Efficient Routing and Communication Algorithms for Wireless Mesh Networks

by

Liang Zhao

A thesis submitted

in partial fulfillment of the requirements of

Edinburgh Napier University,

for the award of Doctor of Philosophy

February 2011

Abstract

There has been an increasing demand to create wireless technologies that allow people to connect their home networks together as well as to the Internet through wireless community networks (WCNs). Although several network architectures have been studied and indeed deployed in practice for WCNs due to evolving technology and locales, wireless mesh network (WMN) has been undoubtedly the most popular architecture that proved clear superiority in different roles. Indeed, there are many advantages to enabling such mesh connectivity and forming a community mesh network. This will improve quality of life, the digital engagement and partnerships of the rural and isolated communities and will promote and facilitate further the seamless economy.

The design of efficient communication is a challenging issue for the success of the next generation WMNs to handle real-time and Quality of Service sensitive applications and to satisfy both service providers and users. Unfortunately, the significant point in the literature is that today’s cutting edge routing standards in WMNs are not perfectly equipped to cater to this task as these standards come with an inherent complexity and suffer from innate problems with respect to efficient communication based applications. Thus, the aim of this thesis is to enable the WMNs to handle various efficiently real-time multimedia applications, in different operating conditions. This is achieved in this thesis by making three major contributions. Firstly, a new load balancing aware multicast routing protocol is designed in order to enhance the performance of multicast communication in WMNs. Secondly, novel unicast and multicast schemes have been devised as centralized routing algorithms in WMNs. Thirdly, new cross-layer routing metrics have been proposed to further increase the efficiency of protocols in WMNs. Finally, rigorous performance analysis in different operating conditions has been conducted to confirm the superiority of the proposed solutions over well-known existing work.

To:

the memory of my beloved grandparents, and all other family members.

Acknowledgement

First and foremost, I would like to thank God for giving me the opportunity to carry out this research and the patience and strength to finish it. Secondly, I would like to express my sincere gratitude to my supervisor Dr. Ahmed Y. Al-Dubai. I thank him for leading me into this exciting research area and giving me help whenever I need. I am grateful for his constant encouragement, support, supervision and guidance during my PhD at Edinburgh Napier University. I would also like to thank my second supervisor Dr. Xiaodong Liu for giving me help and valuable suggestions and comments during my PhD.

Thirdly, I would also like to thank the kind help from other university staffs, Dr. Imed Romdhani, Prof. William Buchanan, Prof. Jessie Kennedy, Prof. David Benyon, Dr. Ken Chisholm, Mr. David Bonas and many others during my study.

Fourthly, I would like to thank my friends at Edinburgh Napier University for the good time and the friendship. In particular, my thanks to Chengpu Li, Lin Li, Zakwan Jaroucheh, Zhe Wang, and many others.

Finally, I would like to thank the support and understanding from my girlfriend Yang Yang, and my parents.

Publications during the course of this research

Journals:

L. Zhao, A. Al-Dubai, and G. Min, "An Efficient Neighbourhood Load Routing Metric for Wireless Mesh Networks," *Journal of Simulation Modelling Practice and Theory*, 2011, accepted.

L. Zhao, A. Al-Dubai, and G. Min, "GLBM: A New QoS Aware Multicast Scheme for Wireless Mesh Networks," *Performance Evaluation and Optimization of Ubiquitous Computing and Networked Systems, Journal of Systems and Software*, vol. 83, iss. 8, pp. 1318-1326, August 2010.

Refereed Conference:

A. Al-Dubai, L. Zhao, A. Zomaya, and G. Min "A Novel Multicast Gateway Based Multi-hop Routing for Wireless Mesh Networks," accepted to be published in *IEEE International Conference on Communications ICC2011*, 2011.

L. Zhao and A. Al-Dubai, "Routing Metrics for Wireless Mesh Networks: a survey," accepted to be published in *2011 World Congress on Computer Science and Information Engineering*, 2011.

L. Zhao, A. Al-Dubai, and I. Romdhani, "Novel QoS-Aware Gateway Centralized Multi-hop Routing for Wireless Mesh Networks," in *Proc. of 10th IEEE CIT*, 2010.

L. Zhao, A. Al-Dubai, and G. Min "Cross Layer Neighbourhood Load Routing for Wireless Mesh Networks," in *Proc. of 24th IEEE/ACM IPDPS*, 2010.

L. Zhao and A. Al-Dubai, "GLBM: A New QoS Aware Multicast Scheme for Wireless Mesh Network," in *Proc. of 23rd IEEE/ACM IPDPS*, 2009.

L. Zhao (Advisor: A. Al-Dubai), "Poster: QoS Aware Multicast Communication Algorithms," *TCPP PhD Forum Poster Presentations, abstract in Proc. of 23rd IEEE/ACM IPDPS*, 2009.

L. Zhao, A. Al-Dubai, X. Liu, "A new multicast algorithm for wireless mesh networks," in *Proc. of the 9th Annual Postgraduate Symposium, The Convergence of Telecommunications, Networking and Broadcasting*, Liverpool John Moores University: EPSRC, 2008.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

Communication traffic over wireless devices has been increasing enormously, handling a wide range of collective communication based applications. The Internet, in future world, will be everywhere, and will be a vastly deeper and more powerful environment than we know today. The information society will be a networked society, with individuals and enterprises always linked locally and to the Internet. These networks aim to provide socially beneficial functions, from monitoring individual health in urban and rural communities to supporting global enterprise with efficiency and resilience.

Even today, the Internet has a significant influence as a requisite part in individuals' life. Many Internet access devices are used nowadays, including portable devices such as mobile phone, PDA, laptop PC, PlayStation Portable [1], etc and non-portable devices such as desktop PC, PlayStation 3 [2], and digital TV etc. At the service provider side, they try to offer services within different features, from webpage browsing to video browsing in wide areas in order to meet various demands. However, such services cannot be afforded by applying traditional networking technology owing to two points: *bandwidth* and *construction cost*. In the bandwidth side, the bandwidth of 2.5G GPRS [3] cellular network [4] [5] is up to 48 kbps, and even 3G [6] cellular network provides data rates at 384 kbps [7]. In fact, this speed is only enough for the web browsing and emailing applications, but not adequate to support neither high quality video streaming nor Internet gaming. It is worth indicating that only the cable broadband such as DSL [8] [9] and Wi-Fi [10] can offer high speed Internet access. On the other hand, as for the construction cost, it is needed to set up base stations for cellular networks (2G [11], 2.5G, 3G and forthcoming 4G [12]) which are costly. In addition, the installation of the cable broadband is also expensive. Therefore, Wireless Mesh Network (WMN) [13] is an ideal candidate to construct scalable high-bandwidth broadband network with low deployment cost.

1.1 Wireless Community Networks

There has been an increasing demand to create wireless technologies that allow people to connect their home networks together as well as to the Internet through *wireless community networks (WCNs)* [14] [15]. There are many advantages to enabling such connectivity and forming a WCN. For example, when enough neighbours cooperate and exchange packets, they do not need to individually install an Internet "tap" (gateway) but instead can share faster, cost-effective Internet access via gateways that are distributed in their neighbourhood. Packets dynamically find a route, hopping from one neighbour to another to reach the Internet through one of these gateways. Another advantage is that neighbours can form community partnerships, beyond their building walls, with neighbourhood groups, schools, and small businesses to develop programmes and create opportunities to promote economic growth. A third advantage is that this technology allows bits created locally to be used locally without having to go through a service provider and the Internet, which results in saving energy and reduction in Internet traffic. In addition, neighbourhood networks allow faster and easier dissemination of cached information that is relevant to the local community. Moreover, these networks typically have low maintenance overhead and high data rates.

WCNs use wireless connections to networking computers/nodes and connect them to the Internet. Because this technology has become very affordable, many communities have begun to build WCN to connect the whole neighbourhood or even the whole city. Connection speeds can easily exceed those of cable broadband or DSL [8] [9]. While some cases, cities have paid organizations to set up these networks to stimulate growth, in other cases, volunteers have formed non-profit associations to deploy these networks, connect far-flung rural areas or to connect poor areas within urban communities. This takes place nowadays worldwide in line with many future broadband strategies and initiatives, such as the e-Europe [16] initiative that aims to develop the network technologies and architectures, allowing a generalised and affordable availability of broadband access to European users, including those in less developed regions, peripheral and rural areas. Worldwide, governments are increasingly realising that broadband access to the Internet will be central to the economic development of their countries. Wide availability of broadband communication would have a significant impact on their economy, and several EU member states have started reviewing the situation regarding broadband on their territories.

Nowadays, there is a list of networks for constructing WCNs, such as Wi-Fi cluster, WMN, Wireless User Group (WUG) [17], and so forth. Although different networking technologies have been studied and indeed deployed in practice such as Wi-Fi cluster, WMN, WUG, WMNs have undoubtedly been the most popular WCN technology used in practice as it has proved clear superiority in different roles and scenarios. Thus, the main focus of this study is on the WMN technology in order to investigate community mesh networks.

WMN has shown a great potential locally and globally to deploy WCNs. Due to its high quality and cost effective performance, WMN has been attractive to different Internet service providers, aiming to fill the UK's broadband black spots. For instance, in Scotland, Speednet Scotland [18] uses a mesh network to provide wireless broadband access to the area surrounding Troon, Ayreshire, where many telephone exchanges were unable to support broadband until recently. LocustWorld [19] is a pioneer within this growing market, offering a range of broadband access services to business and residential customers in different areas in the UK and worldwide. In addition, the Southampton Open Wireless Network (SOWN) [20] is a project that uses WMN to build non-profit WCNs in Southampton. Globally, many zones and cities have been equipped with WMNs to facilitate wireless broadband access services to urban and rural communities.

1.2 Wireless Mesh Network

Due to its promising technology, WMN has recently attracted much attention [21] [22] [23] [24] [25] [26] [27] [28] [29]. It is becoming a major avenue for the next generation wireless networks. This is due mainly to the revaluation of using basic radio frequency physics to provide a robust, flexible, standard-based architecture. Usually, this architecture offers instant, highly flexible, and low-cost mobile broadband communications to different communities through the readily attainable multi-hop connections. In WMNs, signals are routed optimally and nodes can automatically join and leave the network at any time. Furthermore, networks can be established instantly virtually anywhere, even in places with no fixed infrastructure.

WMN can be used for a wide range of applications including, but not limited to, providing broadband Internet access, voice over IP (VoIP) [30] in neighbourhoods, offices, gaming, video surveillance, public safety, industrial monitoring and media repositories, etc. In addition, WMNs can also play an important role in disasters reporting and emergency networking. The deployment of WMNs is usually within a sizable geographic area, for instance, it can be used to construct both medium-sized community mesh networks [31], and large-scale metropolitan mesh networks [28]. Especially, community mesh networks can be deployed by applying the mesh network technology into WCN in order to further enhance the scalability and adaptability. In developing countries and rural areas [26], it is necessity to find a low-cost and easy-deployed network technology. Therefore, WMN is useful with its wireless infrastructure to provide local communication and Internet access. In education, the setup of WMNs can eliminate the cost of burying cables in the old college buildings, keeping all students connected. In healthcare, WMNs can connect all the operate rooms, labs and offices in hospitals. This ensures the doctors and nurses can update and view patient information instantly. In sport venues, WMNs can provide Internet connectivity to fans and game organisers. For example, in a basketball game with such a network, fans can get the real-time game news, statistics and video highlights from their PDAs or smart phones, while game organisers can upload all game relevant information to the database. In the temporary venues such as street fairs and outdoor concerts, it takes only minutes to set up and remove WMNs. It keeps the cable links to a minimum.

In addition, the military forces plan to use WMNs to connect all the computers, laptops, and soldier-end devices together in the battlefield [32]. This enables the command centre to monitor and command the individuals. Besides, WMN is also the architecture of emergency network during the disasters such as fire accident and earthquake. Due to a wide range of WMN applications, there has been an increasing demand to develop such wireless technologies that allow individuals to connect their home/office networks together [22] [23].

Thus, this thesis rises to the challenges posed by the continuous needs for delivering efficient communications over WMNs. In particular, this thesis aims to design and analyse efficient routing algorithms for WMNs in *unicast routing*, *multicast routing* [33], and *cross-layer routing metrics.* These are now considered, in turn, in the following sections.

1.3 Unicast Routing

Unicast communication is widely studied in both wired and wireless networks. As an approach for the network layer in OSI model [34], unicast is the communications between a single sender node and a single receiver node in a network. Unicast routing is usually regarded to describe the discovering of one-to-one routing path from a single sender to a specific receiver among several nodes. Although unicast is one of the most common networking communication technologies, the current research is still limited to supporting high standard unicast-based applications such as file transfer, video chat, etc.

1.4 Multicast Routing

A class of crucially important communication patterns that have already received considerable attention are group communication operations [35], since these inevitably place a high demand on network bandwidth and have a consequent impact on algorithm execution times. In routing layer, multicast refers to the delivery of the same message originating from a given source to a group of nodes in the network. If the set of destinations is reduced to one node, the communication becomes point-to-point unicast, while it becomes a broadcast operation if the destination set consists of all the other nodes. Multicast is one of the most primitive group capabilities of any message passing network. Thus, multicast routing is the term used to describe the finding of paths from source node to multiple destination nodes in a network. It is central-to-many important distributed application in science and engineering which is fundamental to the implementation of higher-level communication operations such as gossip, gather, and barrier synchronisation. Multicast has emerged as one of the most important areas in the field of wireless communication. It becomes a challenging issue due to the necessity of providing communications and coordination among a given set of nodes. There are many ways to apply multicast technology, e.g. video conference, online gaming, online multicast video, long distance education, etc. However, there is very limited research on multicast in WMNs.

1.5 Routing Metrics

In the network layer, a routing metric is also a key element since it is implemented in the routing protocol to judge the superiority of a route over other alternate ones. The routing metrics cover a set of routing constraints. This includes bandwidth, network delay, path length, load balancing, reliability and communication cost and so forth. In addition, the improvement of one aspect normally results in other aspects. For example, a good communication cost routing metric also performs well in increasing throughput, and reducing delay, etc. As mentioned in [36], there is a need for new routing metrics to improve the performance of WMNs through capturing more constraints. Hence, the design of routing metric is very important to improving the overall performance of WMNs.

1.6 Cross-layer Design

In the open systems interconnection (OSI) [34] architecture, the design of protocols are distributed in different layers. In a protocol, the employed algorithms and schemes are not constrained in the same layer. In other words, the algorithms and schemes can cooperate among different layers transparently in a protocol. Information is shared among all layers to obtain the highest possible adaptivity. This design methodology is so called “cross-layer design”. It can yield significantly improved overall performance. For example, a network layer routing protocol can find the best available path with enough bandwidth in a way of interacting with MAC layer. The cross-layer design and protocols are very essential approaches for wireless multi-hop networks including WMNs. However, existing cross-layer algorithms have been designed only for elastic traffic that only tolerates packet delays/losses. They are not, however, applicable for inelastic video traffic with high data rates and tight delay constraints. Indeed, the developments of new schemes, mechanisms, and systems associated with the cross-layer design and protocols will have a significant impact on the next generation of mobile wireless communications and networks.

1.7 Design Constraints

WMN has an outstanding design space in which the designers have to make choices in many dimensions to construct networks. Thus, several design assumptions are made to constrain the design space.

Both the mesh router (MR) and the mesh client (MC) are routing devices that operate based on wireless radio. In the evaluation part, these devices are also assumed to connect to the Internet backbone via gateways. Notice that, by combining the MR layer and the MC layer, the term “mesh node” is used to represent the mesh devices in the description of routing algorithms.

As for the routing algorithms and schemes including routing metrics, all devices in WMNs are assumed to use single radio and antenna, and operate at the same bit-rate, within the same transmission power and transmission range.

When describing the routing schemes, the terms, “routing algorithm” and “routing protocol” are used interchangeably. The term “mesh node” and “wireless router” are used interchangeably to describe the mesh devices, especially during the discussion of the simulation experiments in this thesis. In addition, the network model is assumed relatively static, thereby determining the performance of community mesh networks. As stated in [13], most of mesh nodes are stationary in this network model, whilst only a small portion of mesh nodes are mobile.

1.8 Major Challenges and New Horizons

The network layer [34] determines how data transmits between nodes, delivers packets according to the address and operates congestion control to reduce the cost of delivery. In the network layer, unicast is the most commonly communication operation for one-to-one message delivery. Besides, multicast is extremely important and instrumental group communication operation. Through studying unicast and multicast, both one-to-one and one-to-many operation schemes are investigated in order to implement and handle different services in the application layer. The design of routing metric is a key element to optimise the path selection so as to further improve the quality of unicast and multicast communications.

The main point that is apparent in the literature is that there is a current lack of efficient unicast communication and multicast communication to handle real-time and Quality of Service (QoS) [37] sensitive applications and satisfy both service providers and users in

WMNs. Supporting QoS-aware communications to enable a rich portfolio of real-time and QoS sensitive applications is foreseen to be vital for the success of the next generation efficient WMNs. Unfortunately, existing standards supporting instant communications in WMNs are not perfectly equipped to cater to this task as these standards come with an inherent complexity and suffer from innate problems with respect to QoS provisioning. More importantly, in WMNs, neither unicast communication nor multicast communication have been addressed so far within the context of emerging challenging high-bandwidth applications, such as real-time multimedia applications that consumes growing bandwidth consumption rates. Indeed, these efforts have not resulted in high communication quality provisioning for real-time applications. Thus, there is a growing need for devising and analysing innovative unicast and multicast communication algorithms for WMNs to facilitating efficient communication applications. The long term target of this study is to enable the WMNs to handle real-time multimedia applications, in different operating conditions within the QoS constraints. This study, therefore, will have a significant impact and will open new scientific technological horizons towards optimal communication environment for tomorrow's communities.

1.9 Motivation and Beyond the State-of-the-art

The multicast routing problem, is also known as a constrained Steiner tree problem, has been proven to be NP-complete [38]. Thus, existing study on multicasting focuses on developing heuristic algorithms that can find a near-optimal tree for a single multicast connection based on the current network state under only certain communication quality constraints, such as throughput and delay. However, these algorithms have not addressed the multicast problem within the traffic engineering prospective; thus one can hardly achieve the required performance in WMNs to offer these applications. On the other hand, some recent study supports multicast communication using network coding [39] [40] [41], where all links in the network can be utilised, instead of a tree. Although network coding can achieve the best throughput theoretically, it requires the modification of existing packet forwarding mechanisms, which is hard to achieve [38]. Nevertheless, network coding based multicasting is a promising avenue and is worthy of thorough investigation in our future work, especially within the context of QoS provisioning.

Despite different available communication schemes reported in literature, the routing for WMNs is still infancy and an active research area due to several reasons. Firstly, most existing unicast algorithms and multicast algorithms used in WMNs have been originally devised for other types of wired and wireless networks, omitting the major challenges and specifications of WMNs. For instance, these algorithms treat all network nodes in the same way, omitting the differences between gateways and mesh nodes in WMNs. Secondly, the relay node selection algorithms and routing metrics used in these algorithms have only partially considered the QoS through limited metrics such as hop count, link level, packet loss ratio, *etc*. However, optimising the performance under heterogeneous constraints is still an unexplored issue in WMNs. In other words, existing routing metrics are not enough to capture all constraints and performance parameters in WMNs. Thus, there is a growing need to integrate multiple cross-layer routing metrics, considering the sensitivity of the applications to the QoS and the key role of gateway nodes in WMNs, an issue that is not addressed well by existing work. Integrating cross-layer metrics is one of the aims of this thesis as well as for our future work.

Communications in WMNs can create bottlenecks for implementations of various applications. Regardless of how well the path selections in WMN, routing can severely affect the communication quality. Therefore, efficient routing algorithms including unicast, multicast, and routing metric, are critical to the implementation of high standard network-based applications.

Applications including video chat, file download, and web browsing, etc, are all handled by unicast. To implement these unicast applications in WMNs, strict communication quality is required in order to provide high-bandwidth, low latency network environment. There are many proposals of unicast routing in MANETs such as Ad hoc On-Demand Distance Vector (AODV) [43], Destination-Sequenced Distance-Vector Routing (DSDV) [44]. However, most of them are not capable of working efficiently in WMNs due to the unmatched characteristics between MANET and WMN. Existing algorithms do not take advantage of the fact that all gateways can deal with network management to provide efficient routing.

Similar to unicast, issues such as load balancing routing and high communication quality routing are not considered in existing multicast algorithms [45] [46] [47] [48] [49]. Therefore, both unicast algorithms and multicast algorithms are also required to be designed within either new concepts or re-engineering existing ones.

The design of routing metrics is very important to improving the overall performance of WMNs and wireless multi-hop networks in general. However, many existing routing metrics still work ad–hocly. As a result, this kind of routing metric may only perform well in Client WMNs. In some routing metrics, routing status is measured inaccurately when unicast is used to test the quality of multicast transmission. Some others also use the exchange of probes which may cause large overhead and even worse in large scale networks. In addition, critical routing parameters such as QoS are not provided in existing routing metrics [50] [52] [53]

[54] [55] [57] [58] [59] [60] [61]. Therefore, routing metrics should be designed to adopt the characteristics of WMN to optimise the overall performance. So far, most existing routing metrics focus on the network layer. There is a need to design cross-layer routing metrics in order to enhance the performance of routing protocols.

1.10 Main Contributions of the Thesis

In this thesis, unicast routing, multicast routing and high efficient routing metrics are addressed within the context WMNs. The major contributions of this study are as follows.

- Gateway-cluster based Load Balancing Multicast algorithm (GLBM) is proposed to optimise the traffic load on the shortest path to the gateway during the multicast operation. GLBM exhibits high capability in handling load balancing by distributing traffic load more evenly than the shortest-path multicast algorithms.
- In this thesis, we introduce the Gateway Centralized Multi-hop Routing algorithm (GCMR). To the best of our knowledge, GCMR is the first gateway based unicast algorithm for WMNs. This work shifts the role of the gateway from a simple packet forwarder to a routing orchestrating node. It adopts the leaf-to-gateway update mechanism to keep the gateway updated with instant topology information. The traffic prediction method estimates the future traffic status on the path. It reduces the number of update processes, thus minimises the traffic overhead. In addition, the TTL (time-to-live) setup mechanism is introduced to reduce the radius of update messages in order to further limit the flood of the update process. These mechanisms improve the communication efficiency by reducing both latency and jitter.

- Based on GCMR, Multicast Gateway-Centralized Multi-hop Routing protocol (MGCMR) is proposed to enable the gateway to calculate the best-metric tree-based multicast.
- Neighbourhood Load Routing metric (NLR) is designed to examine the local traffic load within the neighbourhood in WMNs. With this routing metric, the best-metric path is selected to generate the least interference to other nodes.
- Quality of Service (QoS) [62] [63] is a key performance issue in WMNs. Nevertheless, this is addressed lightly in the literature as most of existing work treats all the delivered packets equally. Unfortunately, this affects the real-time applications that are quite sensitive to the communication latency/jitter. Thus, to efficiently handle QoS provisioning in WMN, Packet Priority-Oriented routing metric (PPO) and Packet Priority QoS-aware routing metric (PP-QoS) are presented in this thesis. This is to assign priorities to different applications so as to provide differentiated services. In addition, Packet Priority QoS-aware routing metric (PP-QoS) is designed on the basis of PPO by combining with Interference-Aware Routing Metric (IAR) [59]. This enables PP-QoS to be suitable for WMNs under different traffic load situations.

1.11 Outline of the Thesis

Chapter 2 shows a brief overview of wireless networking technology including WMN, MANET and cellular network. This chapter also accommodates background and previous research efforts relating to routing technologies.

In Chapter 3, a new routing algorithm, namely, Gateway Centralized Multi-hop Routing algorithm is proposed for WMNs. This protocol is designed to improve the communication efficiency through the routing management. New mechanisms are introduced which are the leaf-to-gateway update mechanism, the traffic prediction method and the TTL setup mechanism. The simulation model of this algorithm is implemented in NS-2 [64] and compared with its counterparts in this chapter.

In Chapter 4, Gateway-cluster based Load Balancing Multicast algorithm is proposed to enhance the performance of multicast in WMNs. The experiments conducted by NS-2 are also provided in this chapter.

In Chapter 5, a new novel algorithm namely, Multicast Gateway-Centralized Multi-hop Routing protocol is proposed. MGCMR enables the route management in the gateway with the leaf-to-gateway update mechanism to maximise the communication efficiency. The simulation model and results are also shown in this chapter.

In Chapter 6, Neighbourhood Load Routing metric (NLR) is proposed as a load balancing routing metric to enhance routing efficiency of existing routing algorithms. This chapter also shows the simulation model and results by implementing the proposed routing metric with AODV [43] in NS-2.

In Chapter 7, two new QoS-aware routing metrics, Packet Priority-Oriented routing metric (PPO) and Packet Priority QoS-aware routing metric (PP-QoS) are designed to satisfy different communication requirements of various applications offered by WMNs. The evaluation model and results are also provided in this chapter.

Chapter 8 summarizes this thesis, draws the conclusions and also outlines the future work direction.

Chapter 2

Background

2.1 Core Issues of Wireless Mesh Network

Wireless Mesh Network (WMN) is known as the next generation network connectivity. It has a list of distinguished characteristics such as multi-hop, gateway involvement, low-mobility, multi-radio, low-cost, easy-deployed, self-healing and self-organised and so forth. Each mesh node acts as a host and a router as well, relays packets for other mesh nodes. In other words, a mesh node can obtain Internet access from multi-hop routing of a reachable node. This reachable node is normally a gateway node with network connectivity.

There has been deployment of WMNs in hundreds of cities and rural areas over the world, e.g. cities such as New Orleans, Seattle, and rural areas of Ghana and Zambia [65]. All of existing deployments could have profited from such a network. For instance, it can be applied to provide network access as well as facilitating services in emergency circumstances such as hurricane relief. In addition, places are linked up within the mesh networks, communicate with each other as soon as finding peers within reach, forming an autonomous network [66]. For the above advantages, more and more municipalities are recognizing that facilitating Internet access is part of their responsibility towards citizens. In this context, they are planning to bypass traditional Internet providers, open access to the Internet in a more direct way [67]. As a new cost effective technology, WMN is a natural and desirable candidate for constructing a resilient, locally networked access to communication infrastructure [13] [68].

As illustrated in Figure 2.1, a WMN consists of mesh clients (MC), mesh routers (MR) and gateway. Both MRs and MCs are designed to increase the network coverage by only assembling wireless radio. Gateway uses both wireless radios and fibre optic cables. For this reason, there are three layers in a WMN: *Gateway layer (GW)*, *MR layer* and *MC layer* [13]. Gateway relays messages between Internet backbone, MR and MC. Under each gateway, each MR is connected with neighbouring MRs and gets access to the backbone via the gateway. MCs also connect to each other to form a mesh under each MR. To gain the Internet connectivity, messages from a MC is directed to the backbone through MRs. Routing devices of both MR layer and MC layer use wireless radio and connect to the Internet backbone via gateways. To simplify the description of routing algorithm, MR layer and MC layer are combined. Therefore, the term "mesh node" (MN) is used to represent the mesh devices including both MRs and MCs in this thesis. An illustrated example of such a network structure is shown in Figure 2.2. However, to realize the community-based goals within WMNs, one has to solve many challenging problems including: *cross-layer routing optimisation, relay node selection, efficient communication, capacity, scalability, range enhancement, privacy*, *security*, *self-stabilizing, multi-path routing, auto-configuration,* and *bandwidth fairness*, etc. In the light of the above issues, this study focuses on routing design, analysis and optimisation of efficient communications in WMNs.

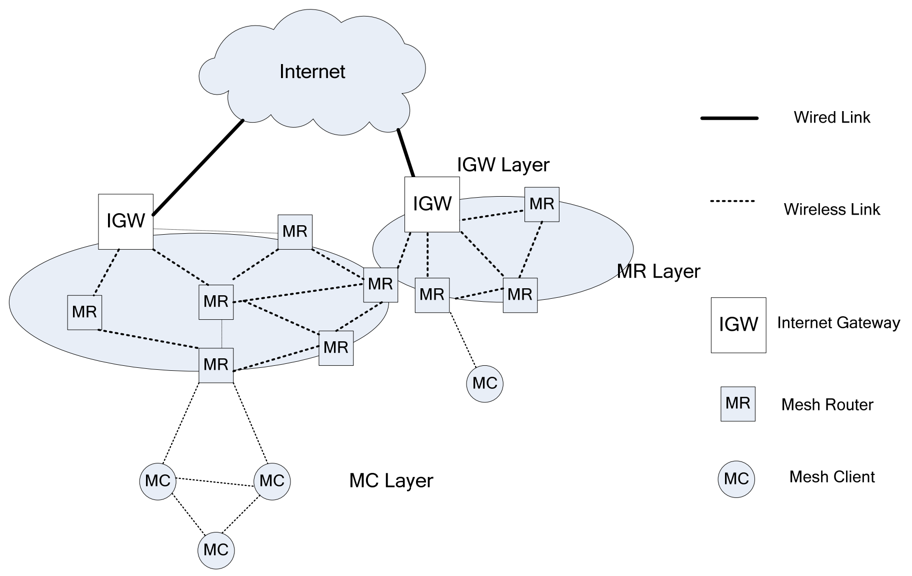


Figure 2.1: The architecture of a typical WMN

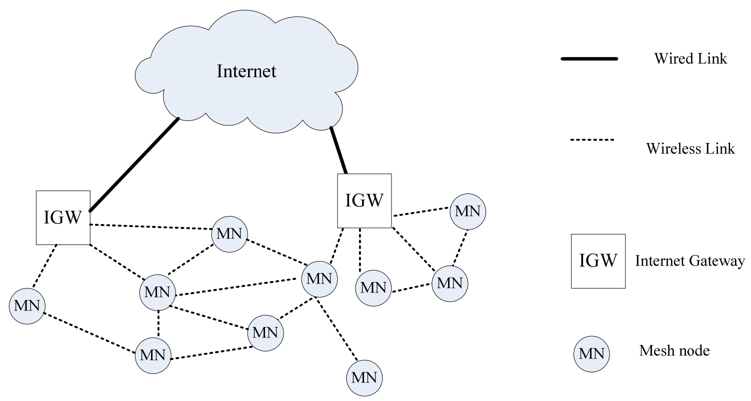


Figure 2.2: The architecture of a WMN presented in this thesis

2.1.1 The Architecture of Wireless Mesh Networks

According to [13], there are different types of WMNs:

- *Infrastructure/Backbone WMNs*: there are two kinds of nodes in Infrastructure WMNs, gateway nodes and mesh nodes. Compared to the network shown in Figure 2.2, gateway of Infrastructure WMN forms an infrastructure for connected mesh nodes only by one-hop distance. All mesh nodes form a mesh with self-configuring, self-healing links among themselves. Conventional clients with Ethernet interface can also be connected to the gateway via wired links. For instance, MIT RoofNet [24] is an example of Infrastructure/Backbone WMNs.

- *Client WMNs*: a Client WMN is similar to a MANET with low mobility nodes. Peer-to-peer connections are set among mesh nodes in Client WMNs. However, there is no gateway in Client WMNs. For example, the One Laptop per Child project [27] deploys Client WMN structure for educational purposes.
- *Hybrid WMNs*: as shown in Figure 2.1 and Figure 2.2, the architecture of Hybrid WMN combines both Infrastructure WMN and Client WMN. Like Client WMNs, a mesh node connects to other mesh nodes by either single-hop or multi-hop in Hybrid WMNs. In addition, the gateway forms an infrastructure to provide the Internet connectivity to all the mesh nodes in a way similar to Infrastructure WMNs. In other words, the infrastructure of Hybrid WMN is highly adaptable to other networks such as the Internet, Wi-Fi [10], WiMAX [25], cellular networks [67] [69] [70] [71], and sensor networks [72] [73]. Therefore, Hybrid WMN is considered as the most applicable architecture. For example, Open-Mesh [74] creates low-cost Hybrid WMN solutions to provide Internet connection to end-users.

From another point of view, WMN is open to all the network capable devices as a combination of the wired network (backbone) and wireless network (backhaul). These devices include desktop PC, laptop PC, mobile phone, and PDA. Furthermore, mesh-enabled camera, GPS, even fridge, washing machine and other smart home devices etc, can access to Internet through WMNs. As a consequence, the design goal of routing algorithms shifts from finding connectivity between source and destination to maintaining high efficient communication, thus it overcomes the bottlenecks in large-scale WMNs. New routing algorithms should also aim to provide scalable implementations of computationally intensive applications.

Since it is similar to Client WMN, Mobile Ad-hoc Network (MANET) [75] is a special type of WMN. In MANETs, each node forwards data for other nodes dynamically to form multi-hopping communication. However, there is no gateway in charge of authentication or security services in MANETs. Compared to the low-mobility nodes of WMNs, nodes are mobile in MANETs. Therefore, nodes are normally lack of sufficient energy. In addition, due to the mobility nature of devices, the topology may change rapidly.

The two major differences between WMN and MANET are *gateway involvement* and *node mobility*. Compared to MANET, most traffic is expected to flow between the mesh nodes and the backbone through gateways. In a WMN, besides handling all the traffic flows between mesh nodes and the Internet, gateway should be a de facto network administrator in charge of orchestrating the internal traffic flows. In MANETs, all the nodes are mobile. On the contrary, most devices are stationary or with limited mobility in WMNs. Only a small portion of devices such as mobile phones, are moving in WMNs. Therefore, for the above reasons, it is necessitated to either modify existing routing algorithms or design new routing algorithms with the aim of enhancing the communication efficiency in WMNs.

Multi-hop Cellular Network (MCN) [76] can also be regarded as a special type of WMN. It is similar to Hybrid WMN, but with high mobility nodes. A MCN has access point(s) acting as a gateway to manage the routing requests for mobile phones. A comparison between WMN (hybrid), MANET, and MCN are shown in Table 2.1.

