. This explains why there is a big jump in the packet delivery ratio when MaxRebroadcast increases from 1 to 2.

When MaxRebroadcast is larger than 2, the packet delivery ratio is not sensitive to the value of MaxRebroadcast. The reason behind this is that a node has no memory of packets it has forwarded. A node removes a packet from its Forward Table after rebroadcasting that packet MaxRebroadcast times, but later, it may receive the same packet from other nodes and save that packet in its Forward Table again for another round of rebroadcasting. A small value for MaxRebroadcast results in a packet being quickly

removed from its Forward Table. Subsequently, since that packet is no longer in its Forward Table, the possibility that this node receives the same packet from other nodes and saves that packet for another round of rebroadcast is high.

Packet Delivery Ratio vs. HelloInterval. Two sets of simulations were carried out to test the effects of HelloInterval on the routing protocol performance. The simulation time for each run was 20,000 seconds. The parameter HelloInterval was varied from 0.2 seconds to 4 seconds. The radio transmission range R in one simulation set was 8 meters, corresponding to $\alpha = 0.047$. In the other simulation set, R was 80 meters, corresponding to $\alpha = 0.47$. The simulation results are shown in Figure 7.12.

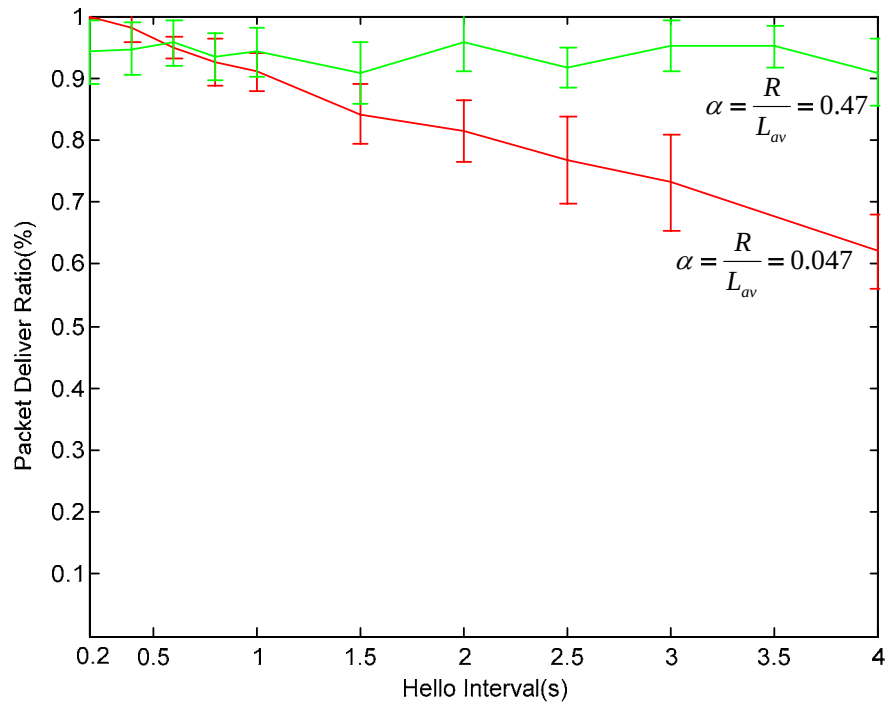


Figure 7.12. Packet delivery ratio vs. HelloInterval.

The figure shows that when HelloInterval increases from 0.2s to 4s, the packet

delivery ratio decreases when $\alpha=0.047$ and does not change much when $\alpha=0.47$. These simulation results verify the observation in Chapter 6 regarding the effects of the HelloInterval on routing protocol performance. The node speed in the simulation is uniform(0, 20) with an average of 10m/s, and the average time that two nodes are within radio range of each other can be approximated by $t_R = \frac{2R}{v_1 + v_2}$, which is 0.8 sec. when $R=8$ meters and 8 sec. when $R=80$ meters. As discussed in Chapter 6, when HelloInterval is larger than t_R , the possibility that packets will be lost due to neighbor information not being updated frequently enough is high, as showed in the simulations when $\alpha=0.047$. Increase of the HelloInterval to values larger than t_R decreases the packet delivery ratio significantly. However, as long as the HelloInterval is less than t_R , increasing the HelloInterval does not have much effect on the packet delivery ratio, as indicated in the simulations when $\alpha=0.47$ and $t_R=8$ seconds. The packet delivery ratio is not sensitive to the HelloInterval when it increases from 0.2s to 4s. Hence, for a specific network using the BBR routing protocol with known value of t_R , the value chosen for HelloInterval should be as big as possible but less than t_R to reduce the hello message control overhead while at the same time maintaining a relatively high packet delivery ratio.

Packet Delivery Ratio vs. Radio Transmission Range. In this simulation, the routing protocol performance as a function of radio range was evaluated. The simulation time for each run was 5000 seconds. The HelloInterval was fixed to be 2 seconds. The radio range was varied from 8 meters to 800 meters. Figure 7.13 shows the packet delivery ratio

relative to the radio transmission range. The vertical dotted line in the figure indicates the point where the radio transmission range equals the average neighbor separation distance L_{av} . Within the 5000 second simulation time, when radio transmission range increases from a small value, the packet delivery ratio initially increases rapidly. After radio transmission range reaches about 80 meters and thereafter, the packet delivery ratio remains constant at about 90%. When the radio transmission range reaches about 800 meters, the packet delivery ratio rises to about 99%.