Therefore, designing a new technology or reinventing existing technology should consider the characteristics of WMNs. For example, to apply a MANET routing algorithm to WMN, different mobility level and battery life should be taken into account.

Network /Feature	WMN	MANET	MCN
Gateway involvement	Yes	No	Yes
Node mobility	Low	High	High
Node energy supply	Sufficient	Constrained	Constrained
Internet access	Yes	No	Yes
Deployment cost	Medium	Low	High
Maintenance cost	Low	High	High

Table 2.1: A comparison between WMN (hybrid), MANET, and MCN

2.1.2 Characteristics of Wireless Mesh Networks

- *Support ad-hoc and Capability of self-forming, self-healing and self-organization*: WMN enhances the network performance through providing low-cost, easy-deployed and easy-configured structure to form mesh connectivity.
- *Compatibility and interoperability*: WMN is compatible with IEEE 802.11 standards [10] [50] [77] to offer both the mesh connectivity and conventional connectivity for mesh nodes. It is also inter-operable with other wireless networks such as WiMAX [25] and GPRS [3] etc.
- *Multi-hopping:* this character enables WMN to expand network coverage without additional radio capability.
- *Gateway involvement:* this feature enables the orchestrating capability of the gateway that situates the mesh nodes and the backbone to forward messages and process various routing tasks.

- *Mobility*: a mesh node can be either mobile or stationary depending on the type of mesh device. However, in community mesh networks or neighbourhood mesh networks, only a small portion of mesh nodes are mobile.
- *Network access*: compared to MANETs, WMNs offer both backhaul access and backbone access.
- *Power-consumption constraints*: most nodes are with sufficient power supply. However, mesh nodes with mobility are normally power constrained.
- *Multiple radios and Multiple Channels*: mesh devices can be equipped with multiple radios to perform routing and access functionalities. This enables the separation of two main types of traffic in the wireless domain. Therefore, mesh nodes can communicate with each other in non-interference channels. This significantly improves the network capacity. In this thesis, all mesh nodes are assumed with single radio and single channel capability. However, these proposals can be converted to multiple radio algorithms.

2.2 Routing Constraints

2.2.1 Load Balancing

A load balancing feature in a network is defined as a situation in which the network is with no uneven traffic load. It can provide high-quality communication with low jitter, low latency, low error rate and high bandwidth. In other words, uneven traffic distribution always leads to problems such as high delay time, network congestion and high packet transmission failure rate. Therefore, optimising the load balancing capability is extremely important to guaranteeing communication quality for WMNs. According to [78], load balancing problems can be divided into two types in WMNs, *gateway load balancing problem* and *centre load balancing problem*.

1) Gateway load balancing problem in WMNs

This issue has been widely studied [79] [80]. As a gateway locates between a WMN and the Internet backbone, overloaded traffic aggregation at gateway may degrade its performance. Consequently, it may limit the overall capability of the network. This problem normally appears in WMN with multiple gateways. The fairly distribution of the traffic load and bandwidths of backbone links among gateways is applied to optimise this problem.

2) Centre load balancing problem in WMNs

The mesh nodes lying on the geographical centre of network may suffer highly from overloaded traffic compared to other nodes. There are mainly two reasons. First, the centre nodes always lie on the shortest path to gateway. Consequently, they forward more packets because of the multi-hop nature of WMNs. Second, since mesh nodes are comparatively static or with low mobility, a centre node remains in the central position for a long time. Existing proposals [78] [81] [82] [83] [84] normally solve this problem by routing the data packets through the node with the lowest traffic load. In such solutions, the load status of each mesh node is either exchanged with all its neighbour nodes, or broadcasted among all the mesh nodes in the network, or can be sent towards the central gateway in Hybrid WMNs. Then, based on the load information, the least traffic load path is selected for relaying packets.

This thesis concentrates on developing algorithms including multicast routing algorithms, unicast routing algorithms and routing metrics to optimise the load balancing capability in WMNs.

2.2.2 Quality of Service

In networking, Quality of Service (QoS) [62] [63] is mainly referred to two aspects. In the first aspect, some existing work [55] [68] [80] [85] [86] [87] considers QoS as a certain level of service quality through gaining the required latency, jitter, error rate, bandwidth and so on. In this thesis, the term "communication quality" is used to define this aspect. In the second aspect, it aims to provide a range of services to fulfil differentiated communication needs. In this thesis, only the second aspect is used to represent the term "QoS". Apparently, new design methodologies such as cross-layer design, may discover the alternate way of achieving QoS.

To achieve the goal of QoS, the admission control mechanism and the traffic control mechanism are normally applied in order to provide preferential services to end users. The admission control mechanism determines which/how/when an application is permitted to use network resource. The traffic control mechanism is used to classify, schedule, and mark packets based on priority of applications. This mechanism distinguishes different service classes and controls traffic in the network. Thus, it can provide differentiated services to various applications. For example, the traffic control mechanism assigns high priority to a video conference application with instant communication requirement to reduce latency. With applying the traffic control mechanism, two QoS-aware routing metrics are proposed in Chapter 7.

Through providing differentiated services, an application is guaranteed to achieve a certain level of bandwidth, latency, jitter, and error rate to implement it in the best network condition.

These four characteristics are also evaluation metrics used to measure the communication quality of a network.

- *Bandwidth* is the transmission rate of traffic flow in the network. In WMNs, it is referred to the term "throughput". It calculates the capability of network to accommodate traffic/messages. A higher throughput of unicast/multicast session determines greater bandwidth provided for the communication.
- *Latency* is measured as the duration time of a packet travelling from one point to another point. In WMNs, it is referred to the end-to-end delay.
- *Jitter* refers to the latency variation in the communications. In networking, jitter is also an important performance metric to evaluate the capability of the algorithms in handling instant communication. It examines the variance of arrival times among packets to weigh packet delay.
- *Error rate* is the rate of system error in the network communication. In networking, packet delivery ratio (PDR) is used to present error rate. It is the percentage of successful packet delivery in a given period of time.

In addition, probing overhead/routing overhead determines the number of request messages generated during path/tree/mesh discovery and maintenance. This performance metric is also applied to study the impact of probing overhead to the network performance.

2.3 Unicast Routing

In wireless multi-hop communication, unicast transmission [33] sends data packets from one source to a single destination. In this section, unicast routing algorithms are divided into unicast routing algorithms for MANETs, unicast routing algorithms for MCNs, and unicast routing algorithms for WMNs. These aspects are reviewed respectively in the following sections.

2.3.1 Unicast Routing Algorithms in MANETs

2.3.1.1 Ad hoc On-Demand Distance Vector

Ad hoc On-Demand Distance Vector (AODV) [43] is a unicast routing algorithm designed for MANETs. As an on-demand algorithm, paths are built only when desired by the source nodes. Sequence number is maintained by each originating node. It ensures the freshness of routes when generating the loop-free, self-starting paths among large numbers of mobile nodes. The route request cycle and the route reply query cycle are applied during path construction. In the route request cycle, when a source node intends to send data packets to a destination without a valid existing path, it broadcasts a route request packet (RREQ) across the network. RREQ contains the source node's IP address, current sequence number, and broadcast ID as well as the most recent sequence number for the destination known by the source node. Once receiving this packet, nodes update the routing information for source node and set up backwards entries to the source node in the routing tables. In the route reply cycle, when receiving a RREQ message, a node only sends a reply packet (RREP) back to the source node by a unicast message if it is either the destination or if it has a path to the destination with corresponding sequence number greater than or equal to that contained in RREQ. Otherwise, intermediate node may rebroadcast RREQ messages. Each forwarder node or each destination node keeps the source IP address and the broadcast ID from RREQ in the routing table and discards the processed RREQ.

Figure 2.3 shows how RREQ packets propagate in a small network to find the best path. On the way of RREP back to the source, the intermediate nodes set up the forward pointers to the destination in the routing table. Once a RREP message arrived at the source node, the source node records the path in the routing table, and begins to forward data packets to the destination. If a RREP message with a greater sequence number or a RREP message with the same sequence number and a smaller hop number is arrived, the source node updates the routing information of destination and sends data packets through the new path. Since AODV uses Hop-count as routing metric, it is also known as AODV-HOP. Figure 2.4 shows RREP follows the shortest path back to the source node. When the source node stops sending data packets, the path is marked as time-out in the source node. Then all forwarder nodes on the path also delete the relevant routing information from the routing tables. When a link is broken during the data packet transmission, the upstream node of the broken link sends a route error (RERR) message. It informs the source node about the unreachable destination. Then, the source node re-initiates the whole route discovery procedure as it receives RERR.

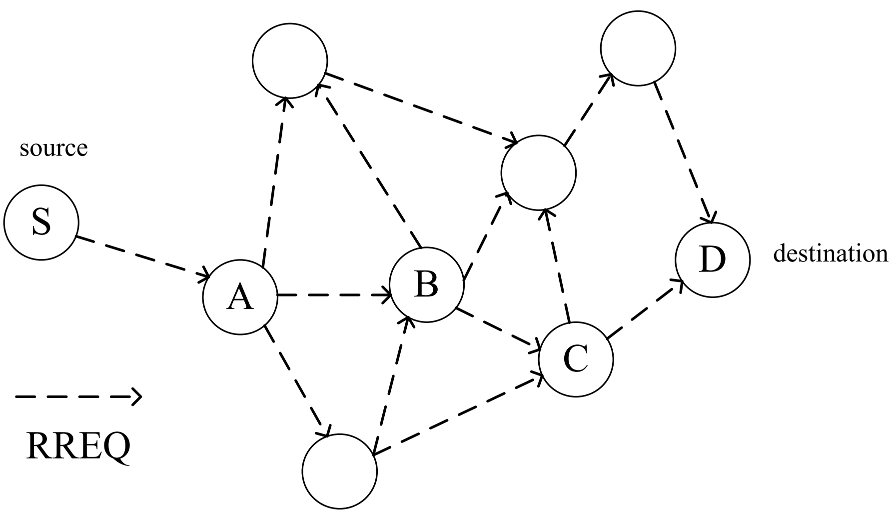


Figure 2.3: RREQ flooding

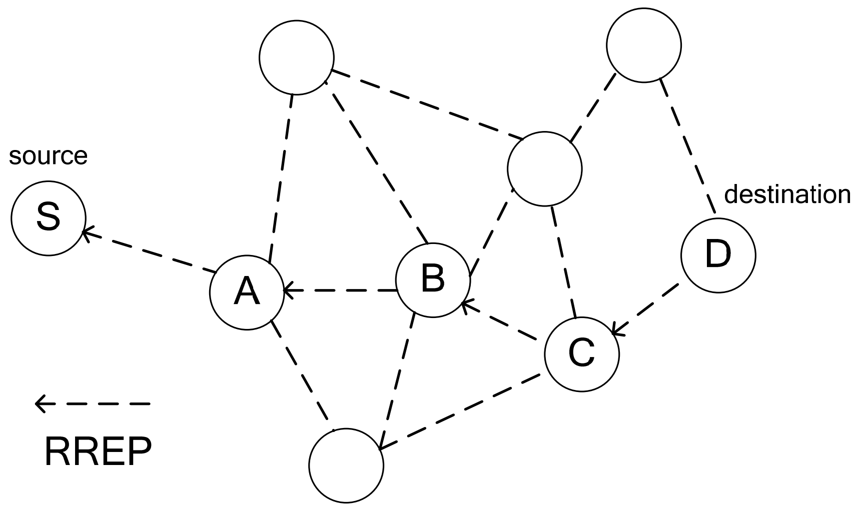


Figure 2.4: Propagation of RREP

2.3.1.2 Dynamic Source Routing

Similar to AODV, Dynamic Source Routing (DSR) [88] is also an on-demand routing algorithm for MANETs. In DSR, a route discovery mechanism is initiated when source node without a valid path intends to send a data packet to the destination node. DSR applies the source routing strategy to broadcast a ROUTE REQUEST message. This includes source id, destination id, a route record with an empty list of addresses of all intermediate nodes and a unique request id towards the destination node. On receiving ROUTE REQUEST, an intermediate node caches the route record. A ROUTE REPLY message may be replied if the destination node is arrived. The destination node stores the route record in the route cache. Then it uses the cached path in the route record for the propagation of ROUTE REPLY back to the source node. Otherwise, if this node is not in the route record of ROUTE REQUEST, it appends its address to the route record and broadcasts the ROUTE REQUEST messages. To avoid the overhead, DSR optionally defines the unique request id for each message in the route discovery mechanism. In addition, these messages are forwarded on a hop-by-hop basis. Unlike other on-demand driven algorithms, there are no proactive periodic probes for neighbour detection, or link status detection in DSR. Therefore, DSR operates truly on-demand to minimise the routing overhead. The route maintenance mechanism is initiated when a node cannot deliver packet to its next-hop node. This node then generates ROUTE ERROR messages towards the source node to find the most viable route. Hence, the broken link is removed from the route cache of the source node.

2.3.1.3 Destination-Sequenced Distance-Vector Routing

Destination-Sequenced Distance-Vector Routing (DSDV) [44] is a table-driven routing algorithm for MANETs. DSDV is designed on the basis of the Bellman-Ford algorithm [89] [90]. In DSDV, each mobile node maintains a routing table with entries of all available destinations. Each entry contains the number of hops to the destination, sequence number to determine a stable route and so forth. Each mobile node exchanges the routing table with the immediate neighbours periodically. In addition, the routing table is also transmitted when there is a change occurred from the last update. There are two ways to update the routing tables, the full dump update and the incremental update. In the full dump update, sending the full routing table to the neighbours may cause network overloaded. In contrast, only entries of available destinations with recent changes from the routing table are sent in the incremental update. When the nodes are relatively static, incremental update avoids extra traffic compared to the full dump update. However, the full dump update is more efficient in the network with high speed mobile nodes. In both update mechanisms, the route update packet is sent with a unique sequence number with the routing information. During the selection, the path with the highest sequence number is chosen as the freshest path. If two paths are with the same sequence number, the shortest path is selected. Table 2.2 lists the routing table of *node_4* on the basis of topology shown in Figure 2.5. As shown in Figure 2.6, when *node_1* approaches to *node_6*, it initiates the update procedure. Meanwhile, *node_2* also starts the update procedure as a link breakage to *node_1* is detected. As the update messages propagate all over the network, all recipient nodes can update their routing tables. Table 2.3 shows the routing table of *node_4*.

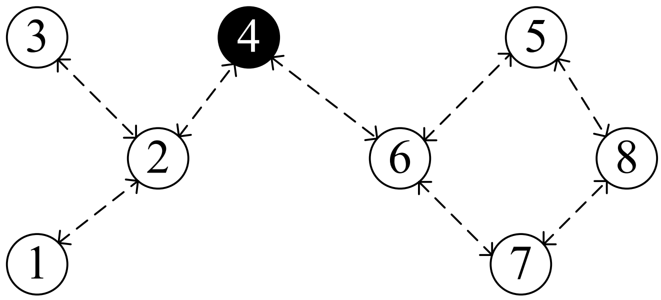


Figure 2.5: An operation example of DSDV

Destination	Next hop	Metric (hop number)	Sequence Number
4	4	0	S400_4
1	2	2	S128_1
2	2	1	S234_2
3	2	2	S124_3
5	6	2	S138_5
6	6	1	S060_6
7	6	2	S078_7
8	6	3	S258_8

Table 2.2: Routing table of node 4

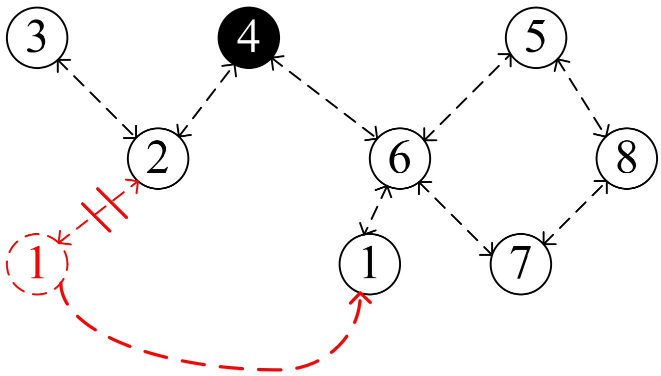


Figure 2.6: An link change example of DSDV

Destination	Next hop	Metric (hop number)	Sequence Number
4	4	0	S408_4
1	6	2	S160_1
2	2	1	S280_2
3	2	2	S134_3
5	6	2	S190_5
6	6	1	S128_6
7	6	2	S090_7
8	6	3	S298_8

Table 2.3: Routing table of node 4 after link change

2.3.2 Unicast Routing Algorithms in MCNs

2.3.2.1 Hierarchical Multi-hop Cellular Network routing

Based on the architecture in [76], Hierarchical Multi-hop Cellular Network routing (HMCN) [69] is proposed by Li et al. Further in [70], Lin and Yu propose a routing algorithm for HMCN by extending AODV. In this algorithm, each mobile node broadcasts hello messages to exchange the instant network topology with neighbour nodes. In addition, nodes also send update messages with neighbour information towards the base station (gateway) periodically.

As receiving update messages, the base station stores the routing information in the routing table. Then it computes the requested routes on the basis of routing table.

2.3.2.2 Base-Assisted Ad hoc Routing

Base-Assisted Ad hoc Routing (BAAR) [71] is a unicast HMCN routing protocol proposed by Manoj et al. In BAAR, the base station (gateway) is responsible for caching paths as well as computing the shortest path for all routing requests. The topology discovery mechanism is implemented on the basis of link update. All nodes including gateway exchange hello beacons periodically. When hello beacons are received, the neighbour node keeps the routing information in the routing table. Each entry of the routing table maintains information of neighbour nodes that includes: *the current received power (Rx_p)* and *the received power last notified to the base station (Rx_{np})*. When there is a significant difference between Rx_p and Rx_{np}, the node broadcasts update messages towards the base station. Therefore, the base station keeps an up-to-date routing table about all the links within the cell. When the base station receives a route request of a path from *node_1* to *node_2*, it computes the shortest path and replies to *node_1*. To avoid the collision in high node density area, BAAR assumes all mobile nodes are capable in multi-channel paradigms. A separate channel is occupied to exchange hello beacons.

2.3.3 Unicast Routing Algorithms in WMNs

2.3.3.1 Wireless Mesh Routing

In [68], Wireless Mesh Routing (WMR) protocol is proposed for unicast routing in WMNs. There are four mechanisms in this protocol: *topology discovery, route discovery, admission control with QoS constraints*, and *bandwidth reservation*. In the topology discovery mechanism, all nodes exchange hello messages to inspect both the connectivity and the hop distance to neighbours. In the route discovery mechanism, for external traffic towards the backbone, packets are routed along the shortest path to the gateway. Then these packets are broadcasted to all nodes in the network for internal traffic. The admission control mechanism is the ensuring of efficient communication deployed in each node. Thus it detects the requested minimum bandwidth for each packet transmission. In addition, it also estimates the end-to-end delay of available paths in the admission control mechanism. This work also presents a bandwidth reservation mechanism. In this mechanism, a mesh node can reserve the channel for each packet transmission flow for a certain period of time.

2.3.3.2 Hybrid Wireless Mesh Protocol

Hybrid Wireless Mesh Protocol (HWMP) is a hybrid routing algorithm which is designed on the basis of AODV. It discovers the optimal path in 802.11s-based small/medium size WMNs [50]. As a hybrid routing algorithm, it consists of two different routing schemes. A reactive routing scheme is implemented for the route discovery, when a proactive tree-based routing scheme is used for constructing the routing tree rooted on the root node (gateway). The Airtime link metric (discussed in Section 2.5.6) is the routing metric implemented for path selection. The paths can be built in either a reactive fashion or a proactive fashion. If there is no root node in the network like Client WMN, on-demand PREQ is initiated towards the destination node in order to discover the best-metric path. If the received PREQ is marked as “Reply and Forward”, an intermediate node with a valid path responds with a reply message (PREP) to the destination. Otherwise, if the received PREQ is marked as “Destination Only”, PREP with the selected path is sent back towards the source node. When receiving PREP, each intermediate node stores the best-metric path in the routing table. It then forwards PREP to the source node. When the source receives PREP, the path to destination is stored. When a root node exists in the network such as Hybrid WMN, it can be configured as two kinds of proactive tree build mechanisms. The first mechanism uses Proactive PREQ to build the tree including root node and all mesh nodes. Proactive PREQ is sent by the root node periodically. Once receiving this message, each mesh node records the path information to the root node. If a recipient mesh node intends to send data packet to the root node, “Proactive PREP” field is set to 0. Otherwise, if there is a change in the path to the root node, “Proactive PREP” field is set to 1. Once receiving Proactive PREQ, the mesh node sends Proactive PREP back to build a fresh path. In the second mechanism, Root Announcement message (RANN) is used. It notifies the mesh nodes about the existence of root node periodically. When receiving RANN, each mesh node creates or updates a path to the root node. It then sends unicast PREQ along the reversed path. When a PREQ message is arrived, the root node may respond PREP to the originating mesh node in order to build a path to the root node. In both mechanisms, root node keeps paths to all mesh nodes, when each mesh node keeps a path to root node. As an example shown in Figure 2.7, during the discovery of the best-metric path, *node_1* (root node) is involved in forwarding route request packet from *node_5*. Then it redirects the request packet to the destination *node_8*. When the request is arrived, *node_8* initiates an on-demand path discovery process by broadcasting a reply towards *node_5*. Finally, the best-metric path is selected when the reply is arrived at *node_5*. Rather than HWMP, Hybrid Routing with Periodic Updates (HRPU) [91] and Hybrid Distance Vector routing (HDV) [92] are also hybrid routing protocols for WMNs. These hybrid proposals do not consider the key role of gateway in managing variety routing tasks either. Besides forwarding the routing requests, there is a need to design centralized routing for WMNs.

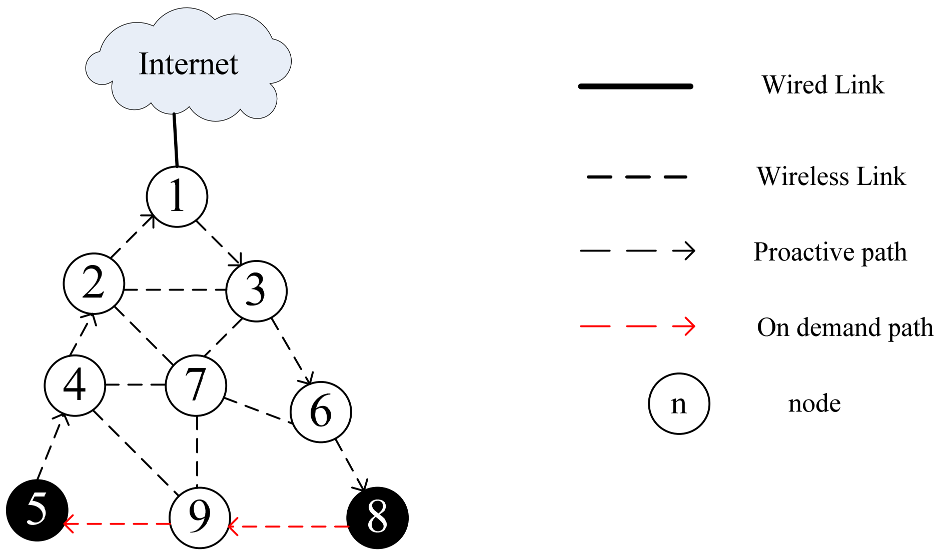


Figure 2.7: HWMP routing with root node inside a WMN

2.3.3.3 Ad hoc On-Demand Distance Vector with Common Gateway Architecture

Ad hoc On-Demand Distance Vector with Common Gateway Architecture (AODV-CGA) [93] is a mesh extension of AODV [43]. It is proposed to support gateway discovery in WMNs. All mesh nodes are connected to a gateway that acts as a proxy to the Internet backbone. On the forward path, the appointed gateway is responsible for sending route reply messages on behalf of hosts in the Internet backbone. The gateway also initiates route requests for nodes within WMN on the backward path.

2.3.3.4 Gateway Source Routing and Gateway Source Routing with Preferred Neighbour

Gateway Source Routing (GSR) [94] is an anycast routing protocol. In GSR, gateway caches and reuses the path information from the route request packets. Anycast is the communication between single source node and several topological nearest nodes in the group. In addition, anycast routing protocols are designed on the basis of unicast technique. Thus, GSR is presented as a particular case of unicast in this section. In GSR, once intermediate node receives a route request packet, the number of forwarding hops between source node and gateway is recorded in the header of each packet. Upon reception of these packets, the paths are stored in the gateway. The gateway then uses the reverse path as the backward path to a mesh node. Once receiving this packet, each node along the backward path updates the routing table. The packet is forwarded to the next-hop of a given backward path before it arrives at the destination node. This protocol has a few disadvantages. Firstly, traffic flow is unidirectional from a gateway toward a mesh node. Therefore, traffic from a mesh node to the gateway is disallowed. In addition, since no topology maintenance mechanism is provided, GSR may suffer from the packet loss problem due to broken links. Therefore, to improve the maintenance, authors propose an improved version of GSR, called Gateway Source Routing with Preferred Neighbour (GSR-PN) [37]. This algorithm selects the next-hop relay node with a preferred signal level. In GSR-PN, mesh nodes are divided into three groups. Nodes with signal level within the prefer range are defined as Preferred Neighbour group (PN). Nodes with signal level stronger than the prefer range are defined as In Group (IN). All other nodes with weaker signal than prefer level are belonged to Out Group (OG). In the propagation of packets, PN nodes prefer to be selected as relay node. If there are no nodes of PN, then OG nodes are chosen. Otherwise, IN nodes are selected to rebroadcast the packets.

Protocol/ Feature	AODV	DSR	DSDV	HMCN routing	BAAR
Routing Category	Reactive	Reactive	Proactive	Hybrid	Hybrid
Periodic Update	No	No	Yes	Yes	Yes
Flood control	Yes	Yes	No	No	Yes
Gateway involvement	No	No	No	Yes	Yes
Routing metric	Hop-count	Hop-count	Hop-count	Hop-count	Hop-count

Table 2.4: A comparison of different unicast algorithms in MANETs and MCNs

Protocol/ Feature	WMR	HWMP	AODV-CGA	GSR	GSR-PN
Routing Category	Hybrid	Hybrid	Reactive	Proactive	Proactive
Periodic Update	Yes	Yes	No	Yes	Yes
Flood control	Yes	Yes	No	No	No
Gateway involvement	No	Yes	Yes	No	No
Route packet from mesh node to Internet	Yes	Yes	Yes	Yes	Yes
Route packet from Internet to mesh node	Yes	Yes	Yes	No	No
Routing metric	Bandwidth, Delay	Airtime	Hop-count	Temperature	Received signal power

Table 2.5: A comparison of different unicast algorithms in WMNs

2.3.4 Conclusion of Unicast

Based on the different adaptive strategies, unicast algorithms can be divided into three routing schemes: *reactive routing (on-demand routing), proactive routing (table-driven routing)* and *hybrid routing*. In reactive routing, a path is calculated when needed. Therefore, the route discovery process is normally initiated by the source node in order to avoid the overhead of periodic information exchange. Since the path discovery may cause heavy traffic load in large scale networks, reactive routing is more suitable for small or medium size network. All of AODV, DSR, GSR, GSR-PN, and AODV-CGA are reactive routing algorithms. Unlike reactive routing, proactive routing calculates paths prior to sending traffic. Therefore, a node sends out the link changes to all the nodes. Each node stores the up-to-date routing information in the routing table. For example, DSDV is a proactive routing algorithm.

Proactive routing is suitable for network with relatively static nodes. In contrast, reactive routing provides the discovery of paths in the networks with dynamic changing topology, in which nodes have high speed. As a combination, hybrid routing has the traits of both proactive routing and reactive routing. In hybrid routing algorithms, the network topology or local topology can be proactively maintained for each period of time. The source can request the path on demand from either central administrator (gateway) or local administrator (cluster head). For example, HMCN routing, BAAR and HWMP are all hybrid routing algorithms. To conclude, hybrid routing is the most suitable routing strategy for WMNs. Both gateway involvement nature and low node mobility nature of mesh networks can be easily applied in the design of hybrid routing algorithm. Table 2.4 and Table 2.5 show comparisons of different routing algorithms in MANET, MCN and WMN.

Routing algorithms for MANET, MCN and WMN are shown in this section. There are limitations of existing routing algorithms such as adaptation problem of MANET and MCN routing algorithms, and the design problem of existing mesh routing algorithms. Due to above reasons, this thesis aims to design new hybrid unicast routing algorithm for WMN considering the key role of the gateway in managing the routing paths.

2.4 Multicast Routing

Multicast is a key technology for a group communication. It is used to send information to multiple nodes in the network. Packets are delivered over each link only once, and copied in the replicator nodes. In contrast, the source node should send packets to multiple receivers for multiple transmissions by unicast for the same result. Multicast can reduce the communication cost, consequently, conserve bandwidth, and reduce latency and network congestion. It overcomes the shortcoming of sharing the same wireless channel as well as the bandwidth scarcity condition of many applications and services, such as service discovery, video conferencing, distributed gaming, etc. Therefore, multicast technology is applicable for future wireless multi-hop networks to provide efficient data communication among a group of nodes. However, existing multicast protocols for wireless multi-hop networks [45] [46] [55] [95] cannot be applied to WMNs for efficient multicasting. These multicast algorithms are primarily designed to be suitable for energy-constrained mobile nodes. In contrast, besides energy and mobility issues, multicast algorithms are required to fulfil the new characteristics of WMNs, such as gateway involvement, low node mobility etc, so as to provide high communication quality to end users.

In the followings, multicast algorithms for multi-hop wireless networks are reviewed. First, multicast algorithms in MANETs are represented in 2.4.1. Then multicast algorithms in WMNs are described in 2.4.2.

2.4.1 Multicast Routing Algorithms in MANETs

In MANETs, multicast technology has been studied in both industry and academia for more than a long while. Since mobility issue and energy consumption issue are most concerned in MANETs, most proposed multicast algorithms operate reactively. They aim to capture the link status of mobile nodes and eliminate the energy consumption caused by handling proactive message exchange. In the following paragraphs, two reactive multicast algorithms for MANETs are described and compared.

2.4.1.1 Multicast Ad-hoc On-demand Distance Vector

As a multicast extension of AODV [43], Multicast Ad Hoc On-demand Distance Vector Routing protocol (MAODV) [45] is an on-demand routing protocol in MANETs. A Route Request (RREQ) is originated if a node intends to either join a multicast group or send a message without a valid path to a multicast group. Since MAODV is a tree based protocol, the structure is composed of all group members and message forwarders. Each multicast group is identified by a unique address and a group sequence number for tracing the freshness of each group. A node broadcasts RREQ with join flag (RREQ-JOIN) messages either if it wishes to find a path to the multicast tree or it aims to become a member of the multicast group. Any node received this message may respond if it has a valid path (based on group sequence number) to the multicast group. If a node is on the tree or it has a path to the tree, it sends a Route Reply message (RREP) back to the source node unicastly. If a node intends to join a non-existent multicast group, this node becomes the leader of that multicast group. Then it is in-charge of the group. As the source node receives multiple RREPs after a waiting period, it may select the best path. To maintain the multicast group, Group Hello message is used to establish the multicast group and update the routing table. Both the unicast routing table and multicast routing table are stored for the group tree structure. This multicast routing table contains the multicast group address, the multicast group leader address, the multicast group sequence number, the number of hops to the multicast group leader, next-hop information and lifetime. There are two types of nodes in a tree structure: *downstream nodes* (more hops from the group leader) and *upstream nodes* (less hops from the group leader). Obviously, a group leader has only downstream nodes. When a member leaves the group, the pruning process is initiated to reconstruct the tree structure. When a link is broken in the multicast tree, the furthest downstream nodes send RREQ-JOIN messages to initiate the repair process. Moreover, a member node notifies the group leader if it intends to terminate the membership.

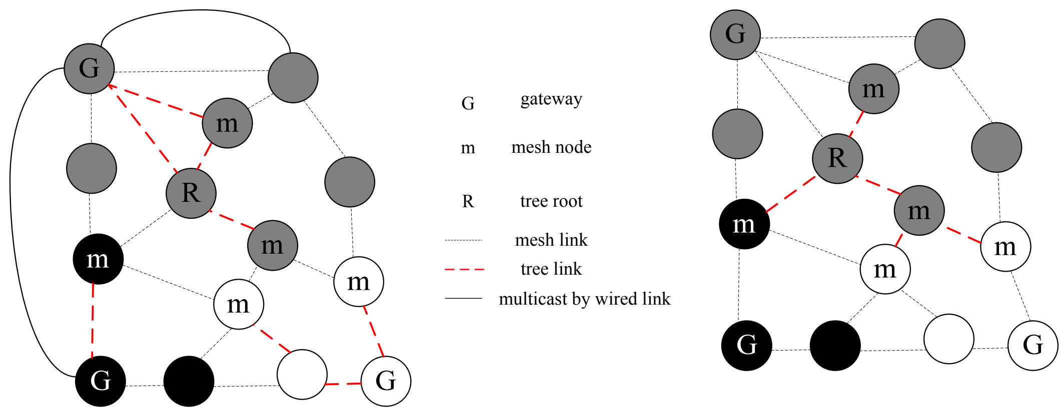
2.4.1.2 On-Demand Multicast Routing Protocol

On-Demand Multicast Routing Protocol (ODMRP) [46] is also an on-demand routing protocol for MANETs. In this protocol, Join Query messages are broadcasted if a mobile node intends to send packets but without a valid path to the multicast group. When receiving Join Query, each node stores the appropriate node id of the message sender and the reverse path back to the sender in the routing table. If the TTL (time-to-live) value is greater than zero, the intermediate node rebroadcasts the message. A group member broadcasts a Join Reply message if it receives a Join Query message. Upon reception of a Join Reply message, a neighbourhood node checks the join reply table to find out whether there is an existing next-hop node with the same source id. If so, the current node is set as a member of the forwarding group indicating it is on the path to source. As a mesh based protocol, there are multiple paths from sender to each receiver in ODMRP. In contrast, MAODV is a tree based protocol with only one path to each receiver. In other words, ODMRP allows a reply message back to the source node via multiple paths. These paths are stored in the source node for the future link breakage. Unlike MAODV, ODMRP is a soft state protocol. Member node can leave the group without a control message.