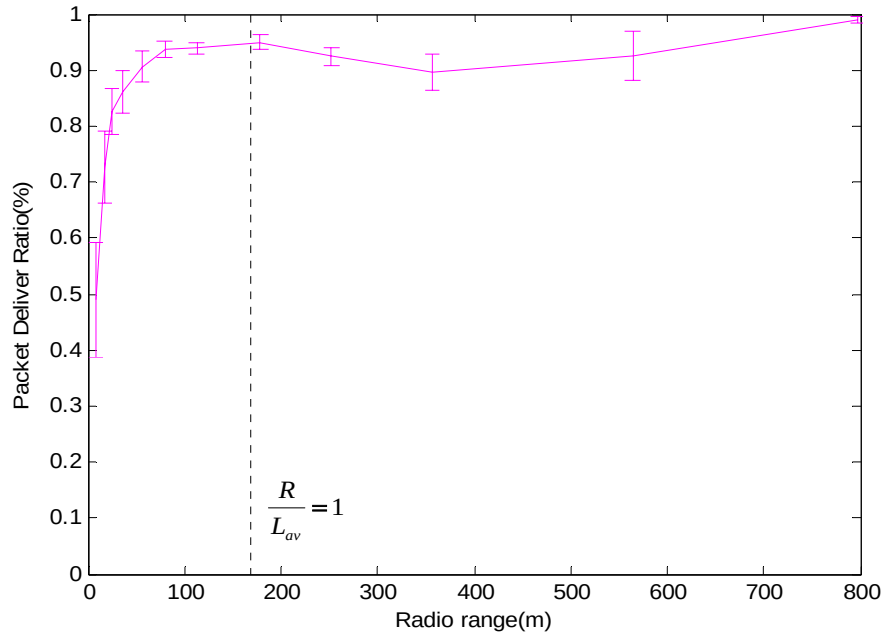


Figure 7.13. Packet delivery ratio vs. radio transmission range.

The low delivery rates at small radio transmission ranges are mainly due to the value selected for HelloInterval, which is fixed to be 2 seconds for all the simulations, and is not small enough to meet the criterion: HelloInterval should be less than t_R . Once the radio range increases to be large enough to meet that criterion, the packet delivery ratio

can achieve a relative high percentage of about 90% in the 5000 seconds simulation time. The conclusion that can be made from the simulation is that for a specific network, if an appropriate value for HelloInterval is chosen, then the BBR routing protocol can achieve a relatively high percentage delivery rate even when the radio transmission range is small, that is, when the network is sparsely or partially connected.

Average Packet Delivery Delay vs. Radio Transmission Range. In the previous simulations, which test the packet delivery ratio under a variety of radio transmission ranges, another performance metric, the average delay per delivered packet was also evaluated. Figure 7.14 shows the average delay per delivered packet relative to the radio transmission range. The average delay decreases rapidly when transmission range increases. When the radio transmission range is larger than the average neighbor distance, the delay is very short.

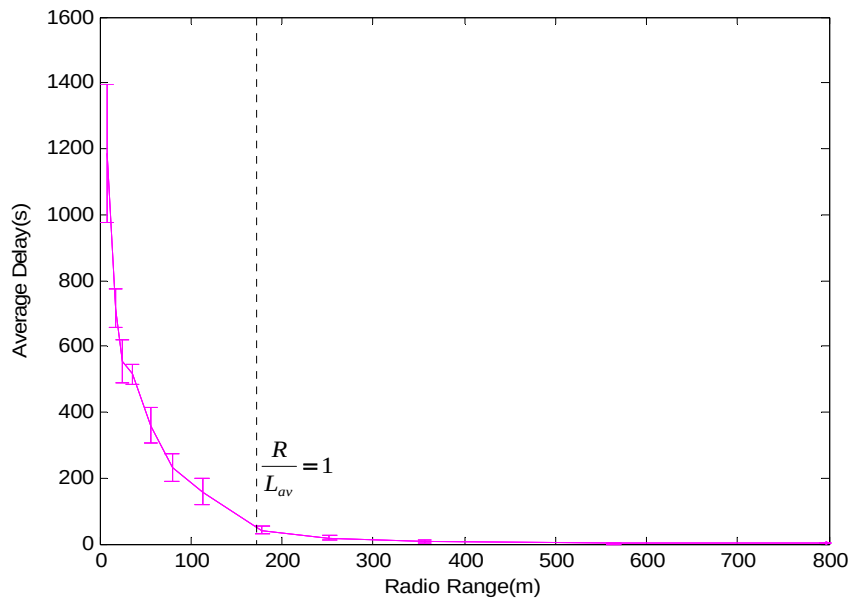


Figure 7.14. Average delay per delivered packet vs. radio transmission range.

The long delivery delay at small radio ranges is expected due to the fact that the network is highly partitioned at those radio ranges. The delivery of data packets under this condition is mainly dependent upon the movement of mobile nodes to carry forward packets instead of using wireless communication among nodes. Once the radio range increases to be larger than the average neighbor separation distance, the network gradually becomes connected. With a more connected network, packets can be delivered mainly through wireless communication among nodes, which significantly shortens the delay time. When the radio range is so large that the network becomes fully connected, delivery of packet can be accomplished within seconds since now the source nodes and the destination nodes are always connected.

From simulation results indicated in Figure 7.13 and Figure 7.14, we can conclude that using the BBR routing protocol, a relatively high and constant packet delivery ratio can be achieved for a fully connected ad hoc network as well as for a partially connected ad hoc network. However, a high packet delivery ratio is achieved with a much longer packet delivery delay when the network is partially connected or highly partitioned.

Comparisons between BBR Routing Protocol and DSR Routing Protocol. This section makes a performance comparison between the BBR and DSR routing protocols. Performances of both protocols are evaluated using the same network configurations. To test the effects of node mobility on protocol performance, each routing protocol is simulated with varying node mobility. The three groups of node speeds applied are low, middle and high node mobility, where the speed of each node in the network is distributed uniformly with ranges (0, 2) m/s, (10, 20) m/s and (25, 35) m/s respectively. A

detailed comparison of these two protocols under different node mobility conditions is given in Figure 7.15 for the delivery rate and Figure 7.16 for average packet delivery delay. In Figure 7.15, the radio transmission range is on a logarithmic x-axis.

1. Packet delivery ratio comparison.

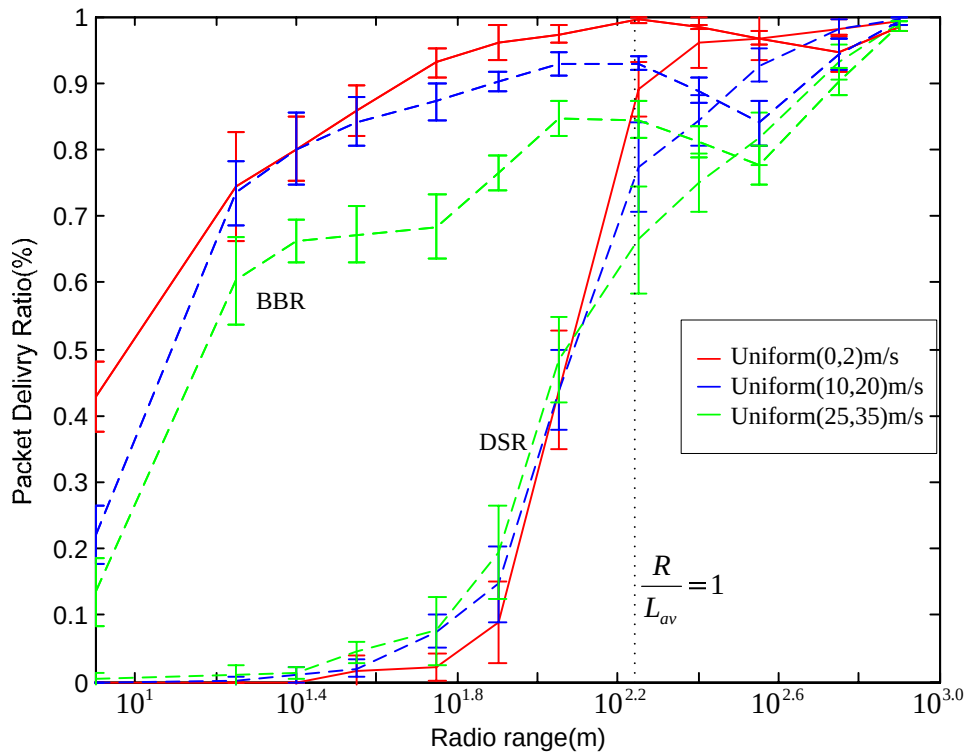


Figure 7.15. Packet delivery ratio comparison between BBR and DSR protocols.

Figure 7.15 shows that when the radio range is very small and the network is highly partitioned, the packet delivery ratio with DSR is close to 0 percent. The packet delivery ratios with BBR are also low at very small radio ranges. However, as discussed above, the packet delivery ratio of BBR is sensitive to the HelloInterval selected. For all simulations above, the HelloInterval is 2 seconds. If instead, a much smaller HelloInterval is chosen for small radio transmission range, the packet delivery ratio can

be substantially improved. Hence, the BBR routing protocol can have much better performance than DSR when a network is highly partitioned. This result is expected since packet delivery using DSR is based on the discovery of connected route from source node to destination node at a specific time, which has very low probability when the network is highly partitioned. While for BBR, packet delivery is based on “carry and forward”. As long as enough neighbors can be met during the simulation time, the probability that packets get delivered is much higher.

As the radio range increases, the packet delivery ratio for both routing protocols increases. However, the packet delivery ratio of BBR reaches a relatively high value well before the point where the network becomes fully connected while for DSR, the packet delivery ratio remains low until the network becomes more fully connected. The BBR routing protocol yields a much higher packet delivery ratio than that of DSR until the radio transmission range is lightly larger than the average distance among neighboring nodes, as indicated by the vertical dotted line in the figure. After this point the packet delivery ratio of DSR is slightly greater than that of BBR. Finally, the packet delivery ratios of both protocols statistically converge to the same point, which is almost 100% when the network is fully connected.

Figure 7.15 also shows that for BBR, the packet delivery ratio is higher at low node speed than at high speed for the whole radio transmission range. In other words, in terms of packet delivery ratio, BBR performs better at low node mobility. The reason behind this is that for BBR, packet forwarding sometimes is not packet by packet. Instead, when a forwarding node meets a destination node and there is a sequence of packets destined to that node in its Forward Table, the forwarding node will forward all that

sequence at one time. High node mobility increases the possibility of packets getting lost in such situations, which is the main reason that packets can't be delivered in BBR routing protocol. For DSR routing, when the radio transmission range is small, node mobility actually helps to slightly improve the packet delivery ratio, as indicated in the figure. At radio ranges less than about 100 meters, the packet delivery ratio with high node mobility is higher than that of low node mobility. However, as the radio range further increases to larger than the average neighbor separation distance and thereafter, DSR has better performance at low node mobility in terms of packet delivery ratio. The reason high node mobility improves the packet delivery ratio at small radio ranges for DSR routing can be explained as follows. When the radio range is small and network is highly partitioned, the possibility that a source node will find a connected path to the destination node is very low if the destination node does not happen to be in the neighbor range of the source node. In DSR, packet delivery is always packet by packet. Within these radio ranges, high node mobility increases the possibility that a source node meets a destination and results in a higher packet delivery ratio. On the other hand, as the radio range increases and the network becomes fully connected, high node mobility will cause an existing connected route to be easily broken, making packets get lost during forwarding, resulting in a lower packet delivery ratio.

It is also noted from the figure that when the radio transmission range increases to that the point where almost all nodes are neighbors of each other, both BBR and DSR can deliver a packet using only one hop. Then the effects of node mobility on packet delivery ratio can be neglected. In fact, at this point, each protocol reaches an almost 100% delivery ratio.