2.4.2 Multicast Routing Algorithms in Wireless Mesh Networks

Compared to MANETs, most nodes are stationary and energy-efficient in WMNs. Therefore, existing multicast algorithms for MANETs cannot be applied to WMNs directly without any modification due to the different characteristics between MANETs and WMNs. To satisfy the new design requirements, several multicast algorithms are proposed for WMNs. For example, LCA and MCM [47] are proposed to solve the problem of multicast channel assignment in WMNs. In the following paragraphs, a list of multicast algorithms for WMNs are described and compared.

In [96], Ruiz et al. introduce the prefix continuity and the construction of Steiner tree [97]. These mechanisms enhance efficient multicast routing in WMNs. First, a discovery mechanism is applied to create default multicast paths towards the gateway on the basis of prefix continuity. Each mesh node is assigned the same prefix to existing gateway only by service provider of WMN. This guarantees a default path between a mesh node and an existing gateway with the same prefix. A mesh node without a prefix does not have the authority to register with the gateway. Similarly, to start a multicast session, source node broadcasts route request messages towards the root node. Once receiving this request, root node replies and builds the minimum spanning tree. Figure 2.8 show (a) multicast tree is constructed on the basis of prefix continuity and (b) multicast tree is not constructed on the basis of prefix continuity. These examples are represented as either a shared tree or a source-rooted tree. The tree root can be either some sort of central point or source of the group. However, this work does not provide the maintenance scheme of multicast tree. This may lead to problems during the link breakage of multicast tree.



(a) with prefix continuity (b) without prefix continuity

Figure 2.8: Multicast tree constructed with or without prefix continuity

In [51], authors propose a multicast protocol, namely, Probabilistically Reliable On-demand multicast protocol (PROD). PROD applies a link quality based routing metric to detect the link cost, Expected Multicast Transmissions (EMT) (this routing metric is described in detail in section 2.5.4). In PROD, all disjointed source-receiver pairs are connected in shortest path to form a temporary multicast tree. EMT measures the link cost to construct a minimal EMT tree as the Reliable Probabilistic multicast tree. During the tree construction, each node broadcasts periodic probe messages to all of its neighbours. It aims to obtain the packet deliver rate of each link to calculate EMT. PROD is a receiver-initiated multicast routing algorithm. In the multicast group, each receiver initiates the path discover procedure. It sends out JoinReq messages to the multicast tree with a minimised EMT value. Each JoinReq packet includes multicast group address, node address, sequence number, time-to-live, neighbour link quality table and link cost. In the link quality table of neighbour, the link quality of all wireless links connected to this node is stored. This allows the neighbours to obtain the backward link quality. In JoinReq packet, the path cost field is initially set to zero. Each forwarder node of JoinReq records the additional transmissions number of establishing the path to the multicast tree. This number is then stored to the field of path cost. The current multicast tree members are responsible to reply a JoinReply packet as receiving JoinReq packet. Upon reception of JoinReply messages, the joining node selects one path with the minimal cost to the multicast session among multiple replies. Then it sends a RouteActivate packet unicastly to the source of JoinReply to complete the multicast join procedure. Moreover, PROD is compared with ODMRP. It was shown that PROD reduces the number of forwarder nodes [48].

In [49], Shortest Path Tree (SPT), Minimum Cost Tree (MST) and Minimum Number of transmissions Tree (MNT), are described and compared as multicast routing schemes of WMNs. In the SPT algorithms, a tree is rooted at the multicast source and spanned over all the multicast receivers. The distance/cost between the source and each receiver along the tree is minimised. Compared to the SPT algorithms, the MST algorithms try to minimise the overall cost of multicast tree. The delivery of a data packet to any neighbours is only required a single data transmission due to the broadcast nature of wireless network. MST has a minimum number of multicast forwarder nodes. It is primarily designed to minimise the number of transmissions, rather than to minimise edge cost. In contrast, MNT generates the lowest number of transmissions than SPT, but with longer path length. The performance analysis between SPT, MST and MNT shows SPT has the outstanding performance. Therefore, author recommends the SPT approach for multicast routing in WMNs. However, as the only drawback of SPT, the experimental results show that SPT causes more packet losses than MST. It works well especially for large multicast group size and high multicast sending rate.

There are other proposals [47] [95] about channel assignment for multicast in WMNs. For example, Zeng et al. [47] propose a Level Channel Assignment (LCA) algorithm and a Multi-Channel Multicast (MCM) algorithm. These algorithms provide the optimisation of throughput for multi-channel multi-interface WMNs. By exploiting the multi-channel nature of WMNs, two network interfaces are implemented in each node. In the multicast session, one interface (RI) is used to receive packets for the upper level data flow, when another one (SI) is used to send packets to lower level nodes. In addition, MCM aims to minimise both the number of relay nodes and the total number of hops in the multicast tree. Consequently, the interference is reduced and the throughput of multicast is improved.

Protocol/ Feature	MAODV	ODMRP	Prefix Multicast	PROD	LCA MCM
Route discovery	Reactive	Reactive	Reactive	Reactive	Reactive
Multicast type	Tree based	Mesh based	Tree based	Tree based	Tree / Mesh
Routing metric	Hop-count	Hop-count	Hop-count	EMT	Hop-count
Reply to source	Unicast	Multicast	Unicast	Unicast	Unicast
Member maintenance	Hard-state	Soft-state	NA	NA	Solf-state
Group maintenance	Hello message	Hello message	Prefix continuity	RouteRepair message	Hello message
Gateway Involvement (request handling)	No	No	Yes	Yes (only for first joiner)	Yes
Gateway Involvement (construct tree/mesh)	No	No	No	No	Yes
Gateway Involvement (data forwarding)	No	No	Yes	No	Yes
Multichannel assignment	No	No	No	No	Yes

Table 2.6: A comparison of different multicast algorithms in MANETs and WMNs

2.4.3 Conclusion of Multicast

To investigate the multicast technologies, two multicast routing algorithms MAODV [45], ODMRP [46] for MANETs are described in Section 2.4.1. Then, the current state-of-the-art multicast routing algorithms for WMNs are studied in this chapter. A comparison of different multicast algorithms is shown in Table 2.6. Since this thesis focuses on optimising the communication performance of WMNs, existing multicast routing algorithms are concluded on the basis of the content of this section as the followings:

- Existing multicast routing algorithm cannot be directly applied to WMNs due to the differentiated networking devices and network deployment structure. In MANETs, multicast algorithms are designed to suitable for the high mobility and energy-inefficient nodes. Therefore, in these algorithms the dynamically changing topology is discovered by broadcast. It may create overhead during the construction of multicast tree or mesh. On the contrary, in WMNs, the design of multicast algorithms should consider features such as the mobility level of mesh nodes and gateway involvement.
- Obviously, there is limited work on multicast routing algorithm for WMNs. Most of existing proposals are concentrated on channel assignment [47], or multicast tree construction [49]. These proposals have not clearly presented the key role of gateway in multicast communication for WMNs. In such proposals, WMNs are treated as flat network with fixed nodes. In other words, this kind of network is as same as MANETs with all stationary nodes. For example, authors abuse the term of WMN in the design of PROD [48] [51]. In contrast, a WMN is hierarchical in real world. Multicast packets are relayed by the gateways between different WMNs through the Internet backbone similar to unicast packets. The gateways are responsible for managing and relaying all the multicast data. For the above reasons, there is a need to develop new multicast routing algorithms for WMNs. These new design should

especially consider the gateway involvement to handle and process the routing request from the potential receivers and multicast sources.

Regarding the above points, this thesis focuses on the design of efficient multicast routing algorithms. These new algorithms consider either optimising load balancing problem or involving the gateway as an orchestrating routing node.

2.5 Routing Metrics

In the implementation of routing protocols, routing metrics are assigned to different paths. It calculates the cost of each path in order to select or predict the best path. This path is then stored in the routing table for future use. Routing metrics are integrated in routing protocols to improve communication quality in terms of bandwidth, error rate, latency, reliability, and cost. The design of routing metrics is important as the limited channel bandwidth in wireless communication. Existing routing metrics can be classified into following types: *distance, latency, traffic load, error rate, multiple-channel, channel usage* and *compositive metric*. The following sub-sections show typical examples of each type.

2.5.1 Distance Routing Metric

Hop-count is the most basic metric applied in existing protocols such as AODV [43], DSR [88], and DSDV [44]. A routing protocol with the Hop-count metric considers the number of hops between source and destination. Hence, it finds the path with the minimum distance. The advantage of Hop-count is that it does not generate any extra overhead with its self-detection mechanism. However, it does not consider other issues such as link quality, transmission rates. Since minimising the number of hops is not usually the performance goal in WMNs, Hop-count may result in poor performance.

2.5.2 Latency Routing Metrics

In [52], Per-hop Round Trip Time (RTT) is designed for Multi-Radio Unification protocol. It measures the round trip delay of unicast probes between neighbours. To implement this metric, each node sends out a probe packet with timestamp to all neighbours. When receiving the probe packet, each neighbour may respond with an acknowledgement. As the sender receives the acknowledgement, it calculates the round trip time between probe sent and acknowledgement received. The advantage of RTT is that the busy channel and the link loss are improved. However, the queue delay exists due to the contention among nodes for low RTT link. RTT also generates high overhead and self interference.

As an improved version of RTT, Per-hop Packet Pair Delay (PktPair or PP) [53] involves queue delay and transmission rate. In PktPair, a node sends out two probe packets to each neighbour every two seconds. In addition, the first probe packet is small (137 bits) and the second is large (1137 bits). This tests the sensitivity of link bandwidth for packets in different sizes. Once receiving probes, each neighbour calculates the delay difference of these two packets and reports the result to the sender. The sender also keeps the delay result of each neighbour for future routing. Although PktPair eliminates the problem of queue delay, it still suffers from the self interference and high overhead.

2.5.3 Load-aware Routing Metric

Load-count [54] is a load balancing metric for wireless multi-hop networks

$$Load-count=\sum_{i=1}^{n}Load_i \qquad (2.1)$$

where $Load_i$ is the traffic load of *node_i* on the path. The load is captured by IFQ (Interface Queue), which is a drop-tail buffer at the MAC layer of 802.11 radios. It contains the outbound frames to be transmitted in the physical layer. The size of IFQ is calculated as the number of the remaining packets in the buffer. The self-detection mechanism enables no probing overhead generated during the selection of the best Load-count path. In addition, it provides more stable paths than Hop-count, especially in the busy networking environments.

2.5.4 Error Rate Routing Metrics

Expected Transmission Count (ETX) [53] [56] is a metric to estimate the number of expected transmissions for the wireless links at the MAC layer. It measures the packet loss rate. A node sends out probe packets to all neighbour nodes every second. When a neighbour node receives a probe, it counts the number of received packets. Based on this information, it calculates the loss rate of packet every ten seconds. The weight of a path is the sum ETX of all links along this path. Therefore, the possibility of successful packet transmission from *a* to *b* on a wireless link is:

$$p=(1-p_f)\times(1-p_r) \qquad (2.2)$$

Then ETX can be achieved as

$$ETX=\sum_{k=1}^{\infty}kp^k(1-p)^{k-1}=\frac{1}{1-p} \qquad (2.3)$$

where p_f is the probability of successfully forwarding a packet; p_r denotes the probability of successfully receiving packets. The limitation of this metric is that it does not measure how data size and transmission rate affect the delivery rate. Furthermore, since ETX applies unicast probes to measure the error rate, it may operate inaccurately due to the broadcast nature of wireless.

Success Probability Product (SPP) [57] is a routing metric that tries to maximise the throughput in WMNs. Based on an energy-efficient routing metric [58], the authors develop SPP to predict the possibility of receiving a packet over a link as following:

$$SPP = \prod_{i=1}^{n} d_{f_i} \qquad (2.4)$$

$$d_{f_i} = 1 - p_{err_i} \qquad (2.5)$$

where d_{f_i} has been already mentioned above in ETX as the possibility of transmission. p_{err_i} is the error rate of link *i*. In this metric, higher SPP implies a good path. The advantages of SPP lie in its capability to generate low overhead due to using the broadcast operation and its suitability for path selection in multicast protocols. As it does not consider packet size and link bandwidth, SPP has the same problem as ETX.

Expected Transmission Time (ETT) [98] measures the MAC layer transmission time of a packet over a link *l*. It considers the impact of link transmission rate and packet size so as to improve the performance of ETX. The relation between ETT and ETX is shown as follow:

$$ETT_l = ETX_l \frac{s}{b_l} \qquad (2.6)$$

where *s* is the packet size and b_l is the bandwidth of link *l*. However, ETT still suffers from the inaccurate measurement of the unicast probing.

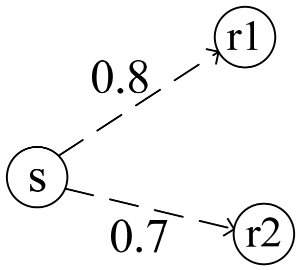


Figure 2.10: Example of EMT

Expected Multicast Transmissions (EMT) [51] is a routing metric to determine the link quality by capturing the loss rates of various links at the MAC layer. Due to the broadcast nature of wireless transmission, the Wireless Broadcast Advantage technology (WBA) presents one single transmission. It can potentially cover multiple neighbour nodes within the transmission range at a retransmission-based reliable MAC layer. With considering WBA, EMT is not the sum of individual unicast ETX values. EMT captures the link quality of multicast more accurately.

$$EMT_i = \sum_{c=1}^{|N_i|} (-1)^{c-1} \sum_{i=1}^{n} \frac{1}{1 - \prod_{j \in S} f_{i,j}} \qquad (2.7)$$

where $f_{i,j}$ denotes the link loss ratio given by $f_{i,j}$=1- $d_{i,j}$. $d_{i,j}$ is the link delivery rate from a source i to a recipient j in a set of nodes, N_i. This routing metric reduces the probing overhead by applying the WBA technology especially for multicast protocols. However, compared to ETX, EMT uses WBA in the proposed routing protocol to eliminate the difference between multicast and broadcast communications. Figure 2.10 shows an example of how EMT utilizes WBA. The EMT value of a multicast transmission between source *node_s* to recipient *node_r1* and recipient *node_r2* is 1.02. In contrast, the sum ETX of two unicast links is 2.68. As a consequence, it builds the minimal EMT tree (discussed in Section 2.4.2).

Airtime link metric [50] determines the routing prospect of each paired nodes. It is defined as the amount of channel resources consumed by transmitting the frame over a particular link. The Airtime link metric of a link is

$$C_a = [O + \frac{B_t}{r}]\frac{1}{1-e_f} \tag{2.8}$$

where O is the variation depends on the channel access overhead in the physical layer, including frame headers, training sequences, access protocol frames, etc. B_t is size of test frame. e_f denotes the frame error rate which is the possibility of transmission error on B_t data size packet at the bit rate r. The main disadvantage of this metric is it generates high probing overhead.

2.5.5 Multi-channel Routing Metric

Weighted Cumulative ETT (WCETT) [98] is proposed by Draves et al. With considering the multi-radio nature of WMNs, two components are added as total transmission time along all hops of a network and channel diversity of selected path. The WCETT of a path p is

$$WCETT(r) = (1-p)ETT_l + p\max_{1\le j\le k} X_j \tag{2.9}$$

where X_j is the number of times that channel j used by path r. p is a parameter as $0 \le p \le 1$. Therefore, $p\max_{1\le j\le k} X_j$ denotes the maximum number of times that the same channel j is occupied along a path. However, a problem of WCETT is that traffic flows may be routed to the dense area. Another important problem of WCETT is that it may generate a forwarding loop during the selection of the best path.

2.5.6 Channel Usage Routing Metric

Interference-Aware Routing Metric (IAR) [59] detects the channel busyness level by capturing the MAC layer information. IAR of a link is

$$IAR(l) = \frac{1}{1-a_{ub}} \times \frac{S}{B} \tag{2.10}$$

$$a_{ub} = \frac{T_{Wait} + T_{Collision} + T_{Backoff}}{T_{Wait} + T_{Collision} + T_{Backoff} + T_{Success}} \tag{2.11}$$

where T_{Wait}, $T_{Collision}$, $T_{Backoff}$, $T_{Success}$ are the time spent in *Wait state*, *Collision state*, *Backoff state* and *Success state* during a packet transmission, respectively. The time durations are captured in the MAC layer. a_{ub} is the percentage of time spent in the *Wait state*, *Collision state* and *Backoff state* compared to the time of completing a transmission. Therefore, smaller IAR presents a path with low traffic.

2.5.7 Compositive Routing Metric

Weighted Cumulative ETT with Load Balancing (WCETT-LB) [60] is a metric proposed by Ma et al. In WCETT-LB, two components are provided to optimise load balancing, congestion level and traffic concentration level at each node on a path. It can be shown as

$$WCETT-LB(r) = WCETT(r) + \sum_{i=1}^{n} \frac{Load_i}{b} + \min(ETT)N_i \tag{2.12}$$

where N_i is the set of children nodes using *node_i* as next-hop node in paths. min(*ETT*) presents the smallest ETT in the network. This routing metric is the combination of WCETT, Load-count, and ETT additionally with considering the bandwidth of links. However, this metric may cause the channel busyness due to broadcasting the results to all the children nodes.

According to [36], based on the protocol layers of each metric working on, existing routing metrics can be classified into the following three types: *single performance parameter metric, single-protocol-layer metric for multiple performance parameters,* and *multi-protocol-layer metric for multiple performance parameters*. In this context, Hop-count and Load-count are network layer routing metrics. They count either the number of hops or the traffic load along the paths. Hence, they are single performance parameter metric. On the contrary, IAR is a multi-protocol-layer routing metric for multiple performance parameters. It considers both the link layer and the network layer to capture MAC handshake time, bandwidth and packet size. Besides above three routing metrics, all other routing metrics mentioned in this chapter are single-protocol-layer metrics for multiple performance parameters.

It is also possible to catalogue routing metrics in the implementation point of view to understand the different characteristics. In this context, routing metrics are divided into probe-exchange based metric and self-detection metric. In the probe-exchange based metrics, probes are sent normally in cluster, group or overall network to detect the routing status. This kind of routing metric normally creates high overhead. On the contrary, self-detection metrics reduce exchange overhead by only measuring the local routing status. Besides Hop-count, Load-count and IAR, all the routing metrics in this chapter are probe-exchange based metrics.

Feature/ Routing Metrics	Layer	Communication quality parameters	Multichannel	Power management
Hop-count	Network	Number of hops	No	No
RTT	Network	Packet loss, delay, contention	No	No
PktPair	Network	Packet loss, delay, contention	No	No
Load-count	Network	Traffic load	No	No
ETX	Network	Packet loss, retransmission, contention	No	No
SPP	Network	Same as ETX	No	Yes
EMT	Network	Same as ETX	No	Yes
WCETT	Network	Same as ETX, plus bandwidth and packet size	Yes	No
ETT	Network	Same as ETT	No	No
IAR	Network , Link	MAC handshake, time, bandwidth and packet size	No	Yes
Airtime	Link	Resource consumed by a packet on a link	No	Yes
WCETT-LB	Network	Same as WCETT, plus traffic load, bandwidth	Yes	No

Table 2.7: A comparison of different routing metrics for WMNs

2.5.8 Conclusion of Routing Metric

According to Table 2.7, there are still several remaining issues in the design of routing metrics for WMNs. First of all, many existing routing metrics still work ad–hocly. Consequently, they may only perform well for a certain type of WMN such as Client WMN. Second of all, some routing metrics measure routing status inaccurately. For example, ETX abuses the broadcast nature of wireless communication as it applies probe unicast to measure the error rate of transmission. Third of all, probe-exchange based metrics may cause high overhead. Therefore, self-detection or limited probes should be considered in the future design. Fourth of all, limited network parameters are considered. Critical parameters such as traffic load of neighbouring nodes and QoS for diverse applications are not captured in existing routing metrics. These points are taken into consideration when designing new metric in this thesis.

2.6 Conclusion

In this chapter, we first review the architectures and characteristics of Wireless Mesh Network. Compared to Mobile Ad-hoc Network and Multi-hop Cellular Network, a WMN is distinguished for its new features which are suitable for providing network access to all kind of mesh-enabled devices. Especially, a community mesh network can connect multiple home and enterprise networks together to offer cheap wireless broadband to consumers. Secondly, both load balancing and QoS are investigated as key elements to improve the stability of WMNs. Then, in the rest of chapter, unicast algorithms, multicast algorithms and routing metrics are reviewed. As a result, we observe that the current study of unicast algorithms, multicast algorithms and routing metrics lack of considering the new characteristics of WMNs. Therefore, this research aims to design new algorithms for unicast, multicast and routing metric to fulfil the efficient communication demand of real-time applications especially within community mesh networks.

In the next chapter, a gateway centralized unicast algorithm is introduced.

Chapter 3

Gateway Centralized Multi-hop Routing for Wireless Mesh Networks

3.1 introduction

Unicast, in which a source node sends messages to a single destination in the network, is one of the most useful communication operations in wireless networks. It has been studied in wireless multi-hop networks including cellular networks [69] [70] [71] [76], MANETs [43] [44] [88] and WMNs [50] [87] [93] [94] [99] [100].

In a WMN, a collection of wireless nodes communicate with each other, where a gateway is deployed as a server. As described in the preceding chapter, there are three kinds of main unicast routing approaches for wireless multi-hop networks including reactive routing (on-demand routing) [43] [88] [93], proactive routing (table-driven routing) [44] [94] [99] and hybrid routing [50] [69] [70] [71] [76] [91] [92] [100] [101]. Due to the use of gateway and the mobility nature of WMNs, hybrid routing algorithms [50] [68] show the distinguished performance against reactive routing and proactive routing algorithms.

For hybrid routing, a source node starts the path discovers reactively in MANETs. Routing information is exchanged among nodes within the range of group or network. The received routing information is then stored in the routing table for handling the future routing requests. Unfortunately, most of hybrid routing algorithms rely on the proactive probing exchange of routing information among nodes. This may increase the traffic overhead and degrade the communication quality. Therefore, this chapter presents a new hybrid unicast routing protocol for WMNs, namely, Gateway-Centralized Multi-hop Routing protocol (GCMR). This algorithm enhances the capability of efficient services towards a wide range of applications. GCMR takes the routing capability of the gateway into account. It only requires a small number of nodes (leaf nodes) to send the routing information which keeps the traffic overhead to a minimum. Besides, a novel traffic prediction mechanism is provided to further reduce probing overhead in the mesh networks. This chapter shows and confirms GCMR is especially suitable for community mesh networks. The reminder of this chapter is organised as follows. Section 3.2 shows the background of Hybrid routing. Section 3.3 describes the preliminaries and the proposed algorithm. Then, Section 3.4 presents the simulation setup and results. Finally, Section 3.5 concludes the chapter.

3.2 Background

In MANETs, existing hybrid routing algorithms such as Zone Routing Protocol (ZRP) [102] generate high traffic overhead. This kind of protocols is required to be reengineered to satisfy the requirements of WMNs. As shown in Table 2.4 and 2.5, in existing hybrid routing algorithms [50] [91] [92], gateways are deployed only for forwarding requests and data packets. These algorithms do not take advantage of gateway involvement to implement centralized routing.

Centralized networking is primitively studied in wired network, where all users connect to a central server which manages all communications within the network. In such architecture, the server keeps both the user information and the communication information. The gateway is normally deployed in the topological centre of a WMN. Unlike distributive route discovery in MANETs, mesh gateway can work as a server to manage communications inside the network. Therefore, the path calculation in gateway should be taken into account to perform centralized routing for WMNs. Despite some proposals [71] [76] [101] which have implemented the gateway centralized routing in cellular networks, this issue is rarely studied within the context of WMN. In fact, most existing work such as [93] [103] has omitted the role of gateway nodes. None of them utilise the gateway as an orchestrating routing node at the provider side to control the message transmissions for WMNs. In these works, all nodes are treated equally as in MANETs, which is not the case in WMNs.

3.3 Gateway-Centralized Multi-hop Routing (GCMR)

In this chapter, GCMR is proposed to deploy the central control mechanism of network resources in gateways. GCMR aims to improve the management of network layer so as to achieve high fairness of sharing network resources in WMNs. In addition, the gateway has knowledge about all the dynamical changes within the network. Based on the knowledge, it computes all the routing paths for clients proactively. One of the most innovative features of GCMR is the achievement of better load balancing, especial in heavily-loaded WMNs. In what follows, in Section 3.3.1, some basic conceptions are defined to improve the readability and presentation of our algorithm. Then both registration and routing procedure of GCMR are explained in Section 3.3.2. Moreover, Section 3.3.3 theoretically compares GCMR with

HWMP in the involvement of gateway, the communication quality control and the flood control.

3.3.1 Preliminaries

Definition 3.1: Given a WMN $G=(V,E)$, and a destination node, $d \in Đ$, as $Đ$ is set of all possible destinations, there is two main route discovery processes between υ and d, $\Re_{in}$ and $\Re_{out}$, inner and outer, respectively: $$\Re = \begin{cases} \Re_{in} & , if\ \upsilon \text{ and } d\ \in \mathrm{V} \\ \Re_{\mathrm{out}}, & \text{otherwise} \end{cases} \tag{3.1}$$

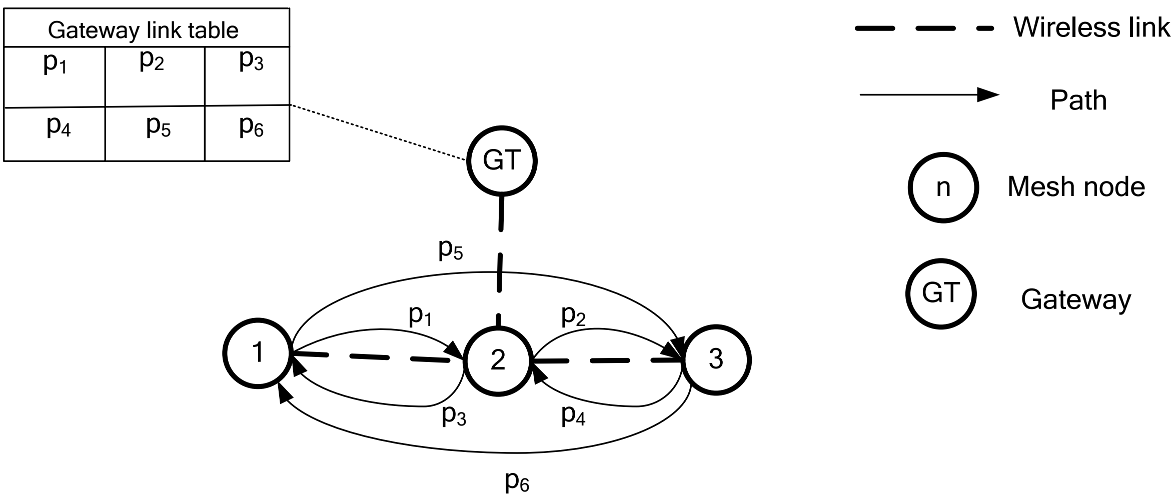


Figure 3.1: An example of paths between three mesh nodes and a gateway

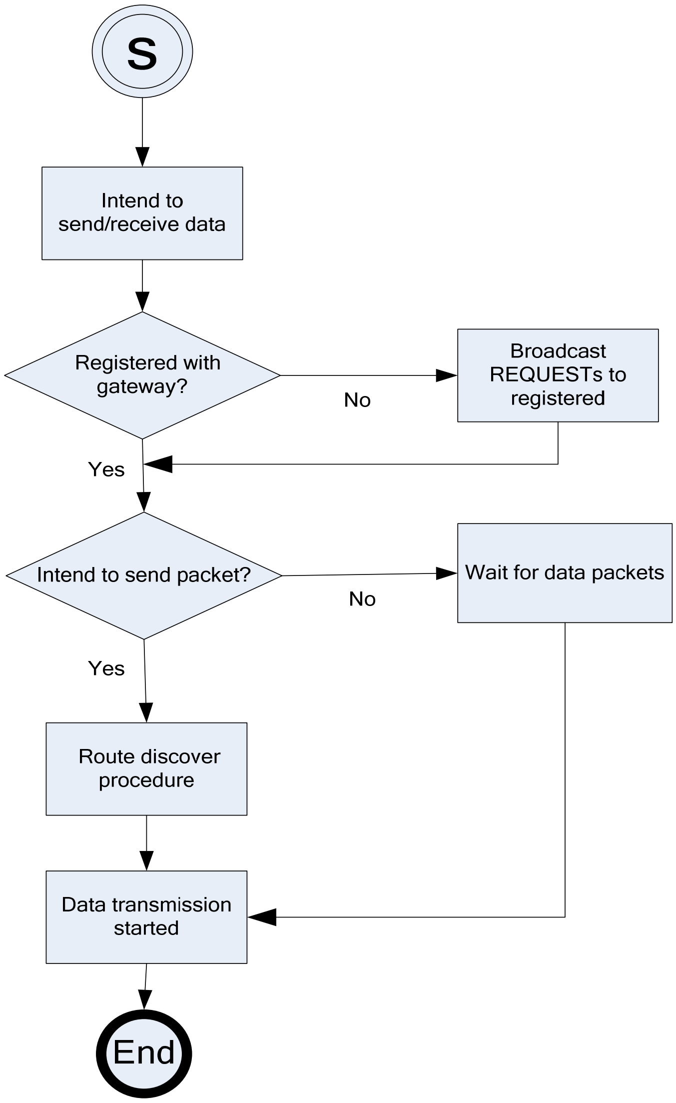


Figure 3.2: The routing process at the mesh node

***Definition 3.2**:* Consider a mesh network $G = (V, E)$, a mesh node, $\forall \upsilon \in V$, is said to be a leaf node if and only if it has only one one-hop neighbour, otherwise it is considered to be a non-leaf node that can be designated as a virtual leaf node.

3.3.2 Registration and Routing

As shown in Figure 3.2, each node intending to send or receive data must first register with the central gateway. In the registration process, a mesh node broadcasts REQUEST messages. When an intermediate node receives a REQUEST message, it checks whether it is included in the path address field of this message. If so, the message is discarded. Otherwise, the intermediate node adds the node *id* and the current communication quality to the packet header. Then this message is rebroadcasted once at least. REQUESTs are propagated along different paths towards the gateway. When a REQUEST arrives at the gateway, all possible links from this packet are recorded in the gateway link table. Then a REPLY message is sent back to the mesh nodes along the best-metric path. Once receiving REPLY message, each intermediate node updates the path to the gateway and forwards the message to the destination node.

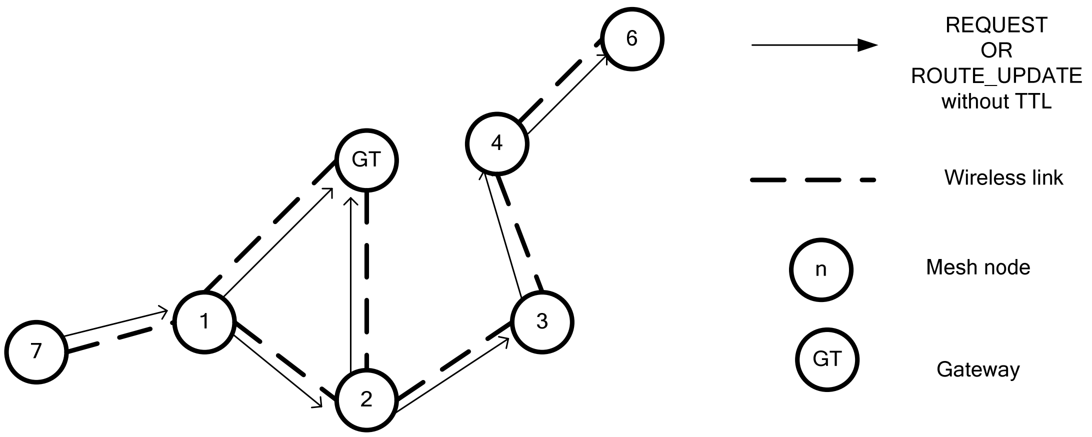


Figure 3.3: Register by REQUEST/ Update by Route discovery

The gateway link table is applied to store all the link information within the network. In the gateway link table, leaf nodes are used to broadcast periodic route update (ROUTE_UPDATE) messages towards the gateway. ROUTE_UPDATE messages flow along all the possible paths from the leaf node(s) towards the gateway. The communication quality of nodes along each path is also recorded in the update message. Finally, this mechanism guarantees that the gateway gathers the link information of the entire WMN with maintaining the gateway link table up-to-date. Figure 3.3 shows an example of propagation of ROUTE_UPDATE/REQUEST from *node_7*. Packets fly through every possible node to record non-duplicate node(s) with communication status until finally reaching the gateway. Once the gateway receives a ROUTE_UPDATE packet, all the passed links decapsulated from this packet are used to update the gateway link table. In some cases, a leaf node later may become a non-leaf node if another node moves closer to become its neighbour. Then the gateway may ask this node to stop sending ROUTE_UPDATE messages if, and only if, there are more than two leaf nodes. This ensures that there is always at least one leaf node to generate ROUTE_UPDATE messages. If a non-leaf node has not received any ROUTE_UPDATE for a period of time, it considers that either the leaf node moves away or it is not reachable by ROUTE_UPDATE messages from the existing leaf node(s). This node then broadcasts ROUTE_UPDATE-ERR until the gateway updates the gateway link table. Finally, if there is no node meeting the conditions of the leaf node, the gateway selects a virtual leaf node to generate ROUTE_UPDATE messages. A virtual leaf node is an end node with the biggest distance to the requested node. Instead of real leaf node, virtual leaf node generated ROUTE_UPDATE packets can pass through all possible paths across the error node. Compared to the updating mechanism of BAAR [71] [101], using the selected leaf nodes to update the real-time topology of the network can effectively reduce the traffic overhead. Figure 3.4 shows a WMN with two leaf nodes, namely, *node_6* and *node_7*. Hence, they are chosen to broadcast periodic ROUTE_UPDATE packets to the gateway GT.