2. Average delay per delivered packet comparison.

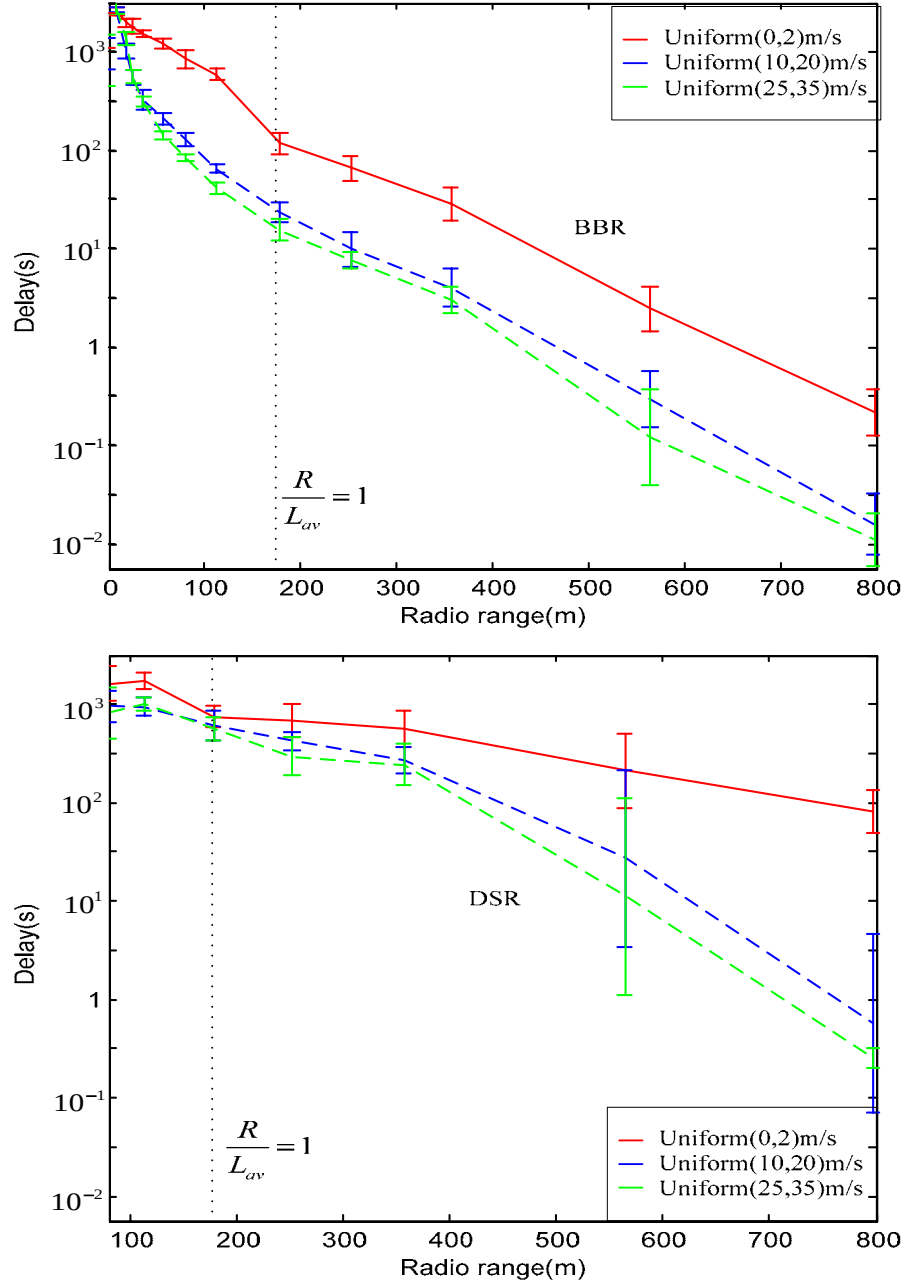


Figure 7.16. Average packet delivery delay comparison between BBR and DSR protocol.

The simulation results for average packet delay using BBR and DSR are shown in Figure 7.16. We note that with DSR, for short radio ranges, the packet delivery

probability goes to zero and the delay becomes infinite. Hence to make useful comparisons the radio range on the x-axis begins at 80 meters. For both protocols, the average packet delivery delay drops rapidly as the radio transmission range increases and node mobility also helps to decrease the packet delivery delay. The high delivery delay for BBR at low radio ranges is due to the network being highly partitioned. The delivery of packets is mainly dependent upon node movement. This also explains why the average packet delivery delay is shorter when node mobility is higher. The low delivery delay for BBR at high radio ranges is the result of the network becoming gradually connected, and packet delivery is more dependent upon wireless communications among neighboring nodes instead of node movement. For DSR, the high delivery delay is due to the low probability of finding an end to end path when the network is highly partitioned. Packets must be queued in the send buffer for long time intervals before the route discovery procedure is successfully completed. High mobility increases the likelihood that a source node gets close to a destination node in a shorter time period, which helps to decrease the average packet delivery delay.

Figure 7.16 also indicates that BBR has a shorter packet delivery delay than DSR for the same radio transmission range and the same node mobility. The delay difference between BBR and DSR is smaller for a fully connected network as compared to a highly partitioned network. The main reason for this is that for a fully connected network, DSR needs extra time to complete the route discovery process before sending data packets while BBR does not rely on such a process. However, for a highly partitioned network, BBR generally can deliver packets albeit with long delay, while for DSR, there is high likelihood that there is no end to end route available within the entire simulation time,

resulting in an infinite packet delivery time.

Several conclusions can be drawn from simulation results shown in Figure 7.15 and Figure 7.16. For both routing protocols, high node mobility helps to reduce packet delivery delay but decreases the packet delivery ratio. For BBR, the packet delivery ratio is relatively constant over radio ranges considered, but the packet delivery delay is much longer when the radio range is small corresponding to a highly partitioned network. For DSR, the packet delivery ratio remains low until the radio range increases sufficiently to make the network connected, which shows that DSR is not appropriate for partially connected networks. When a network is fully connected, DSR outperforms BBR slightly in terms of packet delivery ratio, but with a little bit longer packet delivery delay.

The simulation results indicate that when a network is fully connected, the performance of BBR is comparable to that of DSR in terms of packet delivery ratio and average packet delivery delay. However, BBR has much higher control overhead than DSR. Main control overhead for BBR is related to the neighbor discovery process. This process is executed regardless whether there is any data traffic or not. DSR is a reactive routing protocol. The main control overhead is related to route discovery, which only occurs when there are data packets that need to be sent. When there is no data traffic, the control overhead is very low. Therefore, the DSR routing protocol is preferable to the BBR routing protocol when network is fully connected while the BBR routing protocol is preferable when the network is partially connected.

Highway Scenario

This scenario is designed to evaluate the performance of the BBR routing protocol in

vehicular ad hoc networks, where mobile nodes movement is restricted to roads. The GTI Mobility Model, which is developed in Chapter 5, is used for the study of BBR in this scenario. The simulation area chosen for this scenario is based on the highway system inside Yellowstone National Park. To make a comparison of packet delivery delay with the actual time for a mobile node to move from one specific location to another, the data traffic is specified as follows. Shortly after the simulation begins, a source node located at the West Thumb point within Yellowstone National Park sends out data packets to a destination node located at the west entrance of the park. Data traffic continues for 40 seconds. The packet inter-arrival time is exponentially distributed with an average of 1 second. The packet size is exponentially distributed with an average packet size of 1024 bits. Table 7.2 gives a summary of the basic simulation parameters and values.

Table 7.2. Simulation Parameters for Highway Scenario

Network simulator	OPNET™ Modeler	
Simulation area	Yellowstone Highway System	
Number of nodes	150 (L_{av} =2085 meters)	
Mobility model	GTI mobility model	
Node speed	Uniform(18, 22) m/s	
Simulation time	1.5 hrs	
MAC protocol	IEEE 802.11b	
Data rate	2Mbps	
Data traffic	Packet inter-arrival time	Exponential(1) s
	Packet size	Exponential(1024) bits
BBR configurable parameters	HelloInterval	2 s
	MaxHelloLoss	2 times
	MaxRebroadcast	3 attempts
	TranDelaySlot	3 ms

Figure 7.17 shows a comparison of the packet delivery ratio with BBR and DSR, with the GTI mobility model applied. The results indicate that BBR yields a very high delivery rate that is close to 100% for all radio transmission ranges. For DSR, the delivery rate is close to 0 when the ratio $\alpha = \frac{R}{L_{av}}$ is less than 1. Once α is larger than 1, the delivery ratio increases rapidly to a relatively high point over 90%. When the radio range is sufficiently large so that the whole network becomes fully connected, both protocols exhibit a 100% packet delivery ratio. By comparing Figure 7.17 with Figure 7.15, we note that the simulation results indicate that DSR has similar performance when both mobility models are applied.

For BBR, comparison of Figure 7.17 with Figure 7.13 shows the dependence of the packet delivery ratio on mobility model. The results indicate that with the GTI mobility model, the packet delivery ratio of BBR is more consistent at different radio transmission ranges and has a higher delivery ratio when the radio range is low. The reason for this behavior is that with the GTI mobility model, the node movement has much less randomness as compared to the Random Waypoint Mobility model. The possibility that a source node meet a destination node is higher in the GTI mobility model, resulting in a higher packet delivery ratio.

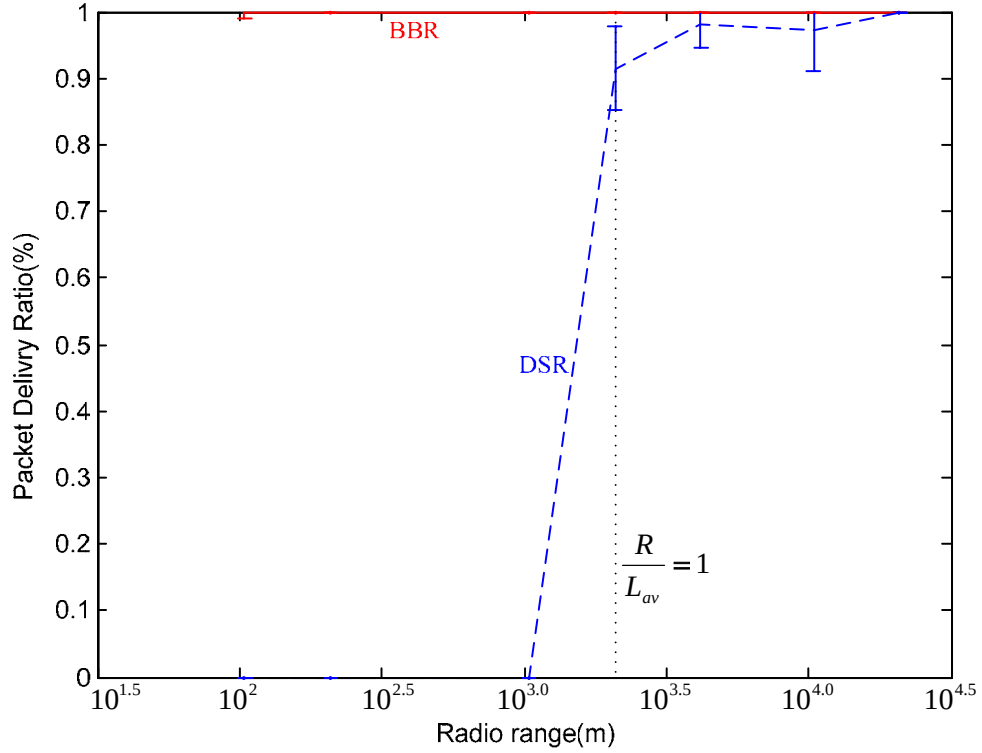


Figure 7.17. Packet delivery ratio comparison between BBR and DSR protocol.