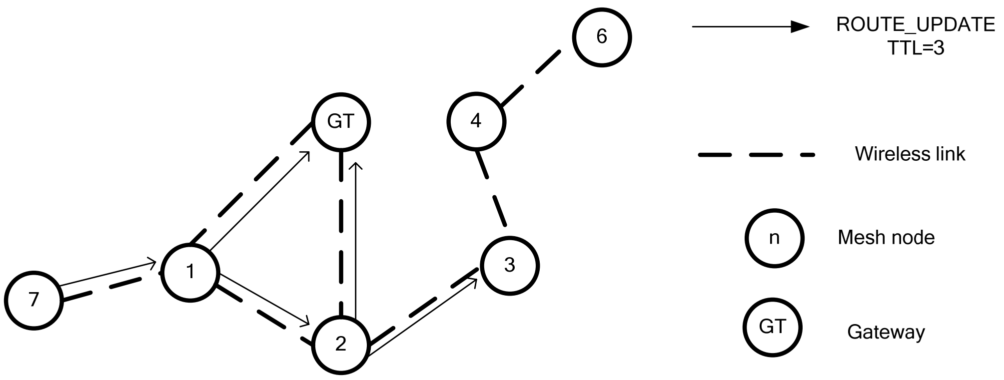


Figure 3.4: Leaf nodes with TTL setup

The traffic prediction method is implemented to limit the flooding of ROUTE_UPDATE messages. To reduce the frequently repeating update processes, this method is the widening of time gap between each two continuous ROUTE_UPDATEs. The time gap is extended to $t \times \Delta$, where Δ is a constant. During this time gap, the gateway estimates the future quality (communication quality) of the path from a mesh node *m* to destination *d* for a route request. It also updates the gateway link table before the destination node informs the source node of the path based on the estimation. The incremental value of the communication status on each node is obtained by:

$$Q\_Incr = Routing\_factor \times \frac{d_{infer}}{d_{node}} \qquad (3.2)$$

where *Routing_factor* denotes the impact value of the routing metric for a single transmission, d_{infer} denotes the average interference distance in the network. d_{node} denotes the average distance between two neighbouring nodes. *Routing_factor* is calculated as

$$Routing\_factor = \frac{\sum_{i=1}^{n} Q_i}{n} \qquad (3.3)$$

where Q_i represents the status based on the communication quality of *node_i* on a given path. When the gateway receives the route request with inner-communication flag

(ROUTE_REQUEST-I) message, it predicts the future link quality of the path as *Q_Incr* and then adds the *Q_Incr* value to the communication status of nodes on the selected path. Finally, the predicted communication status of selected link is stored in the gateway link table.

In order to reduce the probing overhead caused by the broadcasting ROUTE_UPDATE messages, the TTL (time-to-live) field of each message is set as follows:

$$\begin{cases} TTL = S_{node} & if\ S_l = 0, or\ S_l = 1 \\ TTL = \max(h^{\max}_{l->g}, \left\lceil \dfrac{h^{\max}_{l_i->l_j}}{2} \right\rceil) & if\ S_l > 1 \end{cases} \qquad (3.4)$$

where l and g refer to the leaf node and the gateway, respectively. As $_i$ and $_j$ are two leaf nodes in a network, the hop number between l and g along the longest path is expressed by $h^{\max}_{l->g}$, whereas $h^{\max}_{l_i->l_j}$ denotes the hop number of the longest path from l_i to l_j. In addition, l_i is the nearest leaf node to l_j. S_{node} indicates the number of mesh nodes. S_l denotes the number of leaf nodes in the network. This mechanism minimises the probing traffic caused by the ROUTE_UPDATE messages and consequently avoids the duplicated update coverage areas in WMNs. In fact, without adopting such a mechanism, one non-leaf node can be affected by the update messages generated from different leaf nodes. Figure 3.3 shows *node_7* generates ROUTE_UPDATE without TTL setup, while the packets propagate along all the nodes in this case. Compared to Figure 3.3 (without TTL setup), a significant reduction in the number of probes is presented in Figure 3.4 (with TTL setup).

When both of sender and receiver are in the same WMN, GCMR should deal with the routing request as the inner WMN process. If a sender *node_i* intends to send packets to a receiver *node_k* without a valid path, a path discovery process is initiated. After registering with a gateway, *node_i* sends ROUTE_REQUEST-I towards the gateway to request a path to *node_k*. The gateway node implements the Dijkstra's algorithm [104] to calculate the best-metric path. When ROUTE_REQUEST-I is received by the gateway, the best-metric path from *node_i* to *node_k* and the best-metric path from the gateway towards *node_k* are calculated respectively based on the gateway link table. A route request with outer-communication flag (ROUTE_REQUEST-O) message is sent towards *node_k* only if the gateway cannot find a path from *node_i* to *node_k*. In this case, *node_k* is assumed in an outer WMN, i.e. it belongs to the outer route discovery $\Re_{out}$ as described in Definition 3.1. After the request is propagated to *node_k, node_k* responds with ROUTE_REPLY to *node_i* along this path. Once *node_i* receives the reply, it starts to send data packets along the reverse route in ROUTE_REPLY. In case, if the mesh node does not hear any reply from the destination for a certain period, it assumes there is a broken link in the shortest path to the gateway. *Node_i* then broadcasts route request with error flag (ROUTE_REQUEST-ERR) messages towards the gateway to find an alternate path.

Instead of Hop-count, Load-count [54] is used as the routing metric in the process of selecting the best path. Load-count determines the communication quality by detecting the current traffic load of a node. Load-count of a path is calculated as the number of remaining packets in the packet buffer. This is primarily due to two reasons. Firstly, it is a widely used routing metric to optimise load balancing in WMNs. Secondly, compared to ETX [53] [56] and ETT [98], Load-count is a self-detection metric which does not generate extra traffic from exchanging the information.

Algorithm 3.1: Pseudo code of Gateway-Centralized Route Discover algorithm for $\Re_{in}$

Input: mesh node *m*, gateway *g*, intermediate node *f*, destination node *d*

Output: best-metric path *m->d*

```
m.send(ROUTE_REQUEST-I) to g
if f recvd ROUTE_REQUEST-I
 if f!=g and f!=d
 f.forward(ROUTE_REQUEST-I)
 else
 ROUTE_REQUEST-I.path_m_d=g.find_best _path(m->d)
 ROUTE_REQUEST-I.path_g_d=g.find_best _path(g->d)
 g.send(ROUTE_REQUEST-I) to d
 endif
endif
if f recvd ROUTE_REQUEST-I
 if f!=d
  f.forward(ROUTE_REQUEST-I)
 else
 d.send(ROUTE_REPLY) to m
 endif
endif
if f recvd ROUTE_REPLY
 if f!=m
 f.forward(ROUTE_REPLY)
 else
 m.record(m->d)
 endif
endif
if m not recvd ROUTE_REPLY for a period t
 m.bcast(ROUTE_REQUEST-ERR)
endif
if f recvd ROUTE_REQUEST-ERR
 if (f!=g)
 f.foward(ROUTE_REQUEST-ERR)
 else
 g.update_links_table(REQUEST)
 g.send(ROUTE_REQUEST-I) to d
 endif
endif
```

If the source *node_i* and destination the *node_k* are in different WMNs, GCMR should handle the outer WMN process as follows, the gateway first broadcasts route request with gateway flag (ROUTE_REQUEST-G) messages through the Internet backbone. Once an intermediate gateway receives ROUTE_REQUEST-G, the message is rebroadcasted until arriving at the gateway with an entry of *node_k* in the gateway link table. This gateway adds the best-metric path from itself to *node_k* in the route request message with forward flag (ROUTE_REQUEST-F). It then sends the message towards *node_k*. When *node_k* receives this message, a route reply message (ROUTE_REPLY) is replied to the gateway of *node_i*. ROUTE_REPLY is sent along the reverse path from ROUTE_REQUEST-F. When this message is arrived at the gateway of *node_i*, the best-metric path between *node_i* and gateway is also added to the packet before forwarding to *node_i*. Finally as *node_i* receives this reply, data packets are sent along the backward path decapsulated from ROUTE_REPLY.

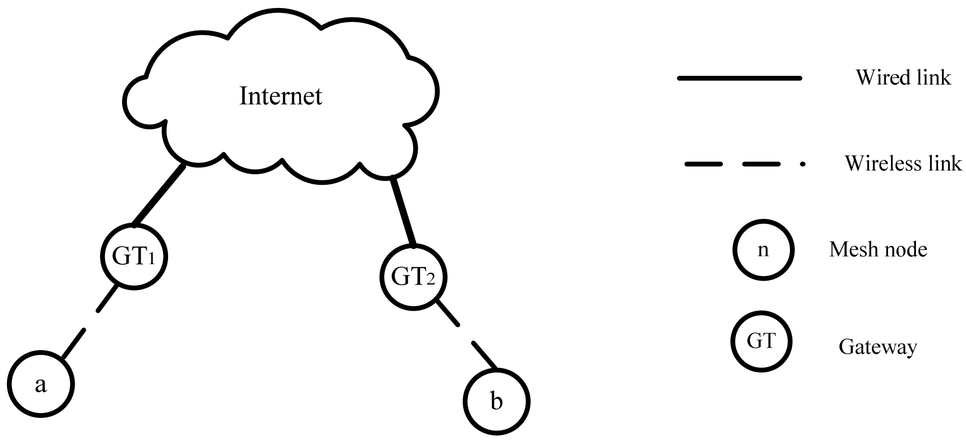


Figure 3.5: Routing among WMNs

Figure 3.5 illustrates an example of the path selection for the outer WMN process. When the source *node_a* intends to find a path to *node_b*, *node_a* first requests the best path from GT_1. Since the gateway link table of GT_1 does not contain any information about *node_b*, GT_1 assumes *node_b* is in another WMN. It then broadcasts the ROUTE_REQUEST-G in the backbone. When GT_2 receives this request, it finds *node_b* in the gateway link table and forwards the request to *node_b* with path $GT_1 \rightarrow backbone \rightarrow GT_2 \rightarrow node\_b$. Upon reception of this request, *node_b* uses the reverse path to send a reply towards *node_a* through the Internet backbone. While receiving this reply, GT_1 adds the best-metric path $GT_1 \rightarrow node\_a$ to the packet header before transmitting to *node_a* along this path. When *node_a* finally receives the reply, it stores the best-metric path to *node_b* which is $node\_a \rightarrow GT_1 \rightarrow backbone \rightarrow GT_2 \rightarrow node\_b$.

3.3.3 Novelty of the Proposed GCMR against HWMP

Compared to HWMP, GCMR has distinguished advantages which are outlined as follows that fit into the capacity of WMNs:

1) Gateway inclusion: both of GCMR and HWMP involve the gateway to relay the request. The gateway only forwards the request from a source node to a destination node in HWMP. It also selects a path from the cache. However, in contrast, the routes are selected on the basis of the link information in the gateway link table. This enables the central management of network routing in GCMR, which is essential for some applications that require control/management at the service provider side. In addition, gateway is required to handle the update packets in GCMR.
2) Communication quality control: although HWMP uses Airtime link metric to select the best radio-aware path, the cached paths in the relay node are affected due to the dynamic

changes of the network. Consequently, the gateway cannot find the best-metric paths all the time. In contrast, in GCMR, gateway chooses the best paths based on the up-to-date information of the network. In addition, most of existing routing metrics can be applied to examine the quality of links in the gateway.

3) Flood control: in HWMP, a node without a valid path to the destination broadcasts a route request over the network. It may lead to a heavy traffic overload caused by these additional flows. Compared to HWMP, GCMR applies the leaf-to-gateway update mechanism. This is a one-way link maintenance mechanism to assign a limited number of leaf nodes to generate update packets. In addition, a mesh node finds the gateway on demand in GCMR. In contrast, HWMP uses a two-ways link maintenance mechanism. In HWMP, the gateway presence is announced by flooding periodical Root Announcement (RANN) messages. Once receiving RANN, each internal node replies to the gateway. Moreover, the TTL setup mechanism further reduces the traffic flows caused by the periodic update messages.

3.4 Performance Evaluation

The performance of GCMR is simulated and examined using discrete event simulation environment created by the NS-2 simulator [64]. In the networking and communication field, NS-2 is popular used worldwide to simulate protocols and communication patterns such as unicast and multicast routing, especially for wireless networks. It contains a list of routing protocols and offers accurate and credible simulation results for both wired and wireless networks. In this chapter, the simulation aims to examine the performance of GCMR compared to AODV-HOP and HWMP with regards to the communication quality.

3.4.1 Simulation Model

As previous study in [14], the simulation first models a network of 45 wireless stationary routers and 5 wireless mobile routers for simulating relatively static community mesh networks. The nodes are distributed over a 1000 m × 1000 m area as shown in Table 3.1. In the second scenario, there are 100 nodes simulated over 2000 m × 2000 m area as shown in Table 3.1. The mesh networks in both grid topology and random topology are simulated. Sources send Constant Bit Rate traffic (CBR) over UDP as transport protocol, consisting of 1024-byte packets with a sending rate of 20 packets per second. The aim is to verify whether the algorithms are capable of driving the applications with high-bandwidth demand. Background traffic is also generated among the nodes to make the network busy enough. Each algorithm is simulated on 10 different grid topologies and 10 different random topologies. Each simulation runs for 400 seconds and the average results over all topologies are presented.

Network size	45 fixed nodes and 5 mobile nodes over a 1000 m × 1000 m area 90 fixed nodes and 10 mobile nodes over a 2000 m × 2000 m area
Gateway location	(500,500) for first scenario (1000,1000) for second scenario
Router transmission power	20 dBm
Interference range	500 m
Transmission rate at physical layer	54 Mbits/s
Physical layer protocol	PHY802.11g
Packet size (excluding header size)	1024 bytes
Queue size at wireless routers	50 Kbytes
Traffic model of sources	Constant bit rate (CBR)
Number of CBR senders	30 or 60
CBR sender's rate	20 packets/s
Simulation duration	400 seconds

Table 3.1: Simulation parameters

In the NS-2 simulation, GCMR-IFQ, GCMR-PRED, and HWMP are implemented. GCMR-IFQ denotes GCMR with Load-count routing metric. GCMR-PRED denotes that GCMR-IFQ additionally applies the traffic prediction method and the TTL setup mechanism. The results of GCMR-IFQ, GCMR-PRED, and HWMP are shown as normalised with respect to that of the original AODV-HOP.

3.4.2 Performance Metrics

Our experiments use three performance metrics: *average packet delivery ratio, average end-to-end delay*, *average jitter* and *probing overhead*. For classification, the average packet delivery ratio refers to the average percentage of packet delivery rate. The average end-to-end delay measures the average time that a packet travels from a source to a destination. The average jitter is also a metric to weigh the various packet delays. It examines the variation of the packet arrival times at destinations. In addition, the probing overhead counts all of the request packets in the discovering of the route as well as in the maintaining of the multicast tree.

3.4.3 Simulation Results

As Figure 3.6 shows, applying the traffic prediction method and the TTL setup mechanism helps GCMR to achieve the highest packet delivery ratio. The results illustrate that the packet loss of GCMR-IFQ is slightly higher than that of GCMR-PRED. However, since the reduction of flood caused by the topology update is concerned, both GCMR-PRED and GCMR-IFQ exhibit good performance compared to other routing protocols.

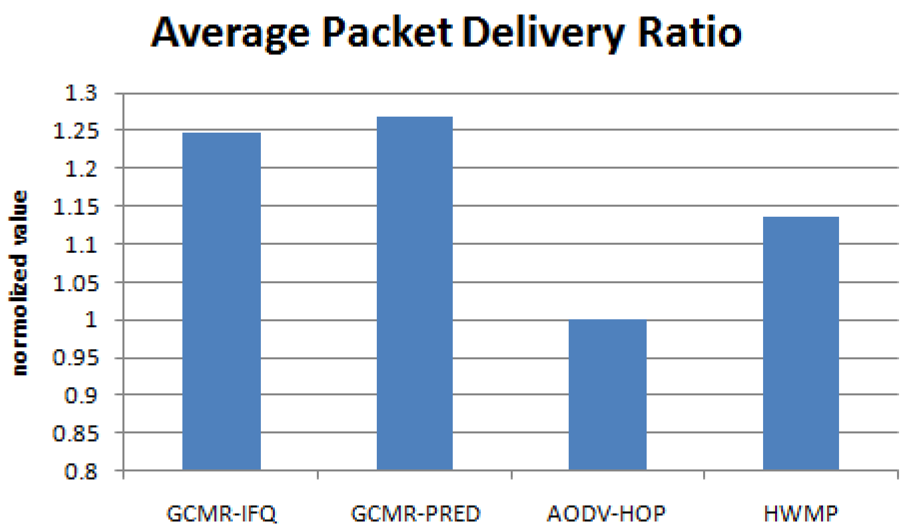


Figure 3.6: A comparison of packet delivery ratio of 50-nodes WMNs

In GCMR, the gateway maintains the up-to-date information of network topology. The optimal path is selected by using the advanced routing metrics. Compared to HWMP and AODV-HOP, the average end-to-end delay is hugely reduced in GCMR-PRED and GCMR-IFQ as shown in Figure 3.7. In addition, GCMR-PRED provides the lowest delay among the four routing algorithms with at most 680% less delay time than HWMP. The results indicate that the leaf-to-gateway update mechanism, the TTL setup mechanism and the traffic prediction method facilitating GCMR to avoid the busy path. On the contrary, the cache-and-forward mechanism is functionalized in HWMP. In such mechanism, after receiving the route reply from the destination node, each forwarder node stores the path for future use until the expiration time is reached. Consequently, nodes may not always keep the up-to-date best path. This yields outdated path that can be selected for transmitting data.

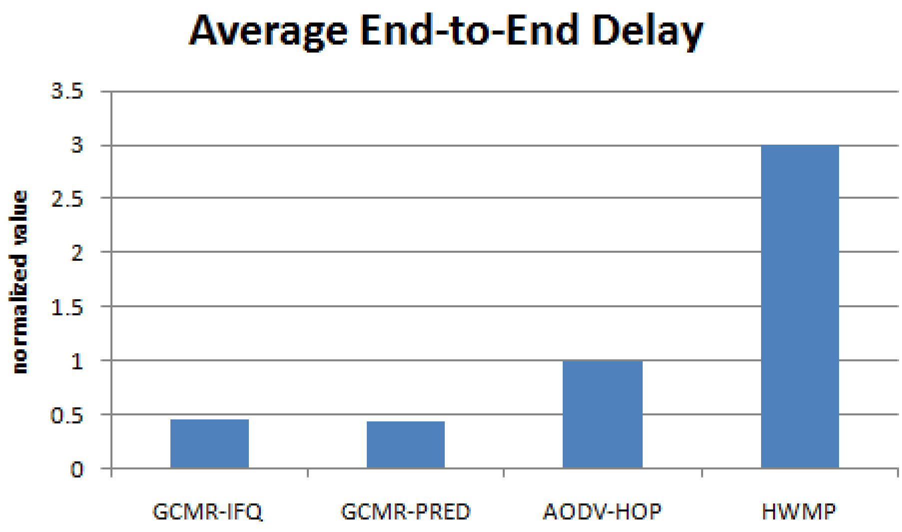


Figure 3.7: A comparison of average End-to-End delay of 50-nodes WMNs

A good path provides stable traffic relay services continuously. In other words, a poor path provides unsteady services in data packet transmission. To qualify for a good path, the instant communication is always highly necessitated in high-demand applications. A high average jitter is not tolerated in services such as gaming and online video conference etc. Figure 3.8 shows that GCMR-PRED exhibits the lowest jitter time, in which GCMR-IFQ is the second best. In both GCMR-PRED and GCMR-IFQ, the average jitter are reduced at least 40% than AODV-HOP and HWMP. All of the above results indicate that GCMR exhibits the highest communication quality. Apparently, GCMR is suitable for applications with strict communication needs in terms of average packet delivery rate, average end-to-end delay and average jitter.

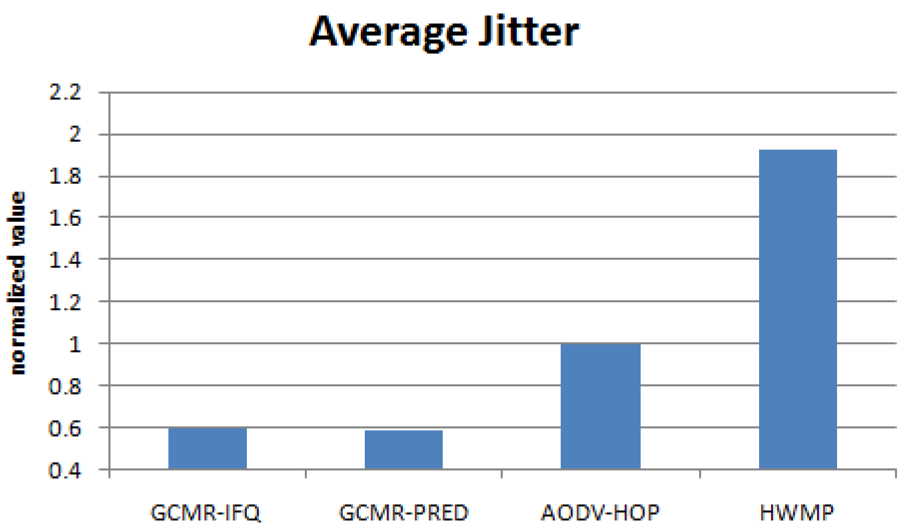


Figure 3.8: A comparison of Average jitter of 50-nodes WMNs

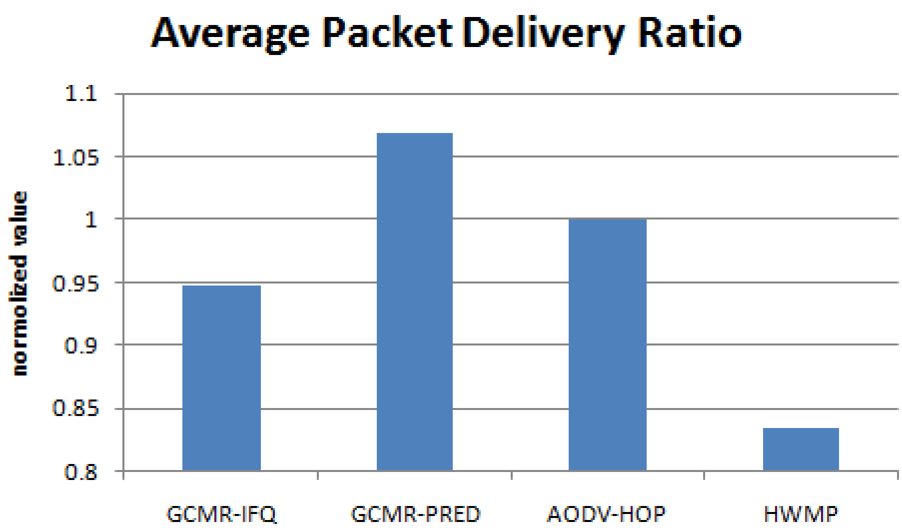


Figure 3.9: A comparison of packet delivery ratio of 100-nodes WMNs

In comparison, the results of 100-nodes WMNs scenario are shown in Figure 3.9-3.11. In Figure 3.9, GCMR with the traffic prediction method and the TTL setup mechanism achieves the highest packet delivery ratio, which is as same as the results in 50-nodes scenario. However, without these two methods, GCMR only performs the third place among these algorithms.

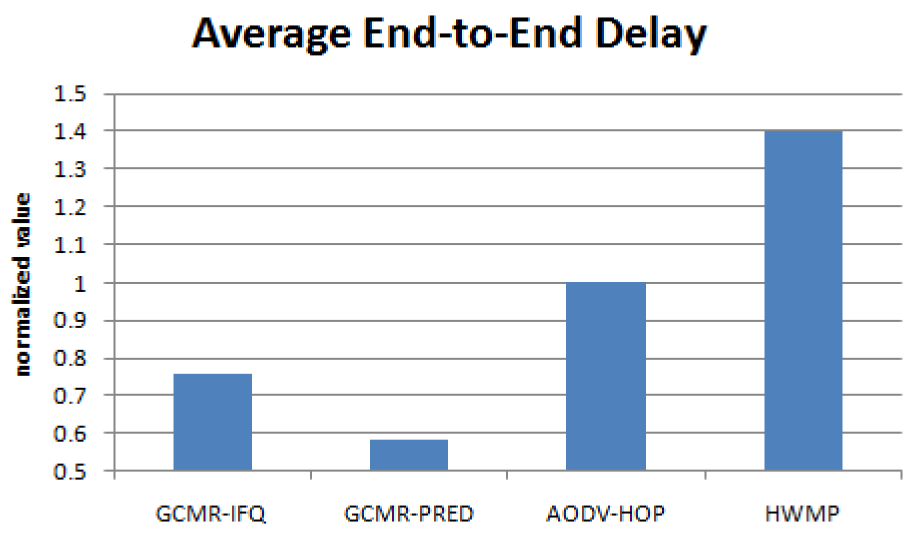


Figure 3.10: A comparison of average End-to-End delay for 100-nodes WMNs

In Figure 3.10, GCMR-PRED provides the lowest delay among these algorithms, where it achieves 42% lower latency than AODV-HOP does. Compared to the results of 50-nodes scenario, GCMR-IFQ provides only 25% lower delay than AODV-HOP. This means GCMR without the traffic prediction method and the TTL setup mechanism does not provide good performance for large scenario. The similar outcomes can be found out in Figure 3.11, where GCMR-PRED provides the lowest jitter time in 100-nodes WMNs. However, GCMR-IFQ performs slightly higher jitter time than it does in 50-nodes WMNs.

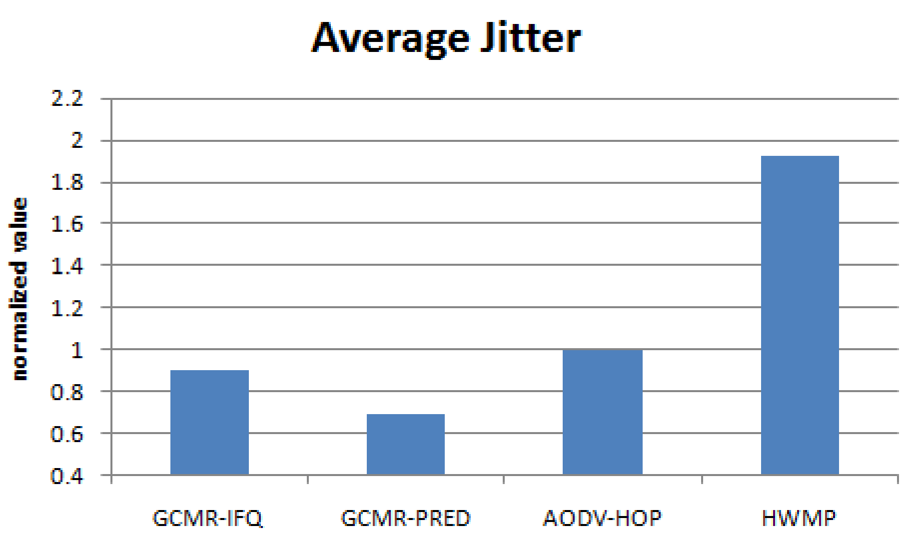


Figure 3.11: A comparison of Average jitter of 100-nodes WMNs

In Table 3.2, we show the comparison of probing overhead in two network sizes, a small network with 50 nodes and a larger one with 100 nodes. As the network size changes, the increasing rate of the probes overhead caused by AODV-HOP, HWMP, GCMR-IFQ and GCMR-PRED with these percentages, 90.6%, 67.8%, 66.2%, 38.9%, respectively. These results confirm that the probing overhead of GCMR is more stable compared to AODV-HOP and HWMP, when changing network size and traffic load, especially with the traffic prediction method and the TTL setup mechanism. In conclusion, GCMR first shows distinguished performance compared to other algorithms in 50-nodes WMNs. We also show a large scenario with 100 nodes. The results show even in large networks, GCMR-PRED still provides efficient communication compared to other algorithms.

Multicast algorithm/ Results	AODV-HOP	HWMP	GCMR-IFQ	GCMR-PRED
%Overhead-50 nodes	3.2	2.8	6.5	5.4
%Overhead-100 nodes	6.1	4.7	10.8	7.5
% Increasing rate	90.6	67.8	66.2	38.9

Table 3.2: Comparative percentage overhead for the different algorithms

3.5 Conclusion

In this chapter, a new novel routing algorithm namely, Gateway-Centralized Multi-hop Routing protocol (GCMR) is proposed to enable the routing management in the gateway with the leaf-to-gateway update mechanism, the traffic prediction method and the TTL setup mechanism. It significantly reduces the flood caused by the broadcasted request messages of the mesh nodes. The results show that the proposed algorithm GCMR outperforms its counterparts significantly, including the well-known AODV-HOP and HWMP routing algorithms. Moreover, in Chapter 5, high-efficiency group communication algorithm is studied by modifying and enhancing GCMR.

In the next chapter, a load balancing multicast technology is introduced for Wireless Mesh Networks.

Chapter 4

Load Balancing Multicast in Wireless Mesh Networks

4.1 Introduction

In the last few years, the demand for group communication based applications has significantly increased. More and more people prefer to deal with a wide of applications to watch football match and TV drama using Internet instead of the traditional TV. As a key communication technology, multicast operation aims to efficiently deliver information from a sender node to multiple receivers, with conserved bandwidth, minimised delay time and low error rate. Thus, multicast is widely used by the service providers to deliver services such as Internet TV, video conference, distance learning, multicast based gaming and so forth to multiple subscribers.

The load balancing issue is a cornerstone of efficient communication patterns within WMNs. In fact, existing multicast proposals for MANETs [45] [46] cannot be applied in WMNs directly without modification as none of these proposals have considered the unique characteristics of WMNs. There are also a number of multicast algorithms proposed for WMNs [47] [51] [96]. However, none of these proposals consider the optimising of load balancing within the context of multicast communication in WMNs. In WMNs, client nodes connect to the Internet via a gateway, which, in turn, acts as a relay node on the multicast tree.

It is worth indicating that, the gateway node and the mesh nodes on the shortest paths from mesh nodes to a gateway can be extremely heavy during certain periods. This is due to the fact that gateways and centre nodes relay both multicast packets and unicast packets as background traffic. Thus, this can result in disastrous consequences on the overall performance of the network. However, with controlling the client-gateway registration, the load balancing can be improved dramatically, leading to a reduction in latency and transmission errors. In the light of this, Gateway-cluster based Load Balancing Multicast algorithm (GLBM) is proposed to enhance the load balancing in multicast communication over WMNs. For the multicast applications with high-bandwidth and instant communication requirements such as video conference and Internet TV, the most important metrics are delay and throughput. For this reason, the new algorithm focuses on high throughput and low delay multicast delivery through achieving load balancing. The reminder of this chapter is organised as follows. Section 4.2 presents the background and motivation behind the proposal of this new algorithm. Section 4.3 describes the proposed algorithms. Then, Section 4.4 shows the performance results of the proposed algorithm compared with existing algorithms. Finally, Section 4.5 provides a summary of this chapter.

4.2 Background

The load balancing within multicast communication in WMNs has been investigated lightly in the literature. First, although existing work [82] [85] [105] tackles the load balancing of multicast in MANETs, these proposals cannot be implemented directly without modification in WMNs. Second, in WMNs, existing work such as [106] is proposed to improve the gateway load balancing. In contrast, to the best of our knowledge, there is no previous study devoted to improve the centre node load balancing for efficient multicast in WMNs. Section 4.3 explains our proposed GLBM protocol.

4.3 Gateway Load Balancing Multicast Protocol (GLBM)

4.3.1 Preliminaries

The problem of communication quality aware multicast is treated as a weighted directed graph. A typical WMN is a network, in which those nodes are connected to each other by wireless links.

Notation	Description
LB_i	the traffic load of *node_i*
$\overline{LB}$	the average load of the network
$\overline{LB(N)}$	the average load of a group of node N after multicast session is initiated
$LB_{t@before}$	the total load of *node_t* before multicast session is initiated
j	the number of nodes are involved in the packet transmissions
i	the number of nodes in the network
$LB(G)_{before}$	the total load of a network before multicast session is initiated
$LB(M)_{before}$	the total load of all the multicast nodes before multicast session is initiated
$\overline{LB}_{G-hop}$	the average load of network after the multicast session is initiated (if the Hop-count multicast algorithm is used)
$\overline{LB}_{M-hop}$	the average load of multicast tree after the multicast session is initiated (if the Hop-count multicast algorithm is used)
$\overline{LB}_{G-load}$	the average load of network after the multicast tree is built (if the Load-count multicast algorithm is used)
$\overline{LB}_{M-load}$	the average load of multicast tree after the multicast session is initiated (if the Load-count multicast algorithm is used)
∂	the load caused on each tree node by handling multicast packet transmission
n	the total number of nodes in network
h	the number of total nodes on the multicast tree for the Hop-count multicast
l	the number of total nodes on the multicast tree for the Load-count multicast
v	the number of multicast packet transmission in a node

Table 4.1: Notation used in the analysis

Therefore, a WMN is presented as $G=(V,E)$, where V denotes the nodes in the network and E is the links between each pair of nodes. The multicast routing is from one single source node to multiple receiver nodes. Then, let $M=\{m_0, m_1, m_2, ...m_{n-1}\}$ be a set of multicast nodes in a network where m_0 is the source node. The following three definitions are presented for load balancing and multicast communication. The algorithm is designed in line with these definitions. Table 4.1 lists notations used in Definitions 4.1, 4.2, 4.3, Lemma 4.1, 4.2 and Theorem 4.1.