Figure 7.18 gives a comparison of the packet delivery delay for BBR and DSR when the GTI mobility model applied. For DSR, once the network becomes connected, delivery delay is short. With DSR, the packet delivery is always accomplished by wireless communication among nodes composed of a complete path from source node to destination node. As long as a complete route can be found, the packet delivery is finished within a minute.

For BBR, when the radio range is small and the network is partially connected, before α reaches 1, the packet delivery delay remains relatively long, (about 3700 seconds for the scenario modeled). This time period is very close to the time interval for a node to move from the source point to the destination point, which is about 3300 seconds

with an average node speed of 20m/s. This simulation result is also in agreement with the simulation results that were obtained in Chapter 5, which also shows that when α is less than 1, the packet transit time is about the same as the time for a node to move from a source point to the destination, though for that simulation the average movement speed is about 13m/s, with a transit time of about 5000 seconds. In Chapter 5, when α is slightly greater than 1, the transit time jumps to 0 due to the ideal routing protocol applied. In Figure 7.18, when α larger than 1, the packet delivery delay decreases gradually, as now the delivery of packets depends partly on wireless communication among mobile nodes and partly on node movement. When the network is fully connected, the packet delay is also in the range of minutes since the packet delivery at these radio ranges is fully due to wireless communication among mobile nodes.

The similarity in performance indicated by the simulation results obtained with the BBR routing protocol and the ideal routing protocol when the network is partially connected also further validates the GTI mobility model.

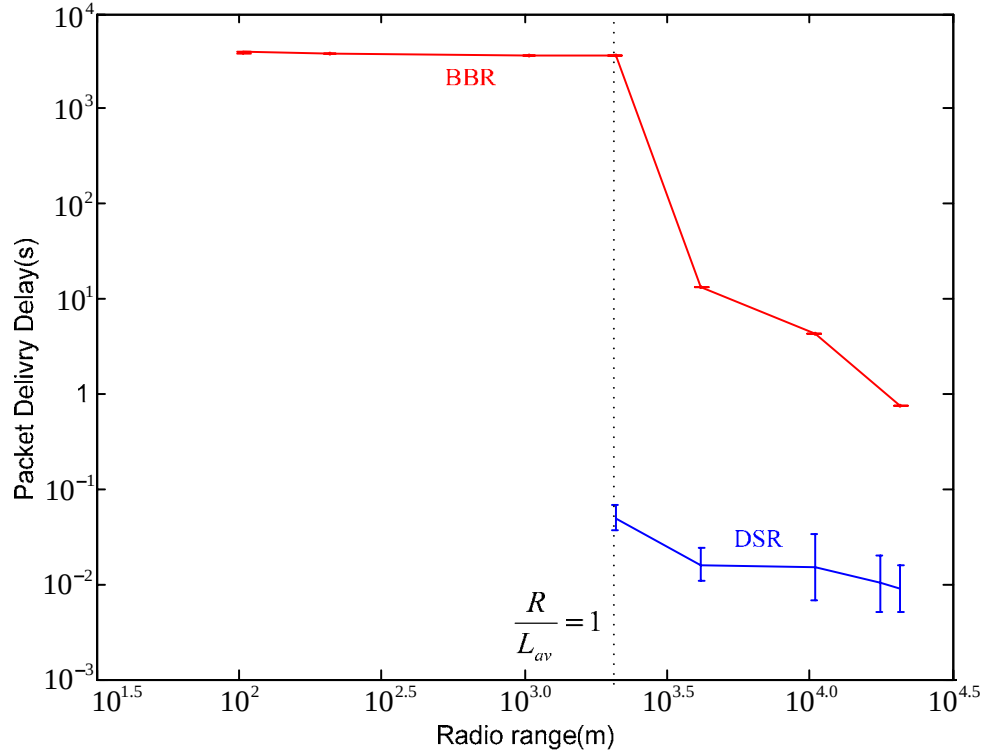


Figure 7.18. Packet delivery delay comparison between BBR and DSR protocol.

BBR vs. Random Node Selection (RNS)

To further verify the performance of the BBR routing protocol, the performance of BBR is compared with a Random Node Selection (RNS) algorithm. The RNS algorithm is similar to the BBR routing algorithm, the difference is that the least common neighbor based border node selection process in BBR is replaced by a simple random node selection in RNS. In brief, in BBR, when a source node broadcasts a packet, each neighbor makes a border node selection to decide whether it is a border node or not. A border node is a forwarding node in BBR. However, using the RNS algorithm, the forwarding node is chosen by random. Namely, when a source node broadcasts a packet, it randomly chooses a node from its current neighbor list to be a forwarding node. The

source node puts the forwarding node designation information in its broadcast packet. A neighbor node receiving such a broadcast packet reads that information and knows if it is a forwarding node or not. Compared with BBR, the RNS algorithm is simpler as it eliminates the procedures associated with border node selection that are executed in BBR. To compare the performance of BBR with RNS, simulations were carried out based on the general parameters and simulation environment summarized in Table 7.1.