Definition 4.1: The load balancing in a network *G*= (*V*, *E*) is achieved, if $LB_i \to \overline{LB}$, $i \in V$.

Definition 4.2: Given a WMN network, *G*= (*V*, *E*), the total load of a network *G* is: $LB(G) = \sum_{i \in V} LB_i$ and the total load of a multicast group *M* is: $LB(M) = \sum_{i \in M} LB_i$.

Definition 4.3: Given a multicast clients $M = \{m_0, m_1, m_2, ... m_{n-1}\}$ in a network, $G = (V, E)$ if the multicast session is load balanced, thus, $\overline{LB(M)} \to \overline{LB(G)}$. Based on Definition 4.1 and 4.2, the traffic load level of each node has a close approximation to the average traffic load of the network. In a load balancing network, the average load of multicast session, $\overline{LB(M)}$ therefore approaches the average load of the whole network, $\overline{LB(G)}$.

Lemma 4.1: Given a network $G = (V, E)$, by comparing the number of hops between two kinds of multicasts, then *h* < *l*.

Proof: the multicast algorithm with Hop-count routing metric finds the shortest path which is the path with minimum hops. Therefore, the number of hops for the Hop-count multicast is less than other multicast schemes.

Lemma 4.2: Given a network $G = (V, E)$, the average load of a network with multicast session is given by $\overline{LB(N)} = \frac{LB(N)_{before} + \partial \times j}{i}$.

Proof: The load of a node t caused by handling multicast packets is: $LB_t = LB_{t@before} + \partial \times v$, while the total number of the packet transmission of all nodes in a network is: $j = v_0 + v_1 + ... + v_{i-1}$; the total load of i nodes before multicast is: $LB(N)_{before} = LB_{0@before} + LB_{1@before} + ... + LB_{i-1@before}$; and the total load of a network after multicasting is started is: $LB(N) = LB_0 + LB_1 + ... + LB_{i-1}$.

Finally the average load of a group of nodes N after initiating multicast is:

$$\begin{aligned}
\overline{LB(N)} &= \frac{LB(N)}{i} \\
&= \frac{LB_0 + LB_1 + ... + LB_{i-1}}{i} \\
&= \frac{LB_{0@before} + \partial \times v_0 + LB_{1@before} + \partial \times v_1 + ... + LB_{i-1@before} + \partial \times v_{i-1}}{i} \\
&= \frac{LB_{0@before} + LB_{1@before} + ... + LB_{i-1@before} + \partial \times v_0 + \partial \times v_1 + ... + \partial \times v_{i-1}}{i} \\
&= \frac{LB_{0@before} + LB_{1@before} + ... + LB_{i-1@before} + \partial \times (v_0 + v_1 + ... + v_{i-1})}{i} \\
&= \frac{LB(N)_{before} + \partial \times j}{i}
\end{aligned}$$

Theorem 4.1: To perform multicast communication in a given network $G = (V, E)$, the Load-count multicast algorithm outperforms the Hop-count algorithm in terms of load balancing.

Proof:

$$\overline{LB}_{G-hop} = \frac{LB(G)_{before} + \partial \times h}{n} \qquad (4.1)$$

$$\overline{LB}_{G-load} = \frac{LB(G)_{before} + \partial \times l}{n} \qquad (4.2)$$

$$\overline{LB}_{M-hop} = \frac{LB(M)_{before} + \partial \times h}{h} \qquad (4.3)$$

$$\overline{LB}_{M-load} = \frac{LB(M)_{before} + \partial \times l}{l} \qquad (4.4)$$

Based on Definition 4.3, in a load balancing aware network, the average load of multicast session, $\overline{LB(M)}$, approaching the average load of the whole network, $\overline{LB(G)}$. We can compare the load difference for each type of multicast algorithm to determine whether it is load balanced. The load difference for the Hop-count multicast algorithm is formulated as $\overline{LB}_{M-hop} - \overline{LB}_{G-hop}$ and the load difference for the Load-count multicast algorithm is represented as $\overline{LB}_{M-load} - \overline{LB}_{G-load}$. Then, by checking whether the load difference for the Load-count multicast algorithm is smaller than the load difference for the Hop-count multicast algorithm, the result is:

$$\overline{LB}_{M-load} - \overline{LB}_{G-load} - (\overline{LB}_{M-hop} - \overline{LB}_{G-hop}) = \frac{LB(M)_{before}(h-l)}{l \times h} + \frac{\partial \times (h-l)}{n} < 0$$

$$\overline{LB}_{M-load} - \overline{LB}_{G-load} < \overline{LB}_{M-hop} - \overline{LB}_{G-hop}$$

Therefore, the load difference for the Load-count multicast algorithm is smaller than that for the Hop-count multicast algorithm. According to definition 4.3, the Load-count multicast algorithm achieves better load balancing.

In this thesis, a new multicast algorithm is proposed, namely, Gateway-cluster based Load Balancing Multicast algorithm (GLBM). The load capture mechanism is implemented to fetch the load status along each mesh node on the multicast tree. The main contribution of this algorithm is the multicast communication with high communication quality by avoiding uneven traffic load. GLBM is a hybrid multicast routing algorithm since the hybrid multicast routing reduces the route discovery overhead compared to on-demand driven multicast [45] [42]. When the multicast source node broadcasts periodic multicast hello messages to all the gateway nodes proactively, the receiver nodes join the multicast session by sending requests to their gateway nodes reactively.

Existing on-demand driven multicast algorithms such as MAODV [45] and ODMRP [46] are all specifically designed for MANETs where the major concerns are energy efficiency and mobility, which is not the case in WMNs. Further, most mesh nodes either exhibit low mobility or operate within fixed positions. Thus, this necessitates the need for implementing hybrid multicast in order to manage both mobile nodes and stationary nodes within the networks.

In this algorithm, the multicast session is divided into four phases outlined in the following sections. The reminder of this chapter is organised as follows. Section 4.3.2 shows how the source node starts a multicast session. Section 4.3.3 refers to the way whereby the multicast receivers join the multicast group. Section 4.3.4 tackles the way in which the receivers leave the multicast group and how the multicast group is maintained respectively. Each node must register with its gateway before sending or receiving any data.

4.3.2 Initiating a Multicast Session

When a source node intends to initiate a multicast session, it broadcasts the multicast sender request (MSR) messages towards its gateway similar to the mechanism of RREQ in AODV [43]. After receiving the MSR messages, the gateway node broadcasts the multicast hello messages (HM) periodically with its routing detail such as sequence number to all the gateway nodes in the backbone. Upon reception of a HM message, the gateway node records the backward path to the source node in its multicast routing table and ignores the old entry if the sequence number is fresher than the existing one.

Algorithm 4.1: Pseudo code of Gateway-cluster based Load Balancing Multicast algorithm

Input: source *s*, gateway *g*, receiver *r*, forwarder node *f*

Phase 1: source s start multicast

1 *s* bcast(MSR) to all *g*

2 *g* recv(MSR) and record MSR in gateway multicast table

Phase 2: receiver joins multicast group s

3 *r*.send(MRQ-J) with IFQ_Agg=0 and s_id

4 **if** *f* recvd MRQ-J

5 MRQ-J.IFQ_agg+=f_IFQ

6 *f*.send(MRQ-J)

7 **end if**

8 **if** *g* recvd non-duplicate MRQ-J with lowest IFQ_Agg

9 **if** *g*.check(s_id)==1

10 *g* record MRQ-J and send(MRP)

11 **else**

12 *g*.send(MRP) to r

13 **end if**

14 **end if**

15 *s* recvd MRP and add *r* in the routing table

Phase 3: multicast receiver leaves multicast group s

16 *r*.send(MRQ-L) with s_id to *g*

17 *g* recvd MRQ-L and *g*.delete(r) in the gateway multicast table

18 *g*.send(MRQ-L) to *s*

19 *s*.delete(r) in the multicast routing table

Phase 4: error repair of multicast group

20 **if** *r* haven't recvd multicast packets for a time *t*

21 *r*.send(MRQ-R) to *g* with s_id

22 **end if**

23 repeat line 4 to 7 with flag *R*

24 **if** *g* recvd MRQ-R and *g*.check(s_id)==0

25 *g*.send(MRR-F)

26 **else**

27 choose MRQ-R with lowest IFQ_Agg and record it

28 **end if**

4.3.3 A Client Node Joins Multicast Session

According to Phase 1 and Phase 2 of Algorithm 4.1, a mesh node broadcasts a multicast route request with join flag (MRQ-J) messages including the destination IP address towards its registered gateway. Once receiving this request, each intermediate node adds node's current IFQ length to the IFQ Aggregation (load aggregation) in the header of message and broadcasts the messages to the one-hop neighbours again. IFQ length is obtained by examining the network interface and traffic load level of the node. When the gateway node receives the MRQ-J requests, it checks the message header for the IFQ Aggregation. Then, it chooses the path with the lowest IFQ Aggregation as the multicast path from the gateway to the receiver. The gateway then checks its multicast routing table. If there is an entry with the same IP address, the gateway node sends a multicast route reply (MRP) combining the path from the request client node to the gateway and the path from the gateway to the source node. Otherwise, the gateway should send a multicast source reply with repair flag (MSQ-R) to this client node. Once receiving a MRP, the gateway of source node records the backward path and starts to forward multicast messages to the client node.

If traffic load on a node is significantly large, the forwarded packets are more likely to cause high packet loss and packet delay. Therefore, the lowest load aggregation path is selected as an edge of the multicast tree in order to avoid the duplicate uses of the busy node. Consequently, as the nodes are equally used, the load of each node approximates to the average load. This can be seen as a centre load balancing problem.

Most of existing multicast algorithms such as ODMRP [47] and MAODV [45] apply the Hop-count as a routing metric in the path selection. An example is shown by comparing the Load-count multicast with the Hop-count multicast in Figure 4.1 and Figure 4.2. In each node, there are two numbers to represent the node id and load status. Figure 4.1 shows a WMN with 24 mesh nodes and one gateway, when *node_6* intends to join the multicast group. Although there are lots of possible paths to the gateway, only two most typical request paths (one is the best load route and the other is the best hop route) are shown. Let the load, caused by the multicast packet transmission, be 1 on each tree node. The Load-count path is illustrated as the solid arrow (6-1-2-3-4-5-10-15) where the aggregate load is 8.

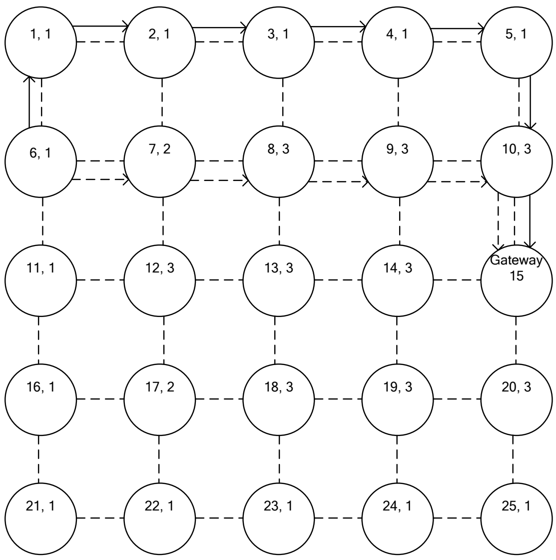
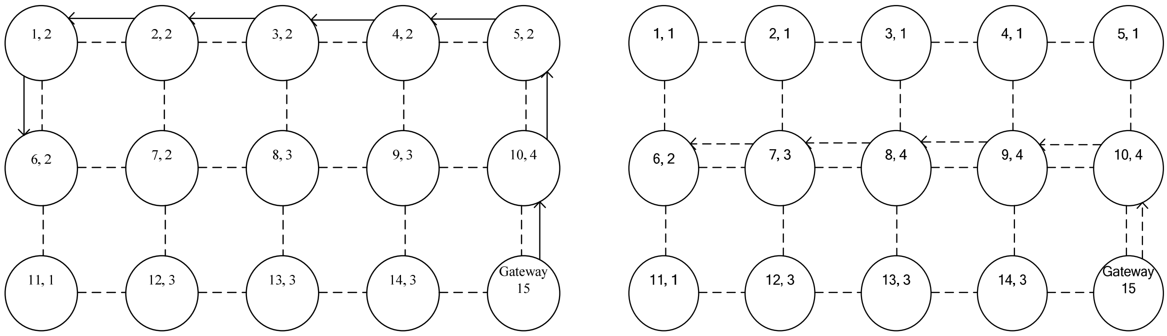


Figure 4.1: Two request routes of node 6

In contrast, the best Hop-count path with dot arrow goes along (6-7-8-9-10-15) with the aggregate load of 11. Figure 4.2 (a) and (b) show the replies by using the Load-count multicast and the Hop-count multicast, respectively. $\overline{LB}$ of network (a) is about 2.54 and $\overline{LB}$ of network (b) is about 2.58. There is one node with load value of 4 in (a), whereas there are 3 nodes with load of 4 in (b). Therefore, the traffic load of (a) is distributed much evenly than (b) according to definition 4.1. This indicates that GLBM achieves better load balancing than the algorithm with Hop-count routing metric.

In Figure 4.3, there are two multicast clients (*node_6* and *node_17*). In the situation of using the best Hop-count paths (6-7-8-9-10-15) and (15-20-19-18-17), the following results are found: $\overline{LB}_G \approx$ 2.208 and $\overline{LB}_M \approx$ 3.555. With the best Load-count path (6-1-2-3-4-5-10-15) and (15-20-25-24-23-22-17), $\overline{LB}_G =$ 2.375 and $\overline{LB}_M \approx$ 2.385. According to the definition 4.3, the multicast session along the best Load-count path is relatively load balanced for this example.



(a) Multicast transmission route selected by the Load-count multicast

(b) Multicast transmission route selected by the Hop-count multicast

Figure 4.2: Comparison between the Load-count multicast and the Hop-count multicast

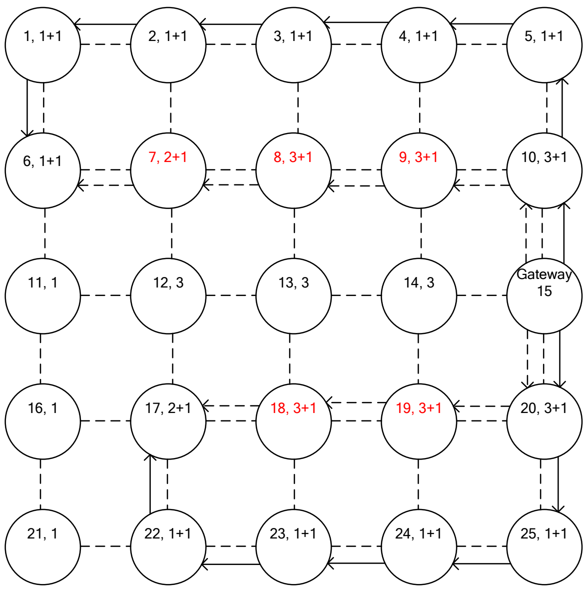


Figure 4.3: Load balancing for multicast tree

4.3.4 Maintaining a Multicast Session

When a multicast receiver intends to leave the multicast session, it sends out the multicast route request with leave flag (MRQ-L) along the default multicast path as reported in the Phase 3 of Algorithm 4.1. As its gateway receives MRQ-L, it deletes the entry of this node in the multicast routing table and then notices the multicast source by using a multicast source reply with the leave flag (MSQ-L). When the source node receives MSQ-L, it deletes the entry for the receiver node in the multicast routing table and stop sending multicast packets to this node.

4.3.5 Repairing a Multicast Session

The Phase 4 of Algorithm 4.1 presents the error repair mechanism of GLBM. When a multicast subscriber *r* does not receive any packet from the multicast source *s* for a certain period of time, *r* assumes that either a link between *r* and the gateway is broken or *s* stops multicasting. Then the source node broadcasts multicast route request with repair flag (MRQ-R) messages to its gateway. MRQ-R request also records the IFQ Aggregation as the MRQ-J does. When the gateway receives MRQ-R, it first checks whether it is still forwarding the multicast packets from a source *s* to this receiver *r*. If so, the best Load-count path is selected and updated. Then, MRP message is replied to *s*. Otherwise, it means that the source node is no longer transmitting. Hence, the gateway sends a multicast route reply with a finish flag (MRR-F) to the receiver.

4.4 Performance Evaluation

GLBM is simulated by using the NS-2 simulator [64] to compare GLBM with MAODV [45] and ODMRP [46]. In the simulation, the aim is to examine whether the multicast algorithm is suitable for the applications with large packet sizes and instant communication requirements. For this purpose, four performance metrics, namely, *maximum throughput, maximum end-to-end delay, jitter* and *probing overhead* are used in the experiments. It is worth indicating that these metrics have been used widely by existing work such as [47] [85] [107], to examine the capability of a multicast algorithm. The maximum throughput is the maximum rate of a successful packet delivery in a time interval. A higher throughput of the multicast session can provide greater bandwidth for the communication. The maximum end-to-end delay is the maximum time a packet travel from a source to a destination. On the other hand, the jitter is the time variation of the packet arrivals at the destination. The probing overhead refers to the total cost of constructing and maintaining the multicast tree.

4.4.1 Simulation Model

As previous study in [57] [85] [107] [108] [109], the simulation models a network of 50 wireless routers uniformly distributed over a 1000 m × 1000 m area with two gateways as described in Table 4.2. To simplify the simulation, the first gateway is set as the multicast source node in one WMN and second gateway with 50 clients in another WMN. To the best of my knowledge, this is the first study that considers the gateway in the multicast of mesh networks. The multicast group size is varied from 10 to 40 members in 50 nodes. According to [57] [110], the nodes are in the fixed position to simulate community mesh networks. The multicast source sends Constant Bit Rate (CBR) traffic, consisting of 1024-byte packets with a sending rate of 20 packets per second. It tests whether the algorithm is capable of handling the applications with high-bandwidth demand. To make the network busy enough, nodes randomly generate the background CBR packets by unicast traffic. It is the test of the algorithm performance in a busy network environment. This also measures the scalability of multicast algorithms. Each algorithm is simulated on 10 different randomly generated topologies. Each simulation runs for 400 seconds and the average results over all topologies are presented.

Network size	50 fixed nodes over a 1000 m × 1000 m area 100 fixed nodes over a 2000 m × 2000 m area
Multicast group size	Between 10 and 40
Router transmission power	20 dBm
Interference range	500 m
Transmission rate at physical layer	54 Mbits/s
Physical layer protocol	PHY802.11g
Packet size (excluding header size)	1024 bytes
Queue size at wireless routers	50 Kbytes
Traffic model of sources	Constant bit rate (CBR)
Number of CBR senders	30
CBR sender's rate	20 packets/s
Simulation duration	400 seconds

Table 4.2: Simulation parameters

In the simulation of wireless nodes, network interface queue (IFQ) buffers packets as a stack to manage all the packets coming in and going out of a node. IFQ length is set to limit the maximum number of packets that can be held in the queue. In the algorithms, IFQ length is used to examine the traffic load of a certain node.

4.4.2 Simulation Results

In well-known applications such as multicast video conference and multicast based gaming, delay is an important performance metric to exam the quality of service. In Figure 4.4, GLBM exhibits smaller maximum end-to-end delay than MAODV and ODMRP in most cases. However, when the multicast group size is as small as 10 nodes, MAODV works slightly better than GLBM and ODMRP. Nevertheless, GLBM provides the lowest maximum end-to-end delay as the group size increases. As these results shows, instant group communication can be achieved with GLBM. This indicates that GLBM works well in scalable networks. Figure 4.5 describes the relationship between the number of multicast nodes and network throughput. GLBM achieves at least 27% more throughput than MAODV and at least 14% more throughput than ODMRP. Therefore, GLBM can better satisfy the high-bandwidth requirement of the instant communication applications.

Figure 4.6 depicts the jitter differences between GLBM, ODMRP, and MAODV. Jitter is an extremely important metric for real-time applications such as multicast video conference. It is normally caused by the network congestion which leads to different delay time between packets. When jitter is high, out of range packets are discarded in the destination. In addition, audio and video problems are heard and seen. GLBM has lower jitter than MAODV. Thus, GLBM can provide sensibly better real-time transmissions.

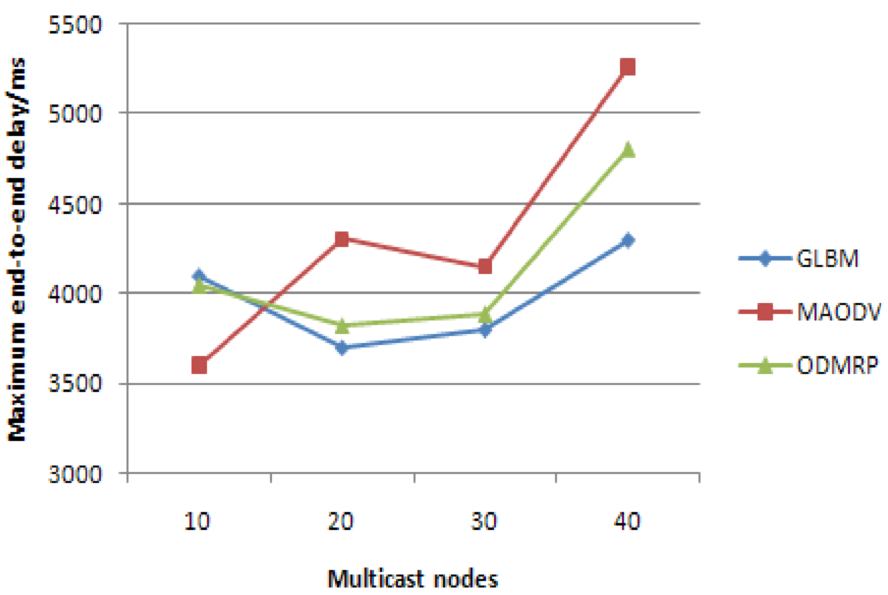


Figure 4.4: Multicast nodes vs. maximum end to end delay

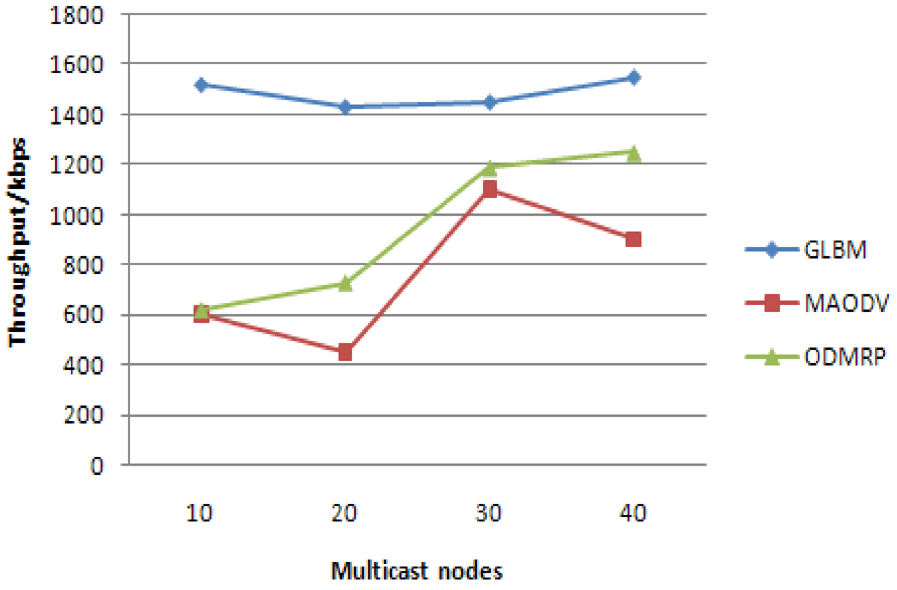


Figure 4.5: Multicast nodes vs. Throughput

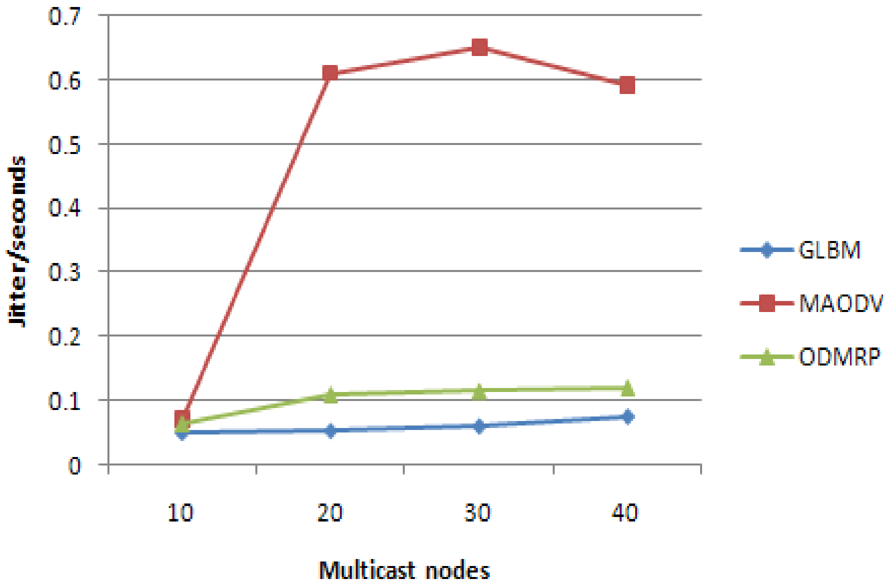


Figure 4.6: Multicast nodes vs. Jitter

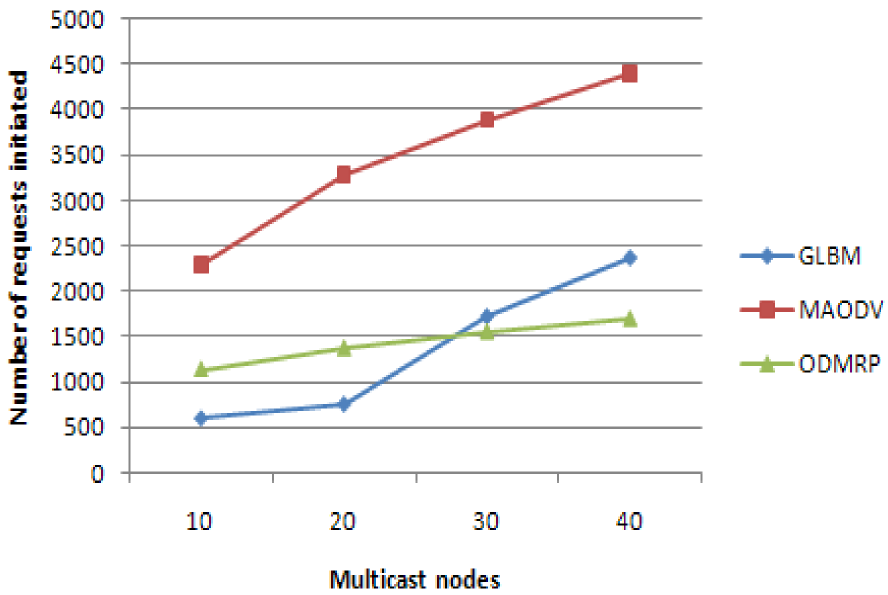


Figure 4.7: Multicast nodes vs. the number of requests initiated

Figure 4.7 shows the routing overhead incurred for constructing and maintaining the multicast tree, expressed as the number of requests sent by the nodes during the multicast session. As depicted in Figure 4.7, GLBM generates much less request packets than MAODV does. However, when group size is 30 and 40, ODMRP works slightly better than GLBM. Thus, both GLBM and ODMRP provide low overhead in the construction and maintenance processes.

We also examine the routing overhead of GLBM, ODMRP and MAODV in a larger mesh network with 100 fixed nodes in order to show the scalability of GLBM. The results are shown in Table 4.3. The first two rows of this table show the percentages that the total probing packets in bytes divided by the total cumulative sum of packets generated by the network in bytes. Then the third row shows the results of average percentage of two different sized networks, for MAODV, ODMRP, and GLBM, which are 10.7%, 5.2%, 3.2%, respectively. This confirms that the probing overhead of GLBM generates lowest routing overhead. In another hand, these results also show the stable performance of GLBM through the changing of network size.

Multicast algorithm/ Results	MAODV	ODMRP	GLBM
% overhead-50 nodes	10.9	5.6	2.7
% overhead-100 nodes	10.5	4.8	3.6
% average rate	10.7	5.2	3.2

Table 4.3: Comparative percentage overhead for the different multicast algorithms

4.5 Conclusion

Load balancing is shown as the paramount priorities in the multicast communication. Unfortunately, load balancing within the context of WMNs has been treated lightly in the literature. As a result, most of existing work has shown deficiency in handling delay sensitive applications such as video conference, online gaming and other real-time applications. Thus, in this chapter, a new multicast algorithm, GLBM is proposed for WMNs. In addition, GLBM is described and compared with well-known existing multicast algorithms. In particular, unlike many existing work, GLBM has been designed to handle the applications with high-bandwidth and real-time requirements. The simulation results demonstrate that GLBM is better than its counterparts in performing these applications.

In the next chapter, Multicast Gateway Centralized Multi-hop Routing is proposed and studied for WMNs.

Chapter 5

Multicast Gateway Centralized Multi-hop Routing for Wireless Mesh Networks

5.1 Introduction

As a group communication pattern, multicasting provides an efficient way to deliver packets from the service provider to multiple subscribers in any networks. Due to the increasing demand of group applications, the multicast communication has been attracting tremendous attention from both industry and academia. Indeed, multicast communication is one of the most primitive collective capabilities of broadband network. It is also central to many important group based applications, and fundamental to the implementation of other group communication based operations. Furthermore, multicast is widely used in many applications such as real-time applications including online TV, distance learning and gaming and so forth. Evidently, multicast reduces the cost of communication compared to sending unicast packets multiple times. Moreover, multicast conserves bandwidth, reduces packet delay time and level of network congestion. However, existing standards for WMNs are not capable of supporting the emerging multicast based real-time communication applications that require efficient gateway routing management. Further, existing solutions [45] [46] [96] offered for providing multicast in WMNs have severe restrictions in terms of different performance metrics, especially jitter and delay. Motivated by the above observations, we present a new multicast algorithm named, Multicast Gateway Centralized Multi-hop Routing algorithm (MGCMR). MGCMR optimises the multicast capability of WMN to support and offer high quality communications required by the real-time multicast based applications which incur high-bandwidth usage.

The proposed MGCMR routing algorithm is a multicast version of GCMR (described in Chapter 3) inosculated with the features of GLBM (described in Chapter 4) in this chapter.

As mentioned in Chapter 3, existing proposals [71] [76] [101] have involved gateway (access point) to offer the centralized multi-hop routing in cellular networks. They aim to either enhance the coverage of network or improve the central management of the network. However, most of existing work [93] [103] treat mesh networks as MANETs. They omit, unfortunately, the role of gateway for the central network resource management at the provider side. In other words, the issue of centralized multi-hop routing in WMNs is rarely studied, especially within the multicast capability. As a new characteristic, gateway should be involved somehow in the construction of multicast tree/mesh in WMNs. For example, with the prefix continuity mechanism [96], the gateways act as relay nodes in the multicast tree. However, besides forwarding and replying packets by the gateway, existing multicast algorithms do not fully involve gateway in handling routing tasks such as computing/constructing the multicast tree.

In contrast, MGCMR aims to optimise the communication performance of multicast for WMNs, so as to provide efficient communications to the real-time multicast applications with high-bandwidth consumption.

The reminder of this chapter is organised as follows. Section 5.2 describes the proposed algorithms. Then, Section 5.3 shows the simulation results of the proposed algorithm compared with existing algorithms. Finally, Section 5.4 provides a summary of this chapter.

5.2 Gateway Centralized Multi-hop Routing algorithm (MGCMR)

In this algorithm, the node registration process is similar to GCMR. Before a node becomes a multicast source or multicast receiver, it must register with the gateway using REQUEST as shown in Chapter 3. Like in GCMR, the leaf-to-gateway update mechanism is applied in MGCMR to ensure that the gateway knows the instant topology of the network.

5.2.1 Initiating a Multicast Session

To start a multicast session, a mesh node *source_s* sends a multicast source request message (MULTICAST_SREQ) to notify the central gateway. As the gateway receives this message, it creates the gateway multicast table for *source_s* and records the new multicast session *id*. After that, the gateway node broadcasts the multicast hello message (MH) periodically with its routing information such as multicast session *id* to all the gateway nodes in the backbone.

5.2.2 A Receiver Node Joins Multicast Session

In MGCMR, the gateway acts as a routing management centre to construct the shortest path tree (SPT) rooted at *source_s*. Any node, for example *node_m*, that intends to join the multicast session, should send a multicast route request (MROUTE_REQUEST) with join flag message towards the gateway. When this join request is received, the gateway first checks whether there is a gateway multicast table with the same multicast session. If so, the gateway confirms the multicast join process as inner WMN and then calculates the best path from *source_s* to *node_m* bypassing the gateway.

Notice that, this path is encapsulated in the multicast route reply (MROUTE_REPLY) with join flag message, and then is sent the join reply towards *source_s* along a pre-calculated path. However, be aware that the gateway does not forward data packets as it is not a forwarder node in the selected path, but only can handle the route request between a receiver and a source for the inner WMN with the aim of keeping the traffic load at the gateway to a minimum. If there is no entry for this source node in the gateway multicast table, the gateway deals with the join request as an outer WMN process. As a result, it encapsulates the best path from *source_s* to *node_m* via the Internet backbone in the message of MROUTE_REPLY with join flag. It then sends the join reply towards *source_s* followed by a pre-calculated path. In both cases, the multicast routing table is updated by *source_s* once it receives the reply. For the selection process of the best path, a new shortest path tree (SPT) is constructed based on the gateway link table if either there is no existing SPT, or the current SPT is reconstructed to be merged with the path from *source_s* to *node_m*.

When the new SPT tree is rebuilt in the source node, *source_s* must inform the next-level tree nodes to forward MROUTE_REPLY message with notification flag to *node_m*. The tree construction process and data packet forwarding process are described in Algorithm 5.1. During the propagation of the notifying reply, a non-receiver tree node is marked as *forwarder*. Likewise, a receiver that has at least one next-level tree node is marked as a *forwarder_receiver*. Both *forwarder* and *forwarder_receiver* are responsible for forwarding the data packets to the next-level node on the tree. Further, a receiver without a next-level tree node is marked as *leaf_receiver*. When multicast data packet arrives in a *leaf_receiver*, it will not be forwarded as the *leaf_receiver* does not have next-level tree node. As soon as MROUTE_REPLY with notification flag message arrive at *node_m*, the multicast join process is completed. Figure 5.1 portrays an inner WMN multicast tree with two *leaf_receivers* (*node_5*, *node_6*), one *forwarder_receiver* (*node_4*) and two *forwarders* (*node_1, node_2*). In this example, *node_6* and *node_7* are leaf nodes to generate ROUTE_UPDATE messages.

Algorithm 5.1: The tree construction & data packet forwarding algorithm for $\mathfrak{R}_{in}$ (Pseudo code)

Input: source *s*, gateway *g*, receiver *r, mesh node m, forwarder f.*

Phase 1: tree construction

```
Procedure source_handling_packet(p){
if s recvd MROUTE_REPLY with join flag{
s.update_src_ mcast_table(p);
s.sends MROUTE_REPLY with notification flag to r; }
}
Procedure mesh_handling_packet {
if m=s {
m.update_src_ mcast_table(p); }
if m=f && m=r {
m.status=forwarder_receiver; }
else if m=f {
m.status=forwarder; }
else if m=r {
m.status= leaf_receiver; }
}
```

Phase 2: data packet forwarding

```
Procedure send_multicast_data_packet();
Procedure recv_multicast_data_packet(p){
if m.status=forwarder_receiver||forwarder {
m.forward(p);
}
if m.status=forwarder_receiver||leaf_receiver {
m recvd p; }
}
```

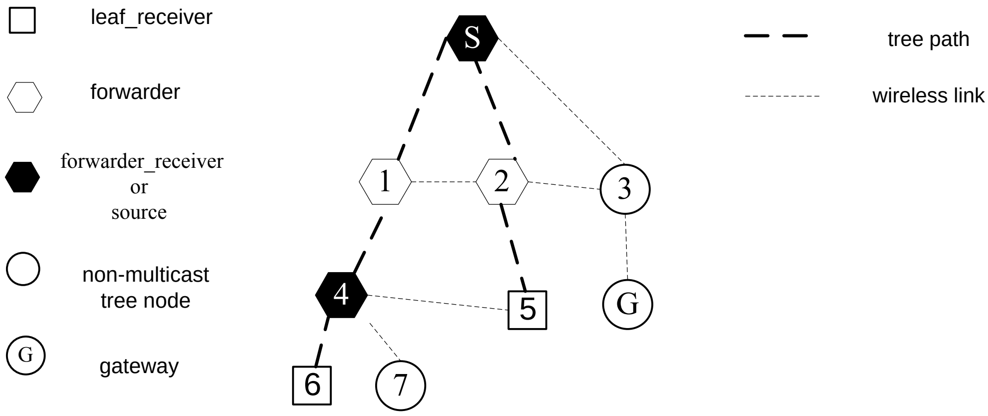


Figure 5.1: Multicast tree construction inside a WMN in MGCMR

5.2.3 A Receiver Node Leaves Multicast Session

A multicast receiver can leave the multicast session at any time. It should send MROUTE_REQUEST with leave flag to the multicast source via the source-registered gateway. As the gateway receives the leaving request, it prunes the node from the tree and updates the multicast routing table. Then the gateway should inform *source_s* to delete the corresponding entry in the multicast routing table of *source_s*.

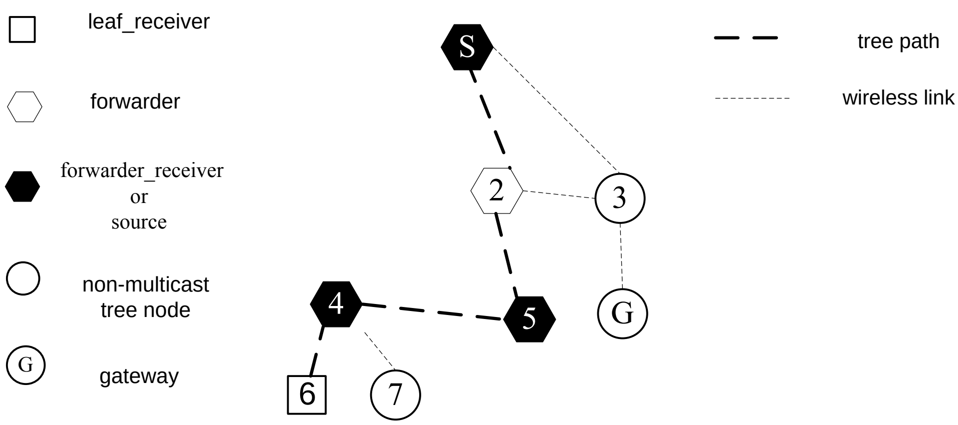


Figure 5.2: Reconstructed multicast tree inside a WMN in MGCMR

5.2.4 Maintenance of Multicast Tree

The maintenance process is achieved from the propagating of ROUTE_UPDATE messages. As same as GCMR, these messages are used to ensure that the gateway keeps the up-to-date paths. Based on the up-to-date gateway link table, the gateway checks the connectivity of all the forwarder nodes in the multicast tree once it receives a ROUTE_UPDATE. If a *forwarder* or *forwarder_receiver* is disconnected, the gateway should reconstruct the SPT tree and passes the new paths to the source by using one or multiple MROUTE_REPLY with error flag message(s). When the source receives the error report, it sends MROUTE_REPLY with notification flag message to notify the affected nodes to complete this repair process. In Figure 5.2, an example is illustrated to show the repair of multicast tree. When a *forwarder node_1* moves away and became disconnected with the original tree (Figure 5.1), the tree is rebuilt. In this case, a former *leaf_receiver node_5* becomes a *forwarder_receiver* and thus, is responsible for relaying packets to *node_4*.

5.3 Performance Evaluation

5.3.1 Simulation Model

The performance of MGCMR is evaluated by using discrete event simulation environment, NS-2 simulator [64]. The simulation aims to examine the performance of MGCMR compared to GLBM, MAODV and ODMRP with regards to the communication quality provision. The following evaluation metrics are used to measure the performance: *average packet delivery ratio, average throughput, average end-to-end delay, average delay jitter,* and *probing overhead*. Besides other four performance metrics, the probing overhead counts all of the request packets in the discovering of the route as well as in the maintaining of the multicast tree.

Network size	50 nodes over a 1000 m × 1000 m area 100 nodes over a 2000 m × 2000 m area
Router transmission power	20 dBm
Radio propagation range	250 m
Transmission rate at physical layer	54 Mbits/s
Physical layer protocol	PHY802.11g
Packet size (excluding header size)	512 bytes
Queue size at wireless routers	50 Kbytes
Traffic model of sources	Constant bit rate (CBR)
Multicast receiver size	{10,20,30} receivers in 50-nodes network {20,40,60} receivers in 100-nodes network
Multicast sender's rate	{50,100,200} packets/s
Simulation duration	400 seconds

Table 5.1: Simulation parameters

As the simulation parameters are shown in Table 5.1, the simulation models a medium-size network of 50 wireless routers uniformly distributed over a 1000 m × 1000 m area, and then a large-size network of 100 wireless routers uniformly distributed over 2000 m × 2000 m area. It is worth indicating that we have used the simulation parameters used widely in previous study such as [107] [108] [109]. Among all wireless routers, 90% of them are set to be stationary and the rest are mobile as to simulate community mesh networks. The gateway is placed in the middle of the network.

The transmission power of the routers is set constant at 20 dBm. The data transmission rate at the physical layer is set to 54 Mbit/s. The two-ray propagation model is used as the radio propagation range is set to 250 m. The data packet size excluding the header size is 512 Kbytes. In the simulation, the various data rates are used to verify whether the algorithms are capable of driving the applications with high-bandwidth and low-bandwidth demand or not. Therefore, the source of multicast session transmits at a constant bit rate (CBR) for 50 packets/s, 100 packets/s and 200 packets/s in order to generate the data rate of 25 Kbytes/s, 50 Kbytes/s, and 100 Kbytes/s respectively. Each algorithm is simulated within 10 different randomly generated topologies with each runs for 400 seconds. The results for the average over all topologies are presented. The results of MGCMR, ODMRP, and GLBM are normalised with respect to that of the original MAODV. In the figures, the simulation results are shown through the increase of data rate, i.e. the discriminative performance of each routing algorithm is examined among various data rates.

5.3.2 Simulation Results

From Figure 5.3 to 5.6, a list of simulation results is illustrated for 50-nodes WMNs. First of all, delay time of GLBM and MAODV are slightly lower than MGCMR when data rates are 50 packets/s and 200 packets/s. In contrast, MGCMR behaves stably in all situations. MGCMR generally achieves the lowest average end-to-end delay in Figure 5.3 and thus ensures the primary prerequisite of real-time communication. Centralized routing allows the best path selected from a big path pool in the gateway rather than in the destination node. This mechanism provides the construction of multicast tree with low latency.

A high jitter is not tolerated in services such as gaming and online video conference where the instant communication is always highly sensitive to delay/jitter. Figure 5.4 indicates that MGCMR reduces the up to 50% of the jitter time compared to others. This guarantees the performance consistency for instant packet transmission. The results of both delay and jitter indicate the leaf-to-gateway update mechanism enables MGCMR to avoid the busy path. The results also show that there is a correlation between delay and jitter as both of them are performance metrics based on the communication cost. The results given in Figure 5.5 also illustrate that the packet loss of MGCMR is the lowest. The high packet delivery ratio results guarantee the packet delivery quality which is particular important in multicast video applications.

As plotted in Figure 5.6, MGCMR exhibits higher throughput than the other algorithms except MAODV with data rate, 200 packets/s. This is interesting because intuitively one would expect MGCMR to perform better than all others, since MGCMR shows always higher packet delivery ratio than MAODV. This is because the relatively high node density in the network and nodes with high data rate typically generate high interference. The interference further leads to more difficulty while updating the instant link status from leaf nodes.

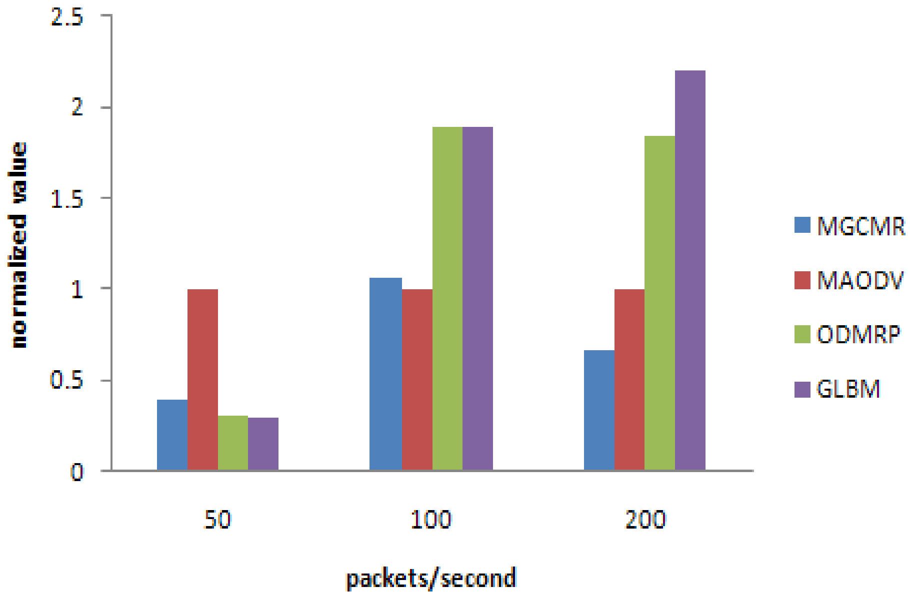


Figure 5.3: Average end-to-end delay of 50-nodes WMNs

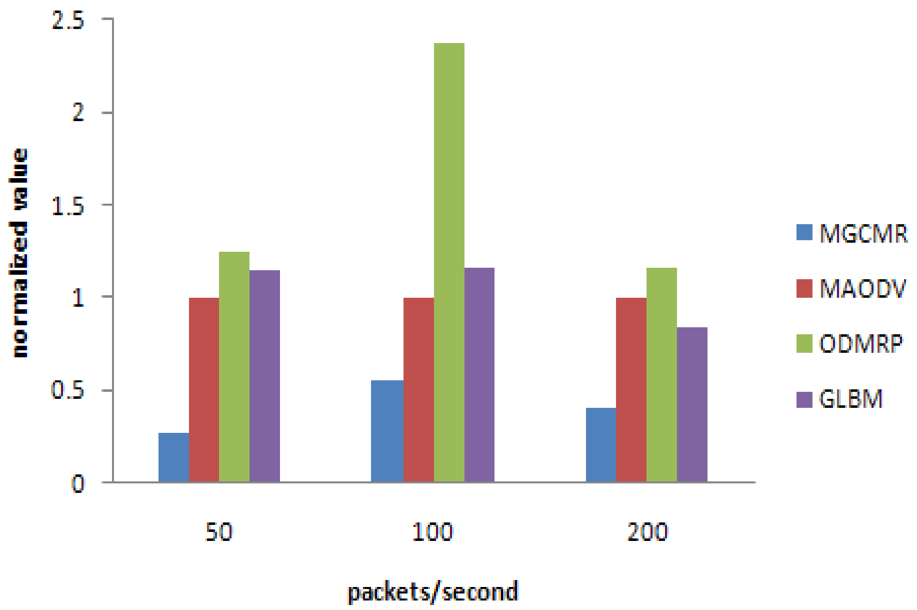


Figure 5.4: Average jitter of 50-nodes WMNs

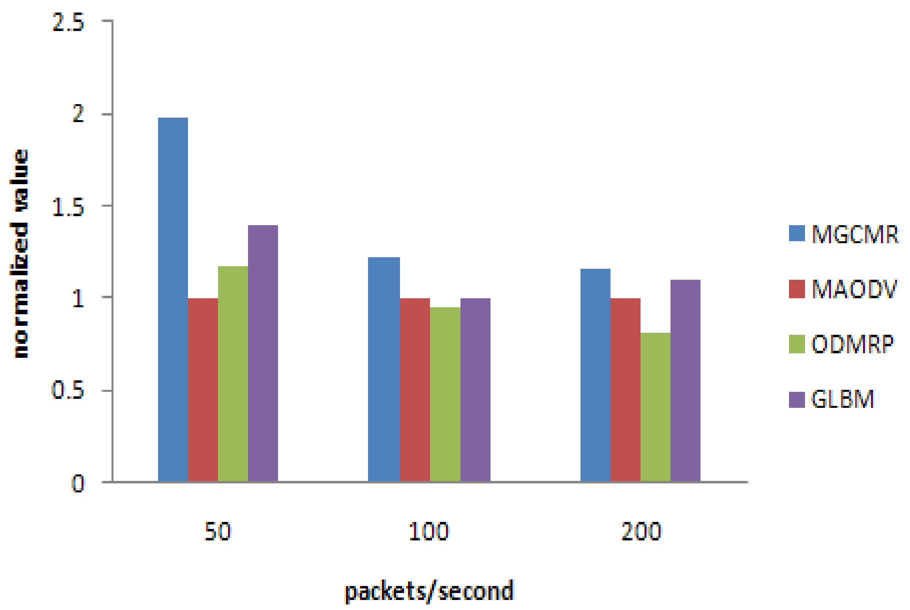


Figure 5.5: Average packet delivery ratio of 50-nodes WMNs

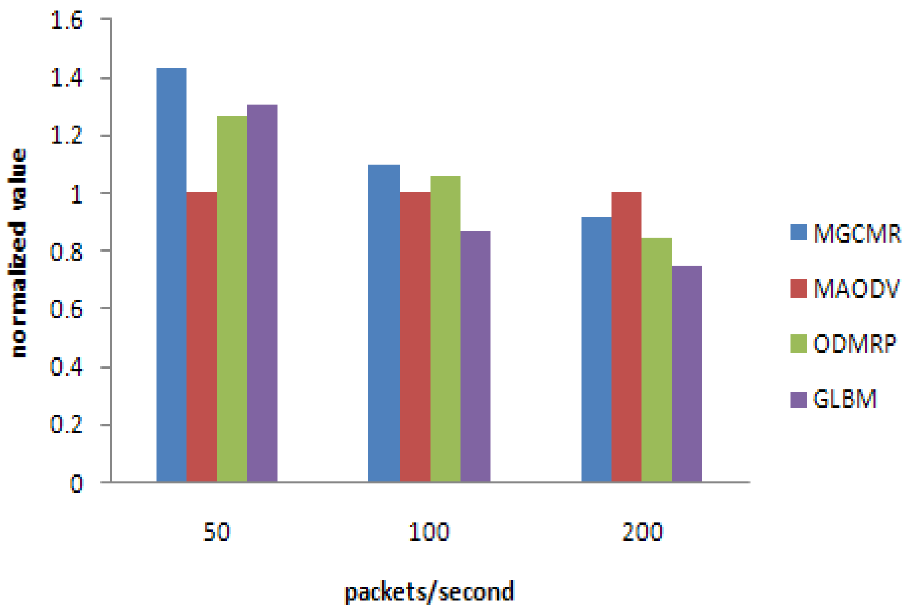


Figure 5.6: Average throughput of 50-nodes WMNs

The outstanding performance of MGCMR is also obtained from the routing metric, Load-count. This routing metric considers the load balancing of the network to provide good path with stable traffic relay services continuously, and avoids poor path to generally increase the results.

The simulation results of a large network with 100 nodes are shown in Figures 5.7-5.10. Compared to the performance in a small network, the average end-to-end delay of MGCMR is more stable with continuously the lowest among the four multicast algorithms as shown in Figure 5.7. In Figure 5.8, MGCMR outperforms MAODV, ODMRP and GLBM in jitter time as reducing the delay differences to at least 50% in large networks. As shown in Figure 5.9, compared to MGCMR, all of MAODV, ODMRP and GLBM present poor packet delivery ratio which do not meet the requirements of efficient multicast communication in large-scale networks. Figure 5.10 shows the throughput gains of MGCMR are comparatively higher than its counterparts in 100-nodes WMNs. In addition, within different network sizes, MGCMR outperforms its counterparts, especially for large networks, i.e., MGCMR has shown high scalability in different operating conditions.

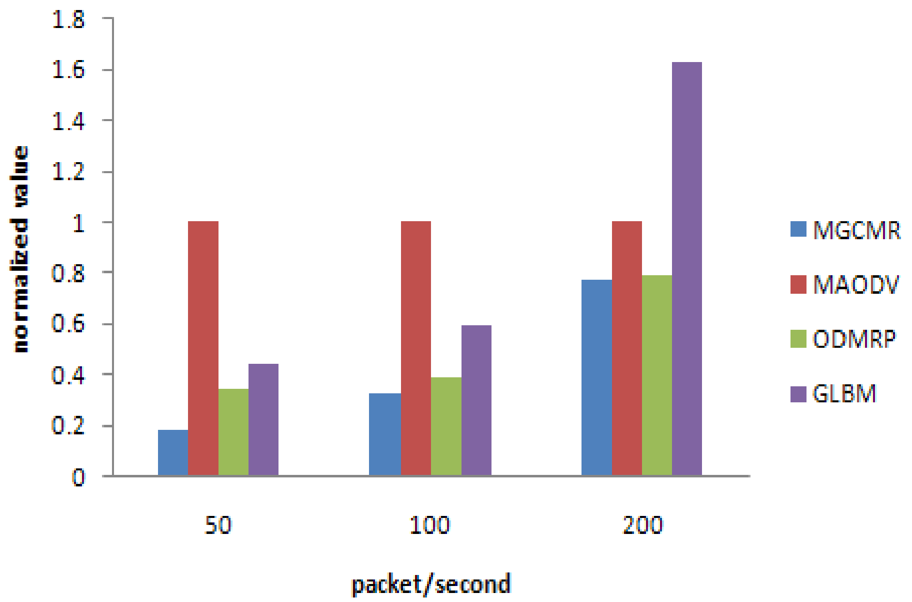


Figure 5.7: Average end-to-end delay of 100-nodes WMNs

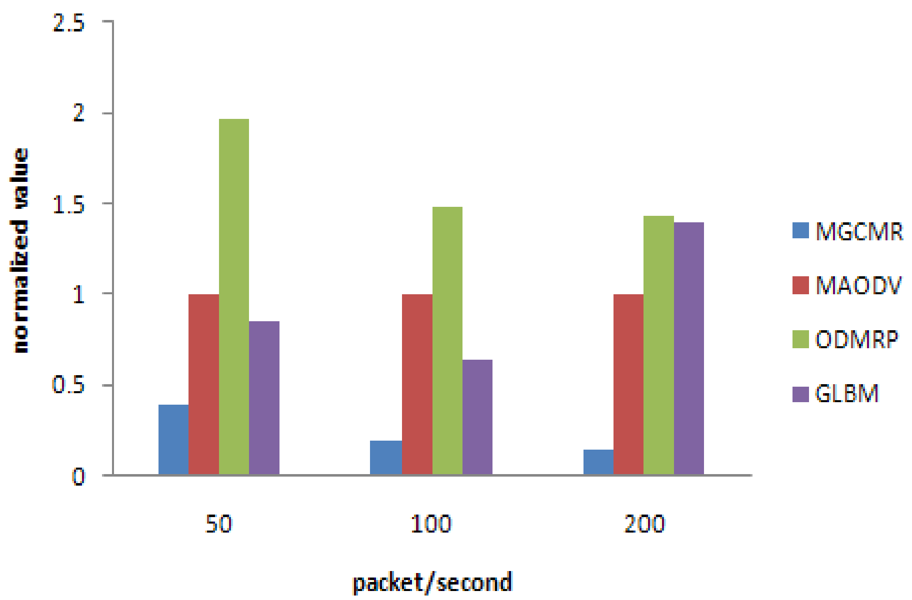


Figure 5.8: Average jitter of 100-nodes WMNs

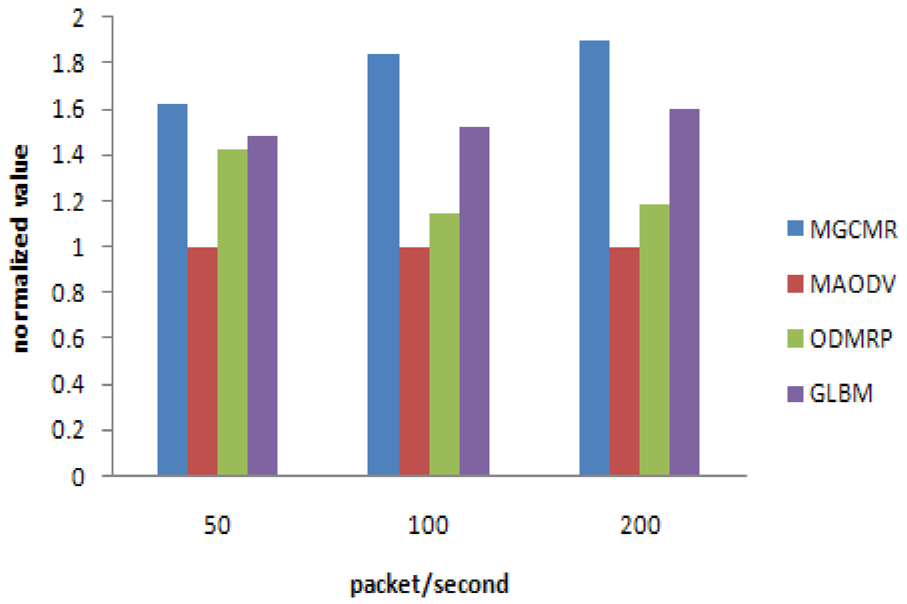


Figure 5.9: Average packet delivery ratio of 100-nodes WMNs

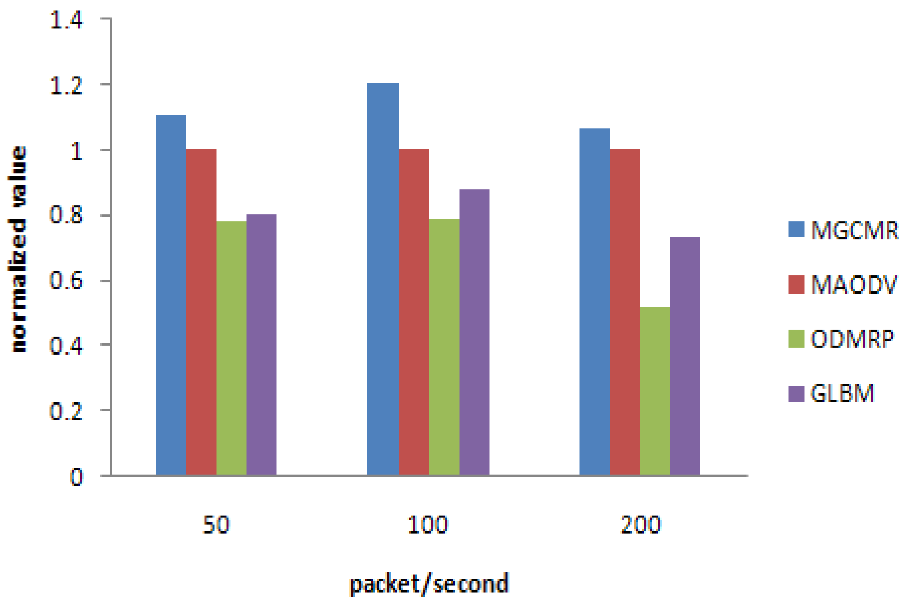


Figure 5.10: Average throughput of 100-nodes WMNs

The total overhead size is considered as the total probing packets in bytes. More importantly, it is shown in Table 5.2 that in two network sizes, a small network with 50 nodes and a larger one with 100 nodes, the probes overhead caused by MAODV, GLBM, ODMRP and

MGCMR increase, from 50 nodes to 100 nodes, with these percentages 127%, 68%, 47.7%, 33.3% respectively. This confirms that the probing overhead of MGCMR is more stable when changing network size and traffic load, where the maintenance cost of MGCMR stays steadily even for larger networks which is not the case in MAODV, GLBM, ODMRP. There are tradeoffs between the probing overhead, the data rate, and the throughput achieved, where a higher probing rate implies more up-to-date information of the network status and, thus, more decision making overhead. The total overhead size is very small against the total network throughput as the size of routing packet is small compared to data packet. We infer that the leaf-to-gateway update of MGCMR does not affect significantly the throughput gains. To sum up the above simulation outcomes, MGCMR is approved as efficient multicast algorithm to provide instant communication in WMNs.

Multicast algorithm/ Results	MAODV	GLBM	ODMRP	MGCMR
%Overhead-50 nodes	1.1	2.2	13.4	5.4
%Overhead-100 nodes	2.5	3.7	19.8	7.2
% Increasing rate	127	68	47.7	33.3

Table 5.2: Comparative percentage overhead for the different multicast algorithms

5.4 Conclusion

This chapter revisits multicast communication in WMNs and then proposes MGCMR as a new multicast routing algorithm for WMNs. To the best of our knowledge, MGCMR is the first algorithm that considers the gateway role of routing manager along with the multicast based communication quality aware applications. MGCMR also provides the leaf-to-gateway update mechanism to optimise the multicast communication. The results show that MGCMR significantly outperforms its counterparts, including the well-known ODMRP, MAODV and

GLBM multicast routing algorithms, in terms of several performance metrics. For instance, MGCMR shows an average jitter improvement of 50% over its above counterparts. For future research, it is anticipated that MGCMR can be implemented in multi-channel based multicast algorithms, which could provide a concrete basis for a number of interesting extensions.

In next chapter, a new routing metric with the consideration of neighbourhood load is introduced.

Chapter 6

Neighbourhood Load Routing Metric for Wireless Mesh Networks

6.1 Introduction

As described in [78], mesh nodes which either lie on the centre of WMNs or situate on the shortest path to the gateway may suffer from extra overloaded traffic compared to other mesh nodes as they always forward more packets. The overloaded traffic causes uneven traffic distribution which may lead to serious problems such as high latency of packet delivery, and high packet loss rate. Different research proposals that investigate load balancing within WMNs [54] [81] [103] [111] [112] have been presented. However, the main point that is apparent in the literature is that there is a current lack of examining the load balancing during the communication phase. This is necessary to evaluate the provision of real-time applications with the aim of satisfying both service providers and users in WMNs. Devising new load balancing metrics to study the performance of a rich portfolio of real-time applications is foreseen as a vital issue for the next generation WMNs. Thus, this chapter presents a new routing metric that aims to optimise the load balancing issue in WMNs.

The routing metric is implemented as a parameter in routing protocols to judge the superiority of a path over other alternate ones. It covers a set of routing constraints such as bandwidth, network latency, path length, load balancing, and reliability. In addition, the improvement of one aspect normally results in enhancing all other aspects. For example, a communication cost routing metric may also perform well in reducing latency, etc. As mentioned in [36], new routing metrics are required to examine and improve the performance of WMNs in dealing with more constraints and challenging applications, such as real-time applications. Hence, the design of routing metrics is extremely important to improving the overall performance of WMNs.

The rest of this chapter is organised as follows. Section 6.2 presents background of load balancing and routing metrics. Section 6.3 outlines the new proposed routing metric. Section 6.4 shows the simulation setup and the results of the proposed routing metric against its counterparts and, finally Section 6.5 summarizes this chapter.

6.2 Background

The network traffic may not be distributed evenly in a WMN. Some nodes are under light traffic load (transmitting or receiving a small amount of packets), while other nodes are under heavy traffic load (transmitting or receiving a large amount of packets). An overloaded node forces the packets in this node to wait for a longer time before the transmission phase. Thus, this may lead to packet transmission delay, and packet loss. All of these problems contribute further to increase the network communication cost.

As described in Section 2.2.1, the load balancing problems of WMNs can be classified into two types, *gateway node load balancing* and *centre node load balancing*. To overcome the centre node load balancing problem, the traffic load is supposed to be evenly distributed. In other words, the main objective that must be achieved is that keeping the load over different nodes relatively equal. Existing load balancing routing metrics [54] [60] only consider the traffic load of each intermediate node on the path. Therefore, there is a need to find alternate way to optimise the centre node load balancing problem.

As an attempt to achieve this goal, this approach consists of three aspects. Firstly, the packets should travel on the path with the lowest aggregate traffic load instead of the shortest path. Secondly, an overloaded node should not be involved as a packet forwarder. Thirdly, by reducing the interference of the network, the average waiting time of transmitting a packet should be decreased, thus, the overall traffic load of all nodes are reduced. Therefore, a new load balancing routing metric is required to optimise the load distribution and reduce the interference simultaneously. For this reason, the proposed routing metric in this chapter should assist routing algorithms by selecting a path by considering the impact of neighbourhood load, i.e. the load of the region within one node's interference range. This design fulfils the communication requirements of WMN.

6.3 Neighbourhood Load Routing Metric (NLR)

In wireless communications, the packet transmissions over one node can influence all other nodes in its interference range which is inevitable due to the broadcast nature of wireless networks [113] [114]. Indeed, this interference can make the whole neighbourhood busy. In this chapter, the load balancing routing metric, namely, Neighbourhood Load Routing metric (NLR) is proposed to reduce the interference level.

Definition 6.1: The region within a transmission range of *node i* is the neighbourhood of *i*, and all nodes in this region are the neighbourhood nodes of *i*.

Therefore, the average load of each neighbourhood is measured to bypass the busy neighbourhood instead of only bypassing the busy node with Load-count routing metric. Moreover, in a heavily-loaded neighbourhood, extra traffic on one node influences the communication of all nodes within its interference range. The transmission of packets in these nodes can be deferred, or dropped. Thus, there is a side effect caused by allowing a packet goes into heavily-loaded neighbourhood. To solve the above problem, NLR is developed to check the summation value of the neighbourhood load over a path which is:

$$NLR = \sum_{i=1}^{k} \frac{Load_i^n}{b_i^n} \qquad (6.1)$$

$$n = \frac{tr}{d_{avg}} \qquad (6.2)$$

where n is the interference radius of neighbourhood in hop number; tr denotes the interference range (here the transmission range is assumed as same as the interference range), and d_{avg} is the average distance between two one-hop nodes. $Load_i^n$ refers to the load of a neighbourhood of *node_i* with a radius n hops and b_i^n is the average transmission rate of this neighbourhood. The neighbourhood size measures average range of the zone with all nodes interference by the centre node. Hence, unlike existing routing metrics, NLR considers three aspects in the selection of the best path, which are *IFQ length of each node*, *neighbourhood interference*, and *transmission bandwidth*.

Definition 6.2: The traffic flow from/towards the outside of the neighbourhood called "external neighbourhood load" is either generated from, destination to one of nodes in this region or propagated through the neighbourhood region. If the sender node, destination node and selected path are all in the neighbourhood region, the generated traffic load is called "internal neighbourhood load".