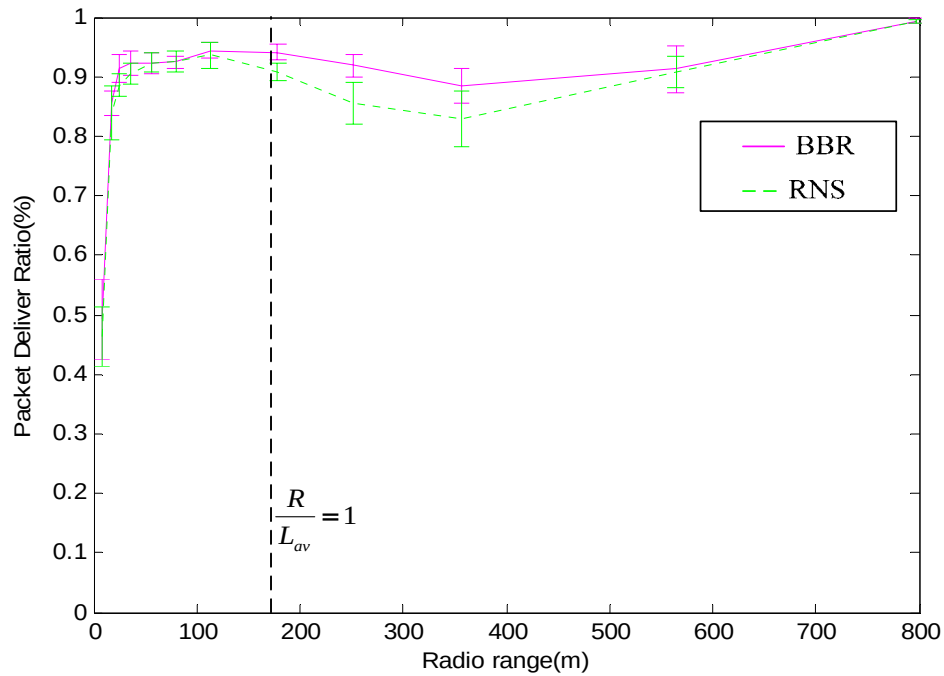


Figure 7.19. Packet delivery ratio comparison between BBR and RNS.

Figure 7.19 gives the packet delivery ratio comparison between BBR and RNS. Figure 7.20 gives the packet delivery delay comparison between BBR and RNS. The figures show that when the network is highly partitioned (when radio range is less than 100 meters) or fully connected (radio range is larger than 600 meters), BBR and RNS

have very similar performance based on packet delivery ratio and packet delivery delay. When the network is partially connected, BBR has a higher packet delivery ratio than RNS and outperforms RNS by about 7% when the radio range is 300 meters. BBR also outperforms RNS in packet delivery delay slightly at that radio range.

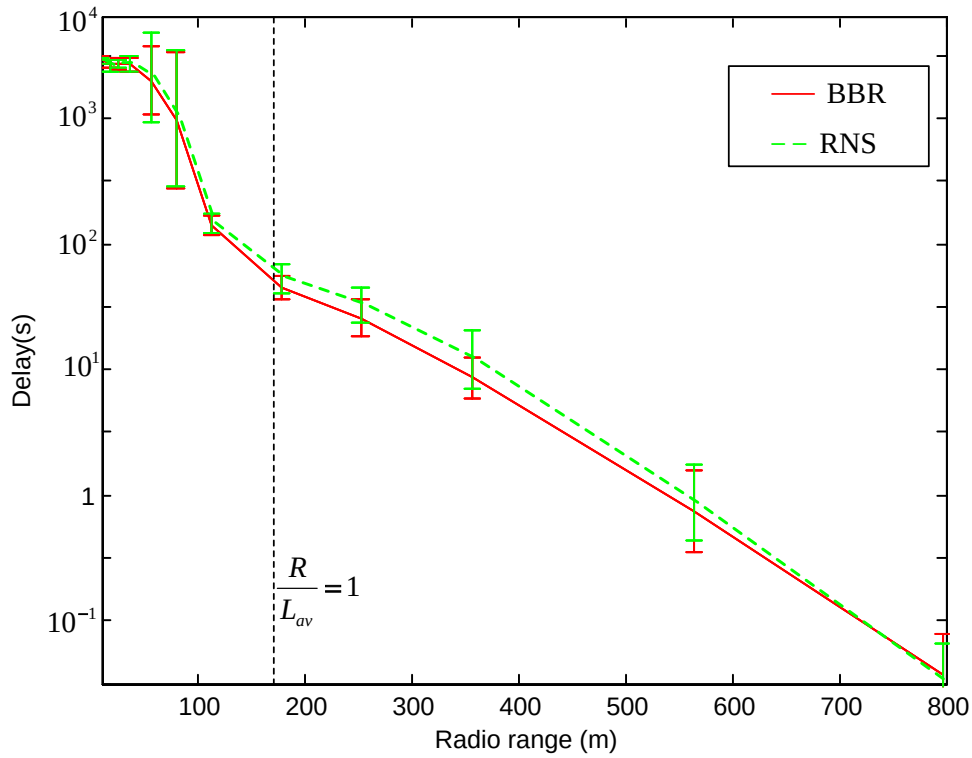


Figure 7.20. Packet delivery delay comparison between BBR and RNS.

The performance comparison between BBR and RNS was also made for the GTI mobility model applied to the YNP simulation scenario. Similar simulation results as those in Figure 7.19 and 7.20 are observed.

The simulation results are in agreement with the observation that when the network is highly partitioned, both algorithms deliver packets dependent on the node, and when network is fully connected, both algorithms deliver packets through wireless

communication. When the network is partially connected, both algorithms depend on the selected forwarding nodes to carry packets. The BBR algorithm can find forwarding nodes more efficiently than the random algorithm in terms of physical location as well as completeness. With BBR, multiple border nodes located in different directions relative to the source broadcast node could be selected as border nodes at the same time. While using RNS, there is only one forwarding node, which is randomly chosen. This explains why BBR has higher packet delivery ratio and short packet delivery delay when the network is partially connected.

Summary

This chapter gives the BBR routing protocol process model developed for use in OPNETTM Modeler. Then using OPNETTM, the protocol performance has been comprehensively evaluated. In the general simulation scenario, different simulations have been carried out to study the effects of the configurable parameters on performance with the Random Waypoint Mobility models applied. The performance of the BBR routing protocol has been compared with the DSR routing protocol with both the Random Waypoint Mobility and GTI mobility models applied. The simulation results with both mobility models indicate that BBR yields better performance than DSR when the network is partially connected and has comparable performance to DSR when the network is fully connected. With the general simulation scenario, the performance of BBR has been also compared with the Random Node Selection (RNS) algorithm, and the simulation results indicate that BBR has better performance than RNS when network is partially connected.