In NLR, a node exchanges messages every period of *t* with its all neighbours including the status of the traffic load and transmission range. Each neighbour records the received message in the neighbourhood routing table for future use. In the design of NLR, it is implemented to neutralize the external neighbourhood load to balance the traffic load among different neighbourhoods in WMNs. Therefore, the change of internal neighbourhood load does not affect the use of NLR. Figure 6.1 shows an example of using NLR to select the best-available path within a WMN. There are thirteen nodes where *node_S* intends to find a path to *node_D*. Suppose that the interference range of this network is one-hop, each link is with the same bandwidth, *b*. *Node_1* is sending data packet to *node_11* through *node_5* and *node_8*. Hence, each of these four nodes has traffic load *l* where l_n>0. All other nodes are not in use. Two possible paths are compared, which are path1: *node_S-node_2-node_6-node_9-node_D* and path2: *node_S-node_4-node_3-node_7-node_10-node_D.* Therefore, the NLR value of path1 is shown as $NLR_{path1} = \frac{3l}{b}$ and the NLR value of path2 is shown as $NLR_{path2}=0$. Since l>0 and $NLR_{path1} > NLR_{path2}$, path2 is selected as the best-NLR path. The selected path, in turn, generates less interference to the existing transmission between *node_1* and *node_11* by alternatively selecting relay nodes with less NLR aggregation.

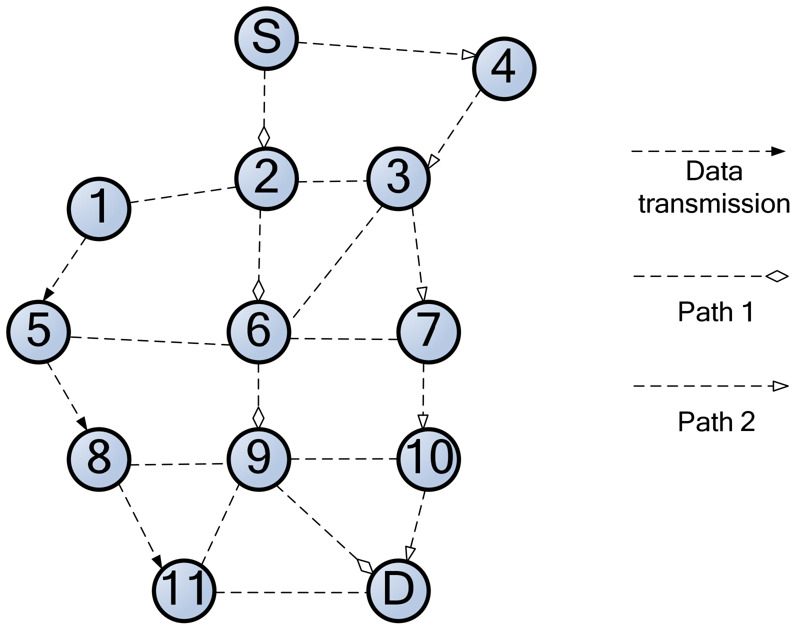


Figure 6.1: An example of using *NLR* in 1-hop neighbourhood

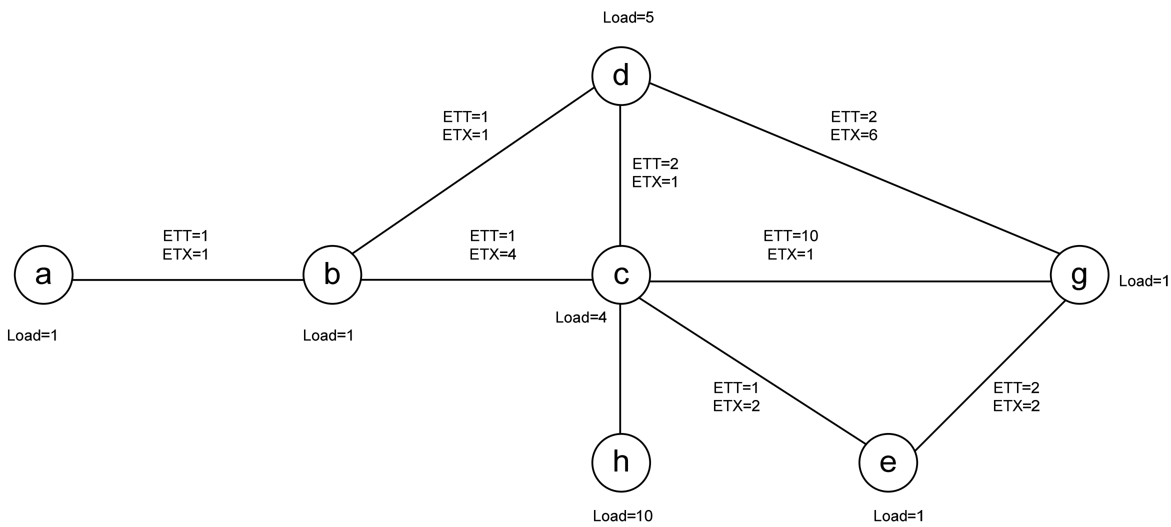


Figure 6.2: An example of different routing metrics in path selection

Figure 6.2 shows another example to illustrate different path selections with different routing metrics, including Hop-count, ETX, SPP, Load-count, and NLR. As illustrated in Figure 6.2, the network topology and each link are characterized by ETT and ETX. In addition, each node is also characterized by the traffic load. The SPP value of a path can be expressed as the product of ETT value of each link. The average interference range is one hop and the bandwidth of each node is equal in Figure 6.2. For instance, the NLR value of *node_b* is shown as NLR_b= ($Load_a$+ $Load_b$+ $Load_d$+ $Load_c$)/4= 2.75.

Table 6.1 lists five possible paths from *node_a* to *node_g*, which are *a-b-d-g, a-b-d-c-g, a-b-c-g, a-b-d-c-e-g*, and *a-b-c-e-g*. Table 6.1 also compares the routing metric values of each path under different routing metrics. The value of ETX can be also seen as the latency of the communication link. ETX and Load-count select *a-b-d-c-g* and *a-b-c-g* respectively, with the high delay value of 14 and 12. Both Hop-count and SPP select two paths, thus, there is uncertainty for the path selections with these two routing metrics. In contrast, NLR chooses the path, *a-b-d-g*, which is with the lowest latency of 4. As shown in this example, NLR is able to select the right path with the lowest transmission delay compared to other routing metrics.

Available paths	Hop-count	ETX	SPP	Load-count	NLR	Delay (sum of ETT)
a-b-d-g	3	9	1	6	5.5	4
a-b-d-c-g	4	4	1	9	9.7	14
a-b-c-g	3	6	0.25	5	6.45	12
a-b-d-c-e-g	5	7	0.25	11	11.2	7
a-b-c-e-g	4	9	0.0625	6	8.45	5
Selected path(s)	a-b-d-g a-b-c-g	a-b-d-c-g	a-b-d-g a-b-d-c-g	a-b-c-g	a-b-d-g	a-b-d-g

Table 6.1: Path selections with different routing metrics

6.4 Performance Evaluation

The simulation modelled by NS-2 [64] aims to evaluate the performance of NLR within the context of the communication quality. In the experiments, the proposed metric is compared to four other single-channel metrics, Hop-count, ETX, Load-count and SPP. Nevertheless, NLR can be converted to multi-channel routing metric like WCETT-LB. Four performance metrics, namely, *average packet delivery ratio*, *average throughput*, *average end-to-end delay* and *average jitter*, are used in the simulation.

6.4.1 Simulation Model

The simulation parameters considered in our experiments have been used widely in previous study such as [56] [57] [108]. The simulation models a network of 50 wireless routers including 45 stationary routers and 5 random mobile routers, distributed over a 500 m × 1000 m area. The mesh nodes are deployed in both grid and random topologies in our experiments. The source nodes send Constant Bit Rate (CBR) traffic with UDP as transport protocol, consisting of 1024-byte packets with a sending rate of 20 packets per second. The capability of NLR is measured by conducting experiments under different traffic load where nodes randomly generate between 10 and 40 different CBR traffic flows. Each algorithm is simulated on 10 different grid topology networks and 10 different random topology networks. Each simulation runs for 400 seconds and the average results over all topologies are presented.

In the simulation, NLR, Load-count, ETX, and SPP are implemented by modifying AODV under both WMNs with grid topology and WMNs with irregular topology. The results are presented using various link-quality metrics normalised with respect to that of the original AODV-HOP.

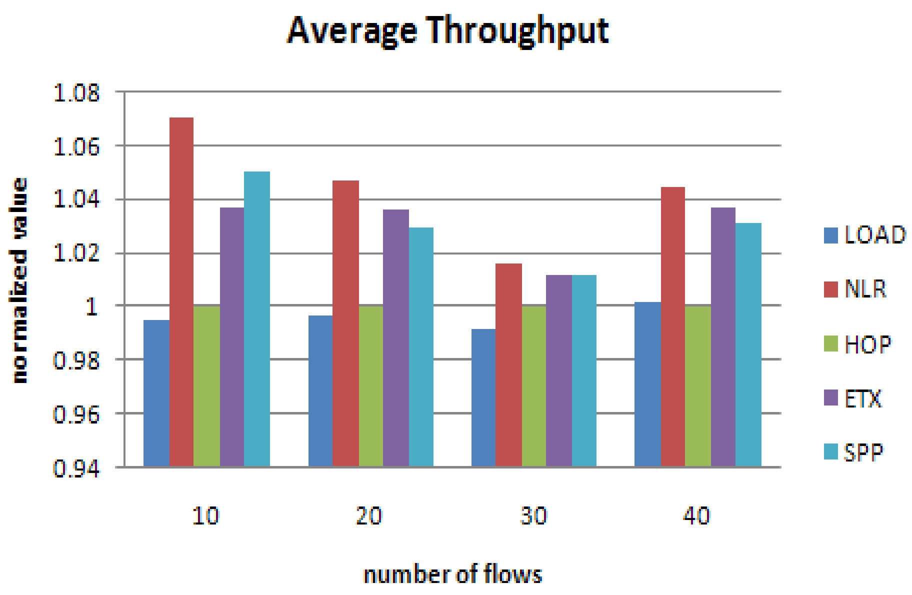


Figure 6.3: The average throughput comparison (grid topology)

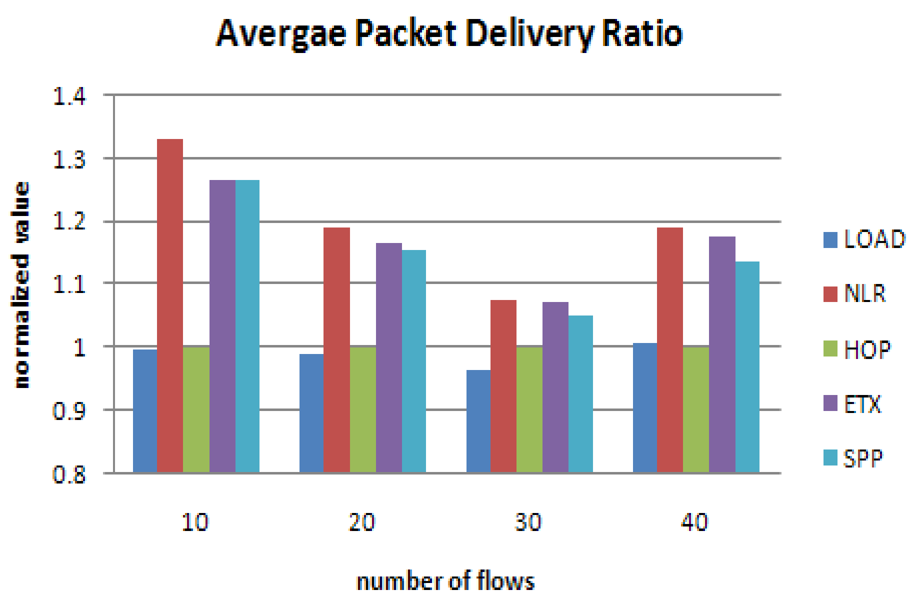


Figure 6.4: The average packet delivery ratio comparison (grid topology)

6.4.2 Simulation Results

a) NLR in a grid topology

NLR is compared with Hop-count, Load-count, ETX, and SPP in terms of average packet delivery ratio, average throughput, average end-to-end delay and average jitter within grid topology WMNs. Figure 6.3 shows the average throughput of NLR is the highest among the five routing metrics in WMNs. SPP has the second best results followed by ETX. NLR also demonstrates the highest packet delivery ratio as shown in Figure 6.4. Both Figures 6.3 and Figure 6.4 illustrate that NLR exhibits the best performance in grid topology networks within various network operating conditions.

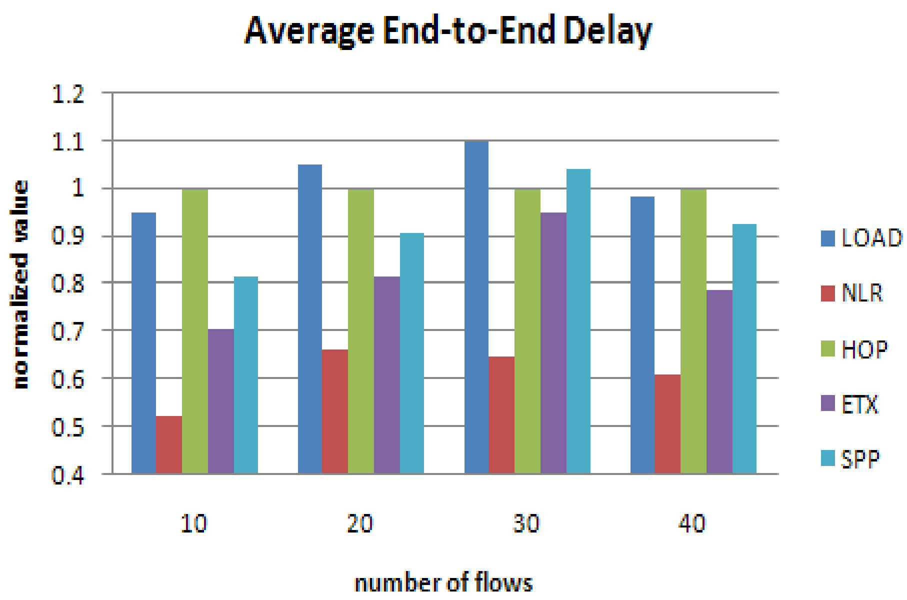


Figure 6.5: The average end-to-end delay comparison (grid topology)

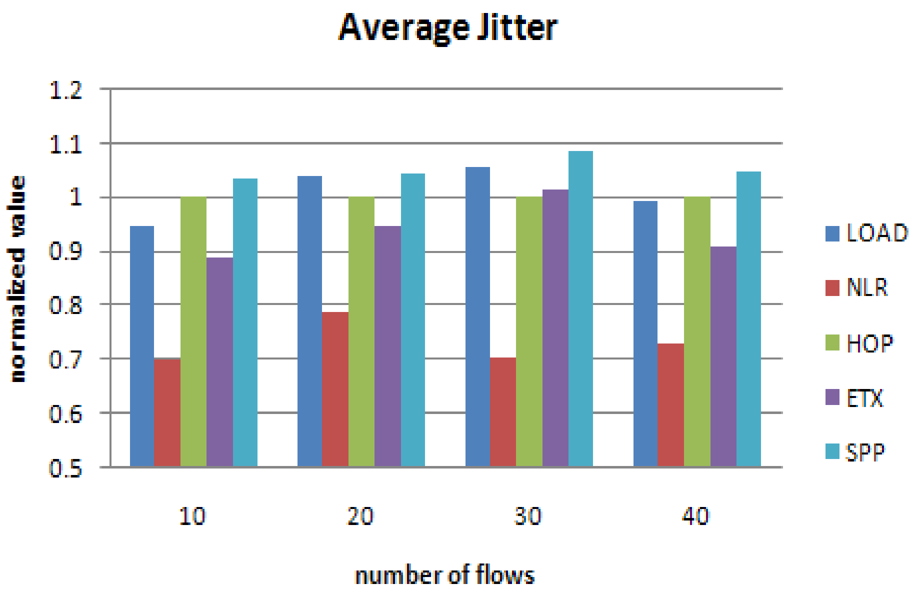


Figure 6.6: The average jitter comparison (grid topology)

In Figure 6.5, it is apparent that NLR outperforms its counterparts by achieving at least 33% lower average end-to-end delay in grid topology networks. Moreover, compared to the other four routing metrics, NLR reduces at least 29.3% of the average jitter. Undoubtedly and more importantly, these two figures confirm the superiority of NLR as a performance metric for instant communication patterns.

As we have illustrated in Table 2.7 in Chapter 2, Load-count only considers traffic load during the path selection. It is inaccurate when the number of traffic flows is low. Therefore, in the light of our above results, Load-count performs poor as Hop-count when the network is not busy enough. However, compared to Load-count, NLR captures more communication quality parameters which ensure its high performance under different traffic conditions. NLR establishes a cost-effectiveness based performance in terms of communication time, such as jitter and delay, which further proves that NLR is apposite to high communication quality demand applications, such as online gaming, wireless TV software and so forth. Unlike other routing metrics, the simulation results depict that NLR has extraordinary advantages to enhance the communication quality of AODV in WMNs with grid topology, by distributing the traffic load and reducing interference.

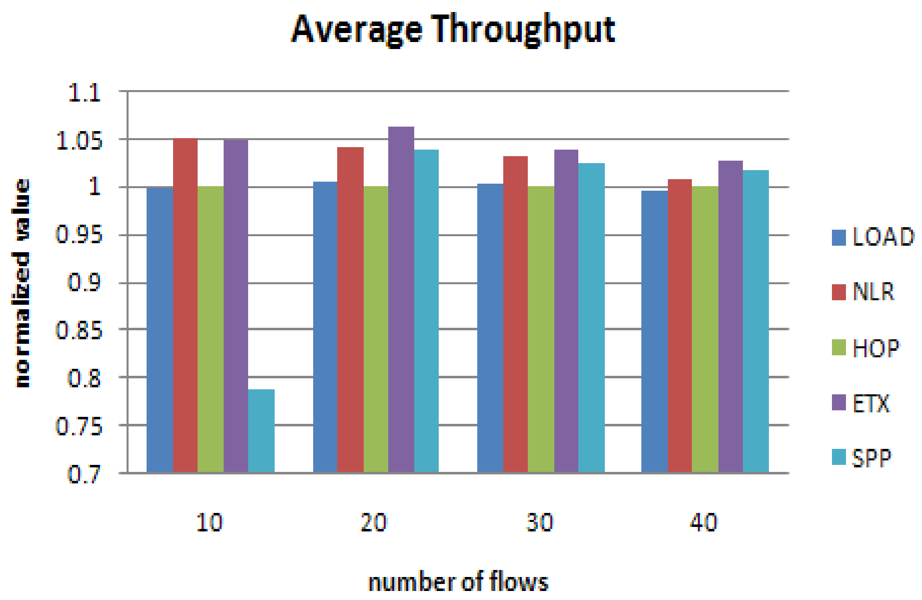


Figure 6.7: The average throughput comparison (random topology)

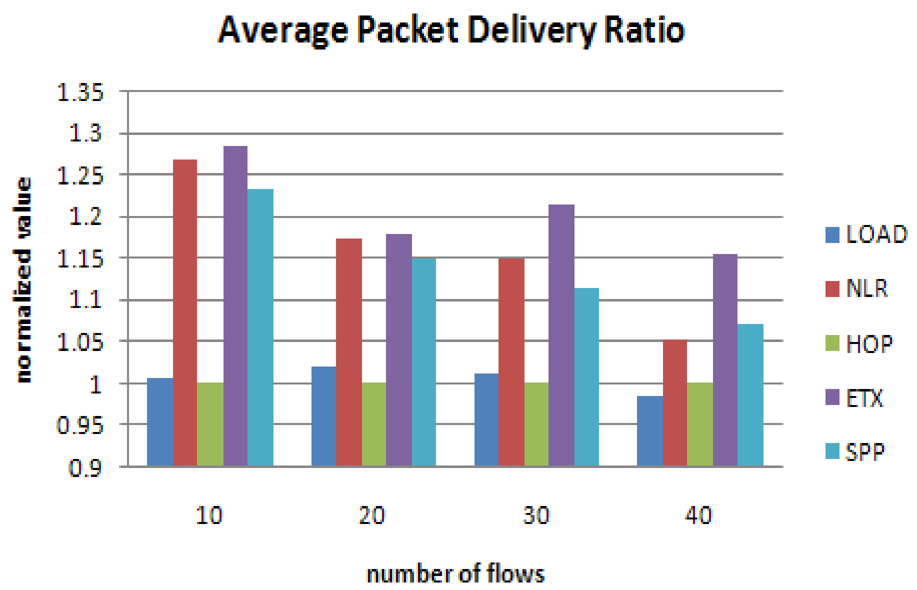


Figure 6.8: The average packet delivery ratio comparison (random topology)

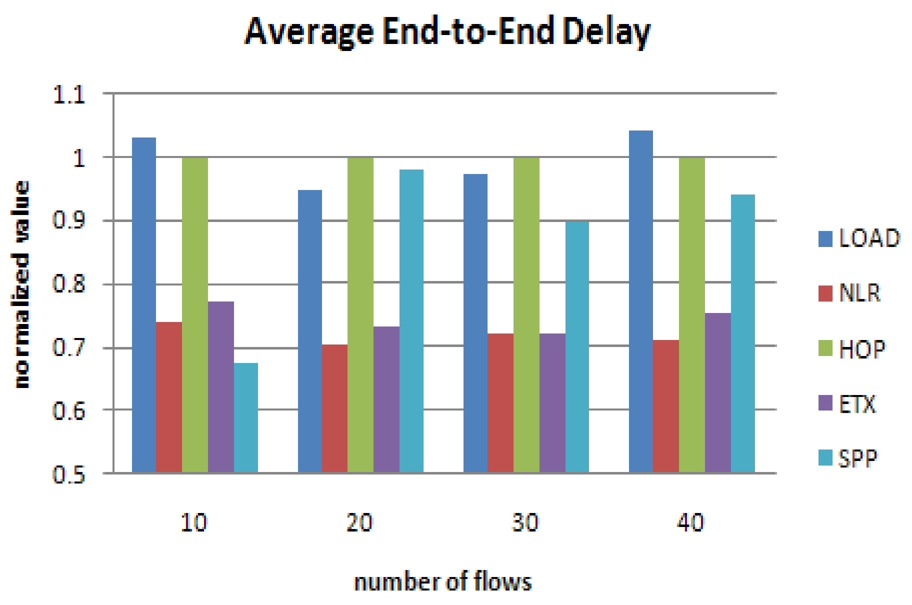


Figure 6.9: The average end-to-end delay comparison (random topology)

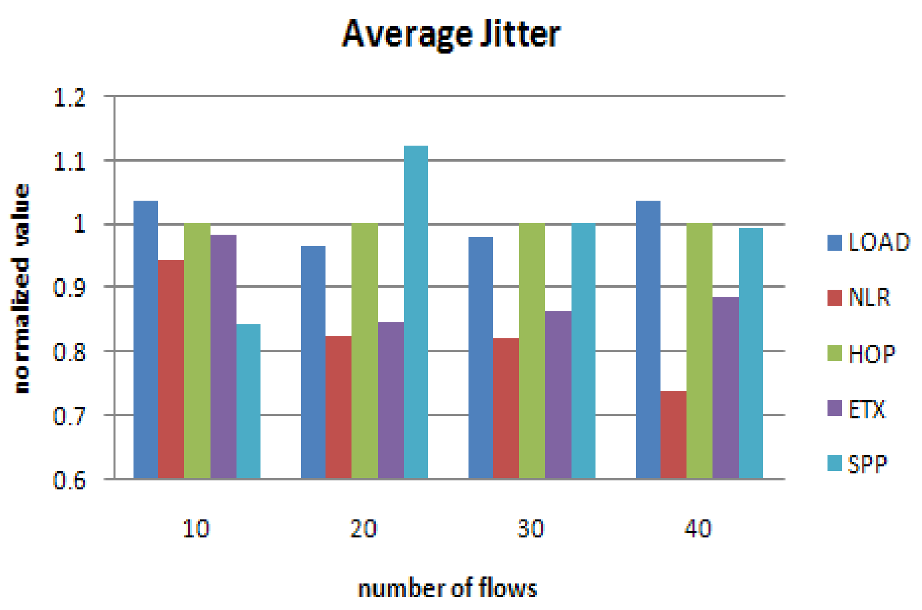


Figure 6.10: The average jitter comparison (random topology)

b) NLR in random/irregular topology

In this part, NLR is compared with Hop-count, Load-count, ETX, and SPP pertaining to average packet delivery ratio, average throughput, average end-to-end delay and average jitter in WMNs with random topology.

In Figure 6.7, the average throughput of NLR is the second place. ETX and SPP perform the best and the worst respectively in random topology networks. The reason behind this can be also found by referring to Table 2.7 of Chapter 2. The path selection of ETX mostly relies on the probe delivery rate of relay nodes. In contrast, NLR considers the average neighbourhood load of each node along the path, where it can be inaccurate due to the irregular node placement.

NLR gains a better average packet deliver rate than Hop-count, SPP, and Load-count, whereas it is slightly worse than ETX as shown in Figure 6.8. In Figure 6.9, although the lowest latency is attained by SPP within the low traffic load with 10 CBR flows, the performance of NLR improves proportional to the growth of the network traffic load. NLR achieves the lowest average end-to-end delay between 20 and 40 flows. In addition, Figure 6.10 shows the average jitter comparison, similar to the results of the average end-to-end delay where NLR demonstrates the lowest jitter in random topology networks between 20 and 40 traffic flows. Nevertheless, SPP suffers significantly from high delay and jitter.

To sum up, NLR clearly displays a distinguished performance in grid topology networks compared to its counterparts. In the random topology networks, ETX ranks first in average throughput and achieves the lowest packet loss rate. NLR is in the second place when examining these two performance metrics. ETX incurs high communication time in terms of both delay and jitter. NLR is the best metric in both average end-to-end delay and average jitter. Therefore, NLR can provide better communication quality compared to its counterparts in general. In addition, NLR outperforms the other routing metrics within the context of instant communication in both grid topology scenarios and random topology scenarios.

It is worth indicating that the good performance of NLR in grid topology networks is due to the following points.

(1) In grid topology, the average distance between each pair of two one-hop nodes is the same. The interference range of each node is set to be approximately equal as all mesh devices (802.11 enabled) are with similar transmission powers. Therefore, the node density of each neighbourhood (except the neighbourhood in which the centre node lies on the borderline) is identical. Since each forwarder *node_i* has the same neighbourhood size *n*, based on equation 6.1 and equation 6.2, the interference radius of each neighbourhood is most likely the same among the entire network. In addition, NLR in grid topology networks is more accurate and precise in the selection of a communication path.

(2) On the other hand, nodes are distributed unevenly in the random topology networks. Therefore, node densities vary among different neighbourhoods. Further, a heavy interference in high node density neighbourhood is experienced, thus degrades the overall communication quality in high node density neighbourhoods and the overall network performance as well.

6.5 Conclusion

Due to the increasing demand for high quality aware communication, load balancing becomes a key element in meeting this demand. It handles efficient communication to enhance the capability of the network. In this chapter, a new load balancing routing metric, NLR, is proposed to integrate into existing routing protocols to provide enhanced routing efficiency required to handle real-time communication applications. In this chapter, NLR is compared with Hop-count, Load-count, ETX and SPP. The results show that NLR achieves the lowest average end-to-end delay and average jitter in both grid topology WMNs and random topology WMNs. In particularly, it offers high standard communication quality provision for grid topology WMNs. In general, the results show that NLR is more suitable in all kinds of mesh networks compared to existing metrics.

In the next chapter, two QoS-aware routing metrics are designed to improve the communication efficiency of differential applications.

Chapter 7

Cross-Layer Quality-of-Service-Aware Routing Metrics

7.1 Introduction

By equipping the network with interface cards, variety devices can be connected to WMNs. This includes the personal computers such as desktop PCs, laptop PCs, and portable devices such as mobile phone, PDAs, and mesh-enabled home appliances such as broadband TVs and game consoles. For a wide range of mesh devices, the applications are designed with different level of communication demands. The high-demand applications require low latency and high data rate, e.g. real-time and multimedia streaming applications. On the other hand, low-demand applications such as business software and web browser require fairly lower communication quality.

Although there are new routing metrics [56] [57] [58] [61] [98] proposed for efficient routing, existing routing metrics still unable to fulfil the QoS requirements of some applications. The current design does not consider the characteristics of WMNs such as gateway involvement and mobility. Hence, it is important to designing routing metrics based on the characteristics of mesh network with the aim of prioritising the different services to guarantee the QoS communication, and thus, stratifying the needs of users as well as the service providers.

Within this context, a list of communication quality parameters has been considered by the previous study [50] [51] [52] [53] [54] [56] [57] [59] [60] [98] is given in Table 2.6 in Chapter 2. For example, this includes hop number, packet loss, delay, contention, traffic load, retransmission, bandwidth, packet size, MAC handshake time etc. However, none of the above parameters consider the effects of various demands of applications. This results in disastrous consequences pertaining to the delivery of QoS-aware applications. In other words, without QoS awareness, transmissions of high-demand applications may be routed in a poor path, where other transmissions of low-demand applications may be routed in a better path at the expense of delivering other QoS-aware applications. In addition, other routing metrics [60] [98] only take the multi-channel facility of WMNs into account and omit, unfortunately, the gateway capability of orchestrating the communication phase in WMN. Thus, in the next section, two routing metrics are proposed. The key point is to prioritise the delivery of packet applications based on the sensitivity to delay, considering the different requirements/demands of these applications including the QoS-aware applications.

The rest of this chapter is organised as follows. Section 7.2 provides the description of the proposed routing metrics. Section 7.3 presents the simulation results of the proposed routing metrics in comparison with existing ones. Section 7.4 concludes the chapter.

7.2 Proposed Routing Metrics

In WMNs, applications are designed with various demands. Some applications such as real-time applications are very strict in term of communication quality, while other applications such as file transfer applications have lower requirements of delay, jitter etc. The WMN technology has been the major avenue for the next generation wireless networked systems as it has the potential to lead to a disruptive change in the landscape of wireless communications. Supporting Quality of Service (QoS) to enable a rich portfolio of applications is important for the success of the next generation WMNs. However, existing standards supporting WMNs are not perfectly equipped to cater to this task. Apparently, these standards come with an inherent complexity and suffer from innate problems with respect to QoS provisioning. Consequently, care needs to be taken when developing algorithms for supporting QoS on top of the standard's mechanisms. Thus, new routing approach is necessary to tackle the challenge for truly enabling QoS in WMNs.

As a major contribution towards this objective, two new routing metrics, Packet Priority-Oriented routing metric (PPO) and Packet Priority QoS-aware routing metric (PP-QoS) are designed in this chapter.

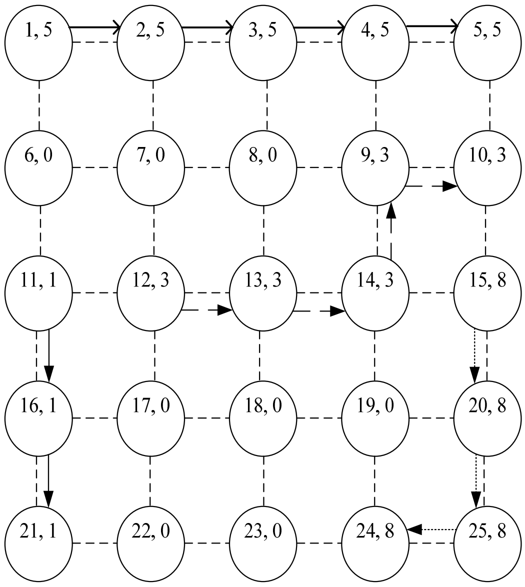
7.2.1 Packet Priority-Oriented Routing Metric (PPO)

In our scheme, the gateway assigns a real-time priority (R) to each application. The value of real-time priority is set between 1 and 9. In the implementation of the proposed routing metrics, there is only one gateway for each WMN in charge of maintaining and distributing R. When a source node intends to find a path to the destination node in order to transmit data for an application, it first checks the R value of this application locally. If the source node does not have the R value of this application, it requests this value from its gateway. Then, the gateway assigns R to this application. An application with instant communication demand is assigned a higher priority, and while an application with non-real-time demand is assigned a lower priority. The routing metrics are described below. A new metric is proposed, namely, Packet Priority-Oriented routing metric (PPO). The PPO value of a route r is obtained as following:

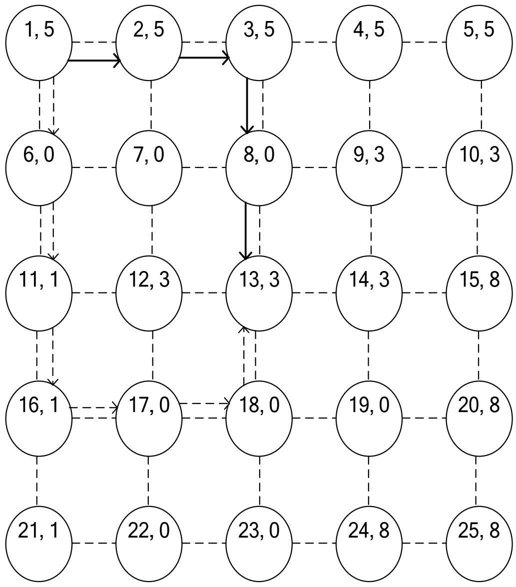
$$PPO(r) = \sum_{l \in r} R_l \ , \ \ l \neq s, l \neq d \tag{7.1}$$

where R_l denotes the real-time priority of the transmissions relayed by node l. s and d refers to the source node and the destination node of route r respectively. In WMNs, some heavily-loaded nodes may forward data packets for real-time applications with high priority. By applying PPO, this kind of nodes is avoided to be chosen as a relay node. Therefore, this mechanism ensures that a node is not involved in relaying multiple transmissions generated by real-time applications.