References

- [7.1] OPNET™ Modeler 10.5, Education version, Node Model: manet_station_adv.
- [7.2] OPNET™ Technologies, Inc, “Understanding MANET Model Internals, Interfaces, and Performance”, Session 1552, *OPNETWORK2005*, pp9.
- [7.3] OPNET™ Technologies, Inc, “Understanding MANET Model Internals, Interfaces, and Performance”, Session 1552, *OPNETWORK2005*, pp63.
- [7.4] OPNET™ Technologies, Inc, “Understanding MANET Model Internals, Interfaces, and Performance”, Session 1552, *OPNETWORK2005*, pp29-32.
- [7.5] Josh Broch, David A. Maltz, David B. Johnson, Yih-Chun Hu, and Jorjeta Jetcheva, "A Performance Comparison of Multi-Hop Wireless Ad Hoc Routing Protocols," *Proceedings of the Fourth Annual ACM/IEEE International Conference on Mobile Computing and Networking*, pp. 85--97, October 1998.
- [7.6] David B. Johnson and David A. Maltz, Dynamic source routing in ad hoc wireless networks. *Mobile Computing*, edited by Tomasz Imielinski and Hank Korth, chapter 5, pages 153–181. Kluwer Academic Publishers, 1996.

CHAPTER 8

CONCLUSIONS

This dissertation explores the potential offered by emerging wireless techniques and develops new networking approaches for fixed and mobile high-speed communications services suitable for rural areas. One part of the dissertation focuses on the development of network models based on the existing and emerging technologies that are suitable for fixed wireless broadband access for rural and sparse areas. Another focus of the dissertation is to develop a routing protocol suitable for very sparse networks where end-to-end connectivity may be intermittent.

The conclusions of the research work reported in this dissertation are summarized below, ending with some remarks about possible future work.

(1) With recent advances in electronics and networking, wireless communications shows promise to play a more important role in meeting the ever increasing demands for broadband access services in rural and sparse areas.

(2) A baseline network model for broadband fixed wireless access for rural areas was proposed and investigated. The simulation results shows that a Wi-Fi based broadband Internet access network is technically and financially viable in a rural area. Selection of proper technology and architecture alternatives can improve the cost effectiveness and enable favorable financial performance.

(3) Multi-hop models for WLANs were investigated and their capacity limitations analyzed. Study results show that for a given node density and average per node offered load, appropriate multi-hop models can be selected to meet the prescribed requirements.

The study shows that there are domains of user node density and offered load where multi-hop strategies provide a more cost effective approach to broadband wireless Internet access than conventional point-to-multipoint architectures.

(4) There are key similarities and differences between a fixed wireless multi-hop and a mobile ad hoc network that can be exploited for communications in sparse areas. In a mobile ad hoc network, an appropriate routing protocol is essential for successful data communication. Conventional MANET routing protocols are generally designed for fully connected networks. The unique characteristics of rural and sparse areas require a new routing protocol for effective use of MANET approaches. An appropriate mobility model is necessary for accurate routing protocol performance evaluation.

(5) A geographic and traffic information based mobility model (GTI mobility model) was developed for modeling the movements of mobile nodes in a vehicular network. This mobility model provides a closer match of mobile node movements with real world scenarios where node mobility may be constrained by roadways or other factors than conventional random mobility models. This mobility model was incorporated into ad hoc network models with different connectivity characteristics, and enabled accurate evaluation of routing protocol performance.

(6) A new border node based routing protocol (BBR) was developed and evaluated for MANETs in rural areas. The routing algorithm is based on the combination of a novel distributed border node selection algorithm, together with a new neighbor discovery algorithm. This routing algorithm combines the store-and-forward concept with border node selection rules, which makes it particularly suitable for a partially connected as well as a fully connected mobile ad hoc network. The border node selection algorithm

places no restrictions on node mobility. It can be used not only for mobile nodes with random movement patterns, but also for mobile nodes constrained to highway movement.

(7) The BBR protocol was developed and implemented in OPNETTM Modeler. Extensive parameter sensitivity analysis for the BBR protocol was carried out. Simulation results show that with appropriate selection of the BBR configuration parameters, data packets can be routed with very high delivery ratio in partially connected ad hoc networks. The performance of BBR is compared with the DSR routing protocol with both Random Waypoint mobility and GTI mobility models applied. The simulation results with both mobility models demonstrate that BBR yields better performance than DSR when the network is partially connected and provided comparable performance to DSR when network is fully connected. Tests show that BBR also provided better performance than a Random Node Selection algorithm (RNS) when network is partially connected.

Future work directions: The research work done in this dissertation has demonstrated that wireless communications show great potential in providing broadband access services in rural areas. At the same time, it also points to several areas for further investigation. For fixed wireless broadband access, new radio technology embodied in the IEEE-802.16 standard, referred as the Worldwide Interoperability for Microwave Access (WiMAX), is now becoming available. The new possibility offered by WiMAX for broadband fixed wireless access in rural and sparse areas needs to be explored. The high throughput and long range offered by WiMAX could be significant advantages for rural areas. Both fixed point-to-multipoint and multi-hop Internet access models could be extended to include WiMAX access and backhaul links. Hybrid WiMAX/Wi-Fi architectures could also be examined.

The fixed and mobile network studies carried out in this dissertation do not take terrain effects into account. The current version of the GTI mobility model ignores terrain effects and assumes free space and line of sight propagation. Further improvement of GTI mobility model to take into account the terrain effects and more realistic radio propagation models could be used to evaluate the effectiveness of routing protocols in sparse networks.

The current version of the BBR routing protocol used a fixed Hello message interval for the neighbor discovery. To reduce the control overhead of the neighbor discovery process while retaining the advantages of BBR, an adaptive neighbor discovery process with a variable hello interval value adjusted based on the local node density needs to be further studied. BBR has better performance than DSR when the network is partially connected and has more control overhead than DSR when the network is fully connected. Further research might also be carried out to develop a technique to automatically switch to the most appropriate routing protocol based on the characteristics of the underlying network.