Figure 7.1 shows an example of how PPO is applied in the path selection compared to Hop-count. In Figure 7.1(a), each node is represented with *node id* and *the current aggregation of R*. Apparently, there are four existing transmissions from four applications. They are *node_1* to *node_5* (1-2-3-4-5), *node_11* to *node_21* (11-16-21), *node_12* to *node_10* (12-13-14-9-10), and *node_15* to *node_24* (15-20-25-24). The real-time priorities of these transmissions are 5, 1, 3 and 8, respectively. In *node_1*, an application whose *R* value is 2 intends to find a path to *node_13* as shown in Figure 7.1(b). The shortest path (1-2-3-8-13) is chosen if Hop-count is applied. In this case, the nodes forwarding data packets are also currently involved in relaying packets for other applications with *R* valued 5, 5, and 0, respectively. The PPO value of this Hop-count path is 16. Therefore, when the packets propagate along the path of Hop-count, the transmission from *node_1* to *node_5* may suffer from higher latency and higher packet loss rate as *node_2* and *node_3* relay packets of two applications simultaneously. Compared to the Hop-count path, a longer path (1-6-11-16-17-18-13) is selected while applying PPO. The PPO value of this route is as low as 12.



(a) Four transmissions among 25 nodes.



(b) How a path is chosen by Hop-count and PPO

Figure 7.1: Using PPO in AODV

In contrast, PPO selects a path with a smaller cumulative sum of *R* which ensures that the selected path bypasses the nodes engaged with transmitting packets for high-demand applications. In other words, suppose *node_a* is forwarding packets for the real-time applications (high-demand) and *node_b* is forwarding packets for the low-demand applications. PPO ensures that *node_a* is not chosen as a relay node of a path for another high-demand application when *node_b* exists as a possible relay node.

7.2.2 Packet Priority QoS-aware Routing Metric (PP-QoS)

To improve the performance of PPO, the channel usage and engagement level are taken into consideration during the selection of a path. Sometimes, a node may experience overloaded traffic due to relaying data packets for multiple low priority applications. Therefore, it is necessitated to avoid a node to play the forwarding role for too many low-demand applications in order to prevent traffic congestion. By reforming Interference-Aware Routing metric (IAR) [59] with the real-time priority factor, a new routing metric is proposed to avoid the overloading problem in the path selection, namely, Packet Priority QoS-aware routing metric (PP-QoS). The combination of PPO and IAR, PP-QoS of a route *r* is shown as below:

$$PP\text{-}QoS(r) = \sum_{l \in r} ((1-p) \times IAR(l) + pR_l)\ ,\ \ l \neq s, l \neq d \tag{7.2}$$

where *p* is a tunable parameter subject to $0 \leq p \leq 1$ and $IAR(l) = \frac{1}{1-a_{ub}} \times \frac{S}{B}$ [59]. In this routing metric, a path with a lower value of PP-QoS represents a better communication quality path. The IAR part of PP-QoS is involved to determine the channel busyness level of a node so as to avoid the overloading situation.

A unicast example shows the benefits of PP-QoS in Figure 7.2, in which a real-time application with the *R* value of 9 aims to find a path from *node_s* to *node_d*. In this example, the parameter *p* is set to 0.5. The path (s-4-d) is the best path for both Hop-count and PPO. On the contrary, another path (s-3-2-d) is chosen while applying IAR. However, the selections of these two paths do not consider both the communication demand of applications and the channel busyness level of the forwarders at the same time. If the path (s-4-d) is used to relay packets, it causes low communication performance of the real-time applications as the traffic load of *node_4* is comparatively high. In addition, a lower IAR value of *node_3* indicates that it is under light traffic load, but a high *R* value of this node also indicates that the current transmissions relayed by this node are related to a real-time application. Since allowing a node to be a forwarder for multiple high-demand applications creates traffic congestion, the path (s-4-d) is not the best path for providing high communication quality in this example. Hence, the path (s-1-2-d) is selected by PP-QoS. This avoids a busy node that may act as a forwarder for multiple real-time applications, e.g. busy nodes such as *node_3* and *node_4* are avoided as relay nodes by PP-QoS.

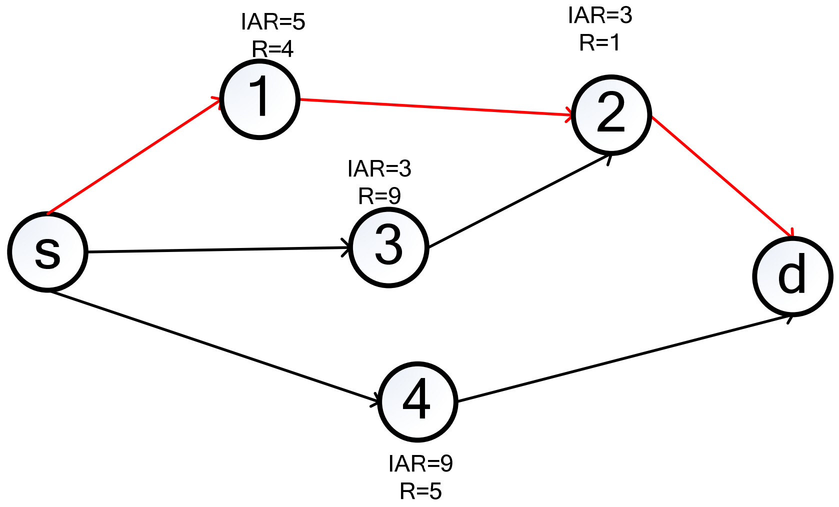


Figure 7.2: A unicast example of PP-QoS

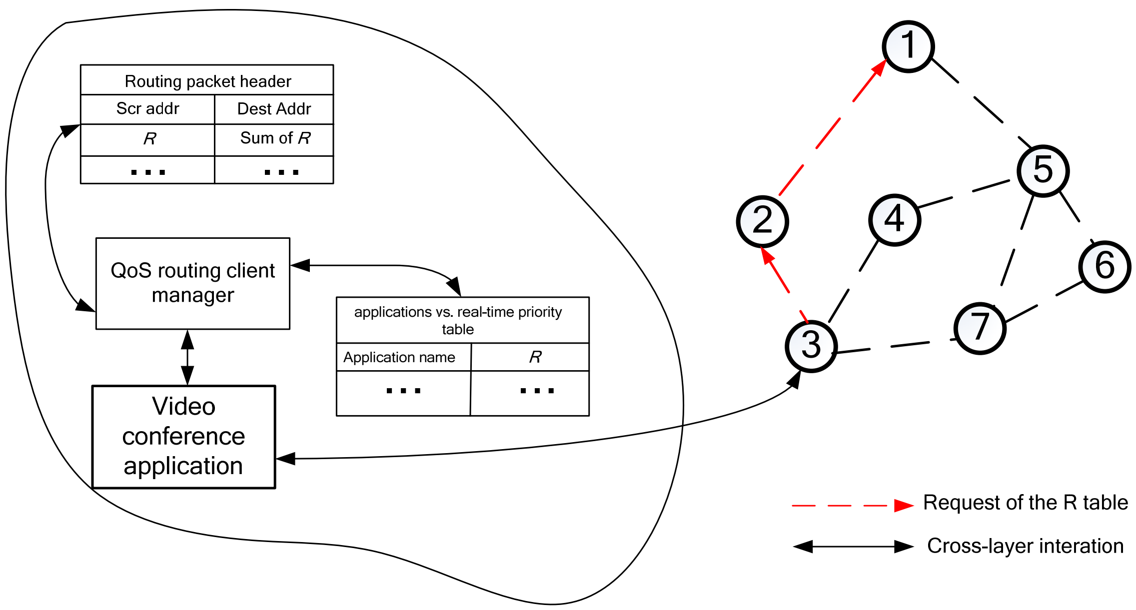


Figure 7.3: The implementation of PPO and PP-QoS

7.2.3 Implementation

From the implementation point of view, this section describes the distribution mechanism of the real-time priority for both PPO and PP-QoS. The gateway node maintains the application vs. real-time priority table which contains the network communication based applications and their corresponding real-time priorities, i.e. the communication quality demand level of an application. The gateway node is in charge of distributing the real-time priority to the mesh nodes. For example, gateway can download the real-time priorities of applications from the Internet service provider. When a mesh node joins the WMN, it is required to download the application vs. real-time priority table from the gateway. Besides, it is also requested to update the table whenever is required, i.e. if there is new version of the application vs. priority table in the gateway.

When the application of a mesh node intends to send data packets to the destination node without a valid path, it first checks whether it has a valid corresponding real-time priority for this application from the local application vs. real-time priority table. If so, the local priority value is set to the *R* field in the route request packet for discovering the new path. Otherwise, the *R* field is set to 1 as default. After setting the sum of *R* field of the route request packet to zero, route request packet is sent out.

An example of the metric implementation has been shown in Figure 7.3. Gateway (*node_1*) maintains the application vs. real-time priority table by downloading/updating it from the service provider side. When *node_3* joins the mesh network, the QoS routing client manager of *node_3* downloads the application vs. real-time priority table from *node_1*. Assume a video conference application of *node_3* intends to transmit data to *node_6*. This application passes the routing request to the QoS routing client manager in the network layer to check the

R value in the local application vs. real-time priority table. After this, the QoS routing client manager passes the *R* value to the network layer and assigns *R* in the route request.

By modifying the request packet, both PPO and PP-QoS can be compatible with existing protocols such as HWMP [50], AODV [43]. When a node receives a route request packet such as RREQ of AODV, it increments the current sum of *R*, channel busyness level (if PP-QoS) in the packet header. When the route request packet arrives at the destination, the best-metric path is selected depending on the combination of the sum of *R* and the channel busyness level (if PP-QoS). Then a route reply packet with the real-time priority of the application is sent back to the source node along the best-metric path. Each forwarder node increments the local current sum of *R* by the *R* value obtained from the route reply packet, and records the real-time priority, source node, and destination node in the forwarder table. Once the route reply packet arrives at the source, the source, in turn, records the best-metric path and starts sending data packets. When the path is no longer occupied by an application, the relay nodes are notified. Each relay node decrements the current sum of the real-time priority by the *R* value of this application only if it is forwarding packets for an application with the same source and destination found in the forwarder table.

7.3 Performance Evaluation

7.3.1 Simulation Model

The performance of the proposed routing metrics is evaluated by simulation experiments using the NS2 simulator [64]. In our experiments, the aim is to examine the capability of the proposed two routing metrics in handling QoS-aware real-time applications. Because both PPO and PP-QoS are single channel routing metrics, these two metrics are compared with three other single-channel metrics, which are Hop-count, ETT and IAR. To do so, these metrics are implemented with AODV. In addition, both PPO and PP-QoS can be converted to multi-channel routing metric like IAR. The performance metrics, *average throughput* and *average end-to-end delay* are used in our evaluation. They are common performance metrics to examine the efficiency of a routing metric. Besides, the average end-to-end delay particularly examines the performance of the proposed routing metrics within the real-time applications in order to determine whether PPO and PP-QoS can provide better communication quality to these applications.

Network size	50 nodes over a 1000 m × 1000 m area 45 fixed nodes and 5 mobile nodes
Gateway location	(500,500)
Router transmission power	20 dBm
Interference range	500 m
Transmission rate at physical layer	54 Mbits/s
Physical layer protocol	PHY802.11g
Packet size (excluding header size)	1024 bytes
Queue size at wireless routers	50 Kbytes
Traffic model of sources	Constant bit rate (CBR)
Number of CBR senders	30
CBR sender's rate	50 packets/s
Simulation duration	400 seconds

Table 7.1: Simulation parameters

Like the previous setup in [56] [57] [98] [108], the simulation models a network of 50 wireless routers distributed over a 1000 m × 1000 m area. There is one gateway node placed in the middle of the network. Beside 45 stationary nodes, there are also 5 mobile nodes with various speeds during the simulation to create real scenarios of community mesh network. This also evaluates the performance of the metrics in handling mobility. The mesh networks are simulated with both grid topology node deployment and random topology node deployment. As shown in Table 7.1, the sources send Constant Bit Rate traffic (CBR) over

User Datagram Protocol (UDP) as transport protocol, consisting of 1024-byte packets with a sending rate of 50 packets per second. The interference range is set by default 500m. Therefore, the packet size and the data rate are large enough to examine whether the routing metrics are capable of handling the applications with high-bandwidth demands. For each comparison, each algorithm is simulated on 10 different WMNs with grid topology and 10 different WMNs with random topology. Each simulation runs for 400 seconds and the average results are presented.

In what follows, the performance of different metrics is studied overall (for all transmitting applications) under heavily-loaded WMNs with grid topology in a). Then the results are also shown and discussed for heavily-loaded WMNs with random topology in b). In c) and d), WMNs with light traffic load are set up to examine the performance of the routing metrics. For each section from a) to d), the performance of high-demand applications are also examined. Finally, in e), PP-QoS is compared with IAR under different level of traffic load.

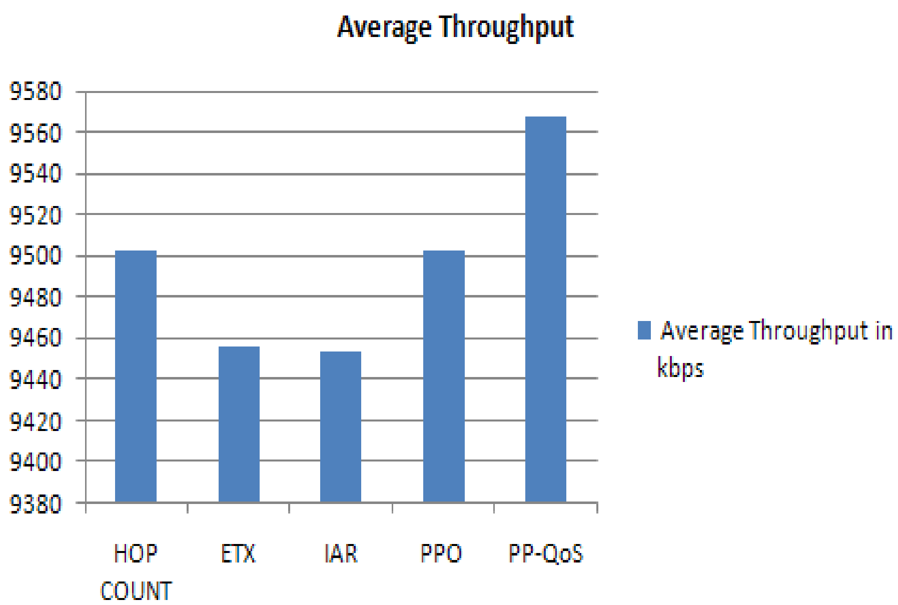


Figure 7.4: Average Throughput of heavily-loaded WMNs with grid topology

7.3.2 Simulation Results

a) The grid topology WMNs under heavy traffic load

In a) and b), there are 30 nodes randomly chosen, where each node is with an application, transmitting data packets in order to simulate a busy (congested) network. This tests the performance of PPO and PP-QoS in the heavily-loaded scenarios. In Figure 7.4, the results show the average throughput of PP-QoS is the highest among all the other routing metrics in grid topology WMNs. This means, even in heavily-loaded networks, PP-QoS outperforms other metrics. In contrast, PPO performs badly with respect to the average throughput since it does not consider the effects of traffic overload generated by the low priority applications.

Figure 7.5 illustrates the superiority of PP-QoS over its counterparts in terms of the average end-to-end delay. This further implies that PP-QoS can provide low latency for variety communication demands in grid topology networks. Figure 7.6 particularly shows the average end-to-end delay of the transmissions generated from the high-demand applications in heavily-loaded WMNs with grid topology. In this case, only the transmissions from applications with the *R* value, greater than 5, are considered as high-demand communication. PP-QoS is observed as attaining the lowest latency for high-demand applications with 22% less delay than the second best routing metric, IAR. This demonstrates the superiority of PP-QoS in providing the routing paths to the real-time application.

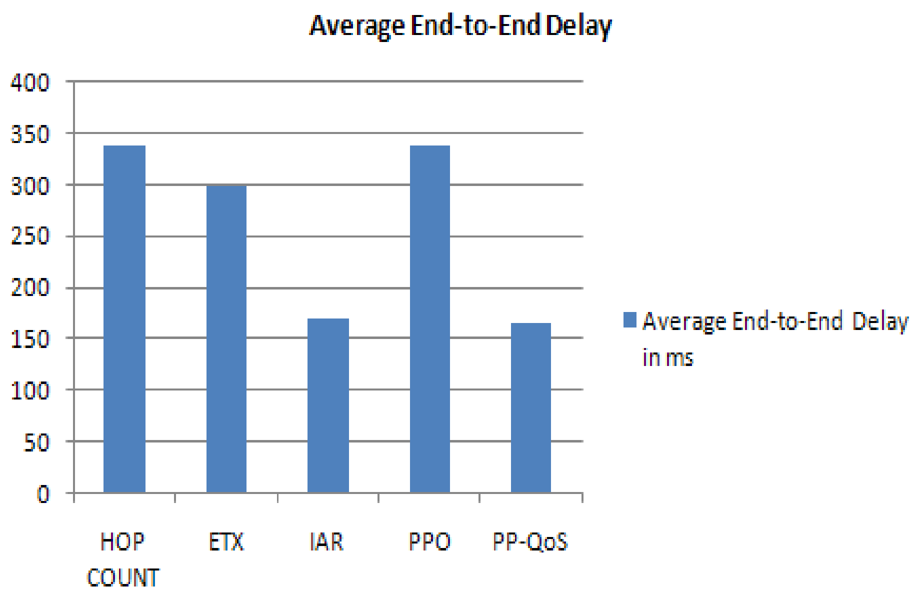


Figure 7.5: Average End-to-End Delay of heavily-loaded WMNs with grid topology

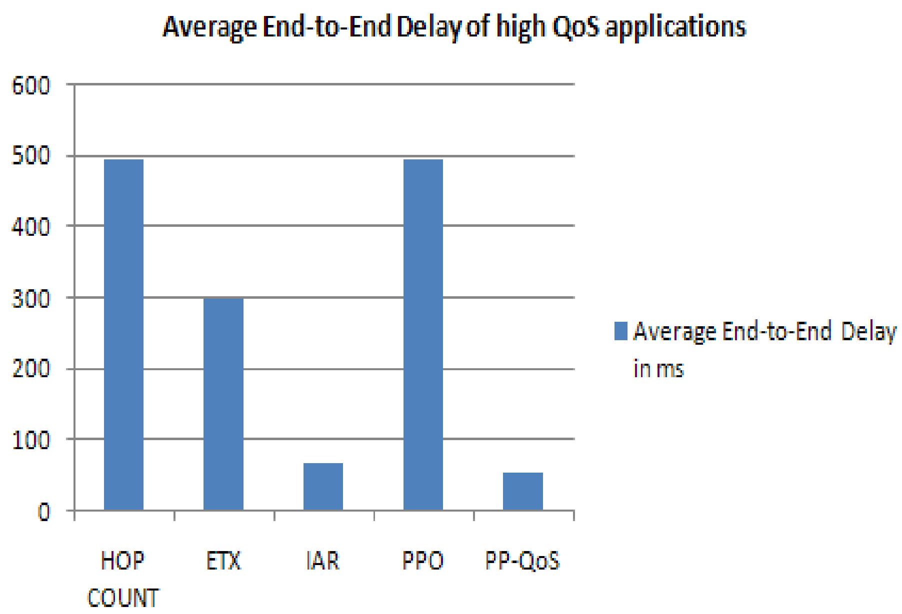


Figure 7.6: Average End-to-End Delay of high-demand applications in heavily-loaded WMNs with grid topology

b) The random topology WMNs under heavy traffic load

Figure 7.7 depicts that PP-QoS provides the highest throughput among these metrics in WMNs with random topology, while PPO ranked second. The results of average end-to-end delay are also shown in Figure 7.8, where PP-QoS is ranked as the second best overall performance among the routing metrics. The advantage of PP-QoS for high-demand applications is clearly shown in Figure 7.9. PP-QoS reduces the communication latency for high-demand applications about an average of 10 times. Thus, both PP-QoS and PPO outperform other routing metrics and both can better provide efficient services under the umbrella of high-demand applications within heavily-loaded WMNs.

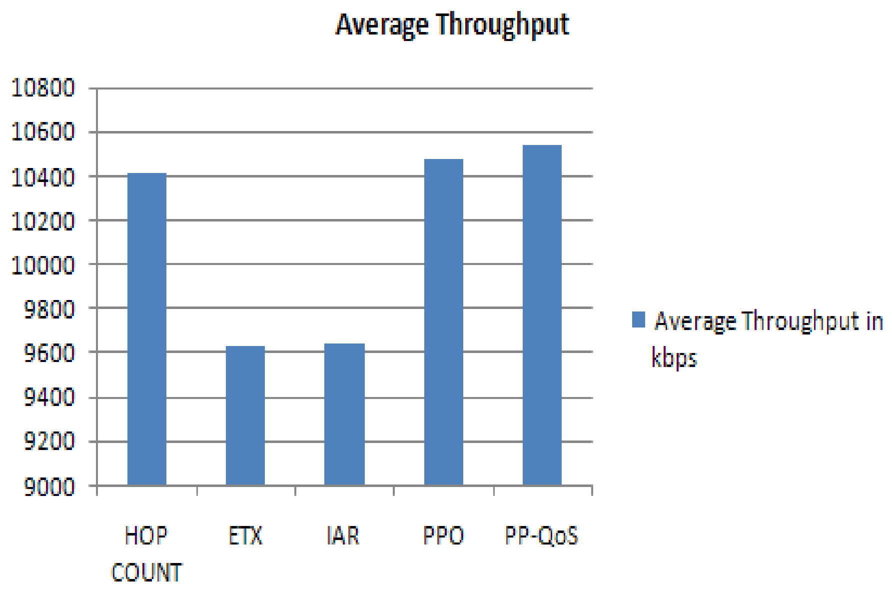


Figure 7.7: Average Throughput of heavily-loaded WMNs with random topology

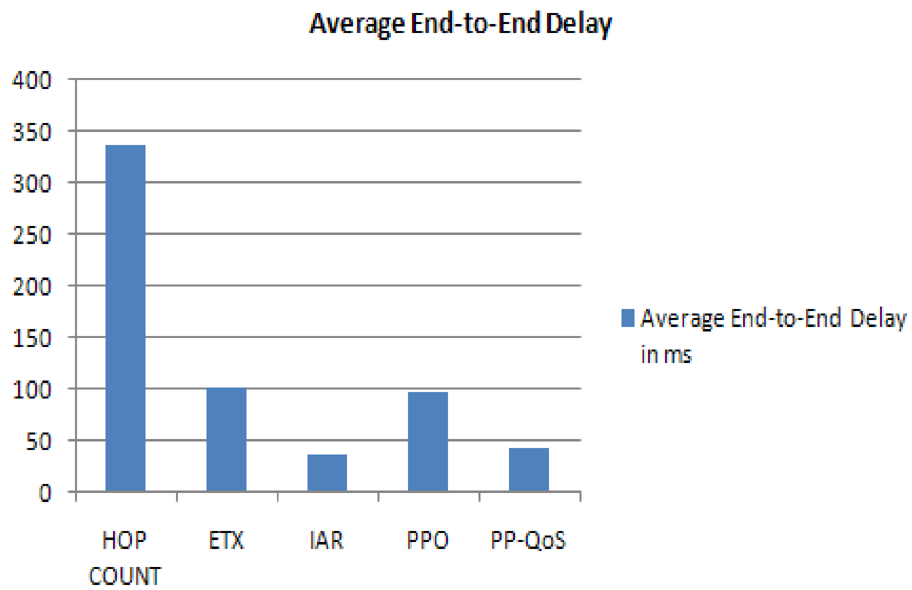


Figure 7.8: Average End-to-End Delay of heavily-loaded WMNs with random topology

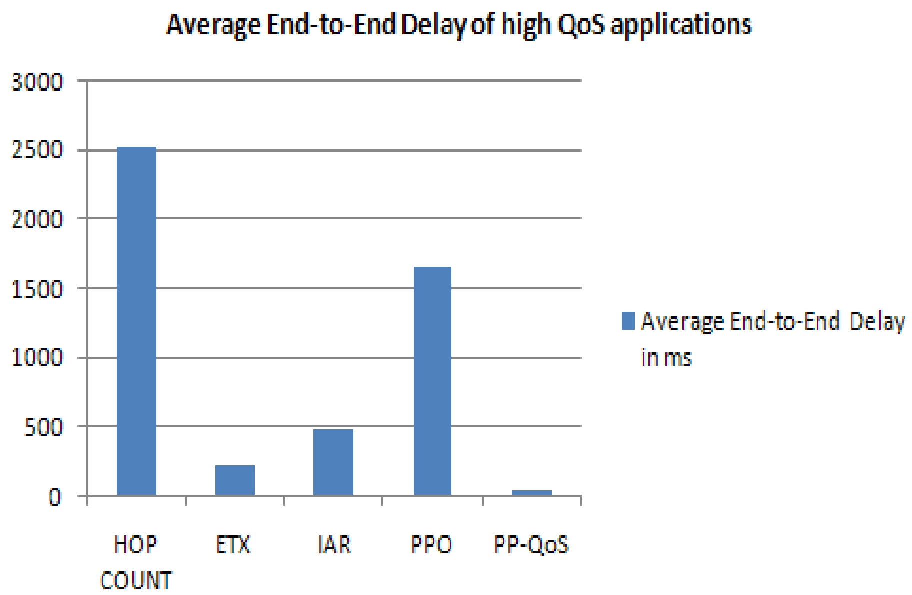


Figure 7.9: Average End-to-End Delay of high-demand applications in heavily-loaded WMNs with random topology

c) The grid topology WMNs under light traffic load

In this subsection, the performance of routing metrics is compared in the scenarios of lightly-loaded WMNs within grid topology. There are 10 different nodes deployed to generate only 10 CBR flows randomly, so as to set up lightly-loaded scenarios. As shown in Figures 7.10, 7.11 and 7.12, PP-QoS is ranked third in terms of the average throughput and the second best within the other two performance metrics. This is because, in a lightly-loaded network, there are only a limited number of nodes relaying packets due to few applications that transmit data. Consequently, the amount of R value of the possible relay nodes is relatively low to select a good path by using PP-QoS. In contrast, IAR is more sensitive for path selection in the lightly-loaded WMNs with grid topology.

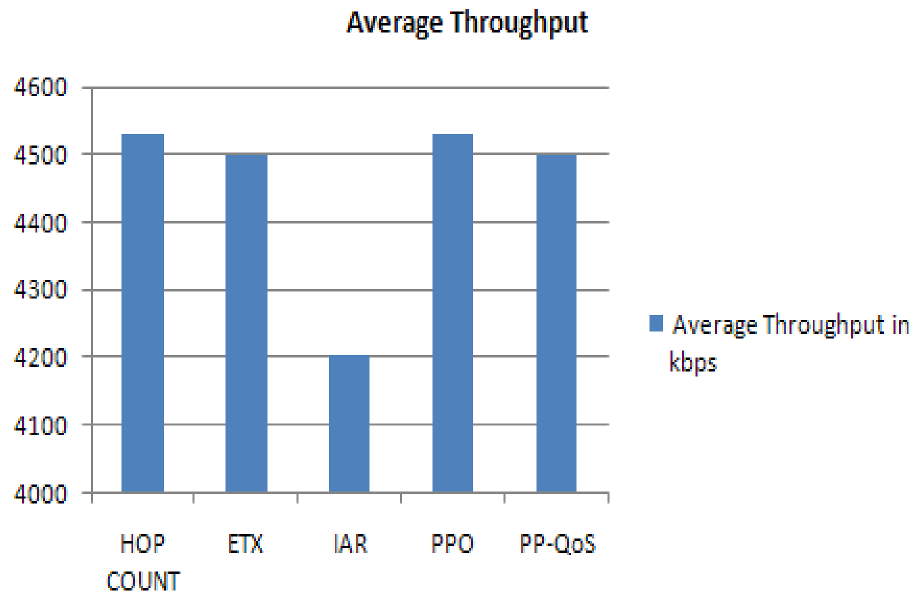


Figure 7.10: Average Throughput of lightly-loaded WMNs with grid topology

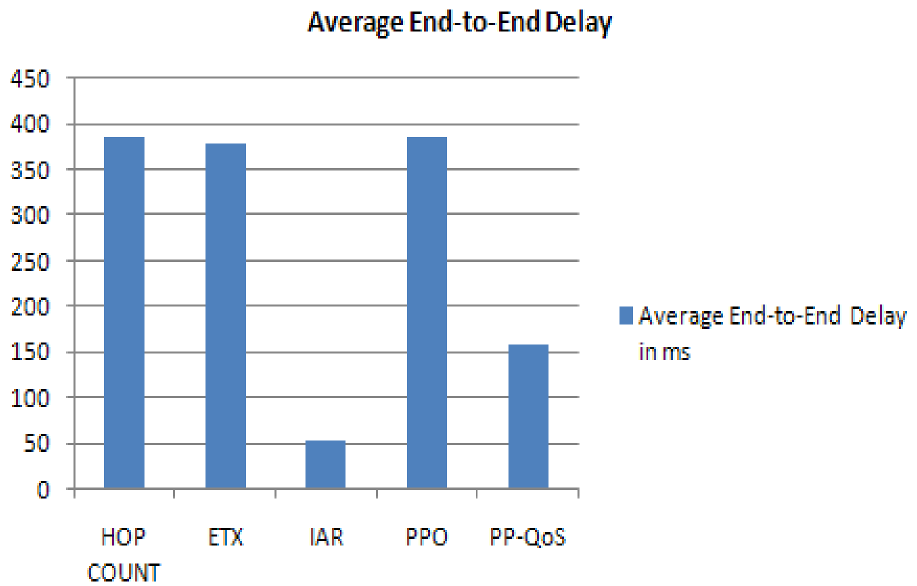


Figure 7.11: Average End-to-End Delay of lightly-loaded WMNs with grid topology

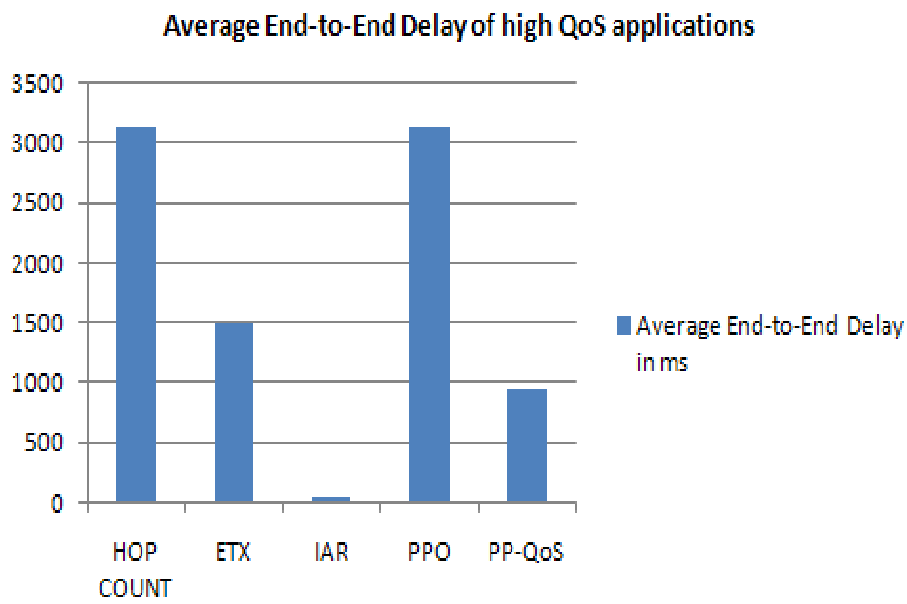


Figure 7.12: Average End-to-End Delay of high-demand applications in lightly-loaded WMNs with grid topology

d) Random topology WMNs under light traffic load

To study the performance of metrics in lightly-loaded networks, nodes are placed randomly to construct irregular topology networks under light traffic load (10 CBR flows). Figure 7.13 shows the overall average throughput result of PPO is the best in the lightly-loaded networks. In addition, PP-QoS still acquires the lowest latency with saving at least 11% communication time as shown in Figure 7.14. In addition, IAR ranked the best and PP-QoS is the second best with respect to the delay time for high-demand applications. For the same reason of c), PP-QoS and PPO do not always achieve the best performance in the lightly-loaded networks. However, the results still show the good adaptability of PP-QoS and PPO, where they still provide efficient routing even in lightly-loaded networks.

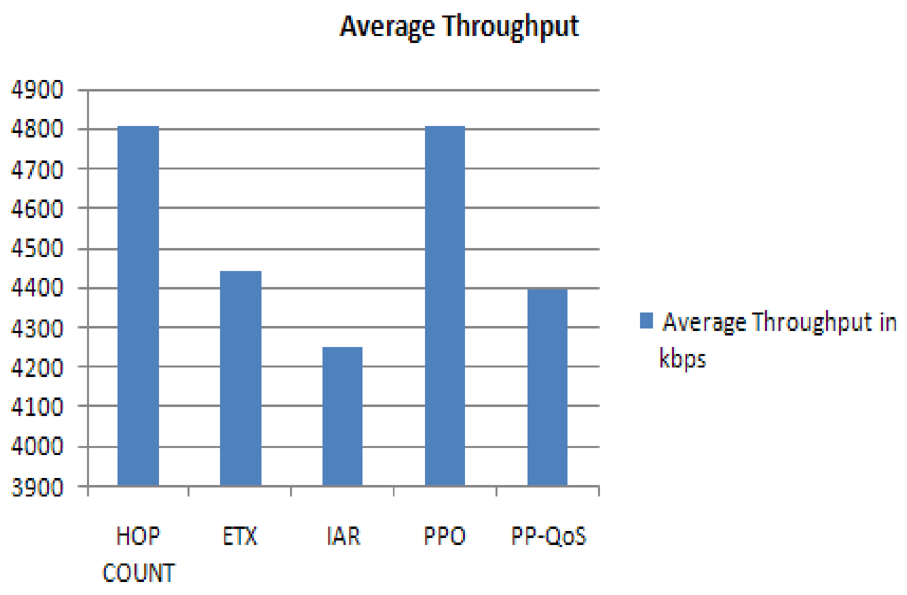


Figure 7.13: Average Throughput of lightly-loaded WMNs with random topology

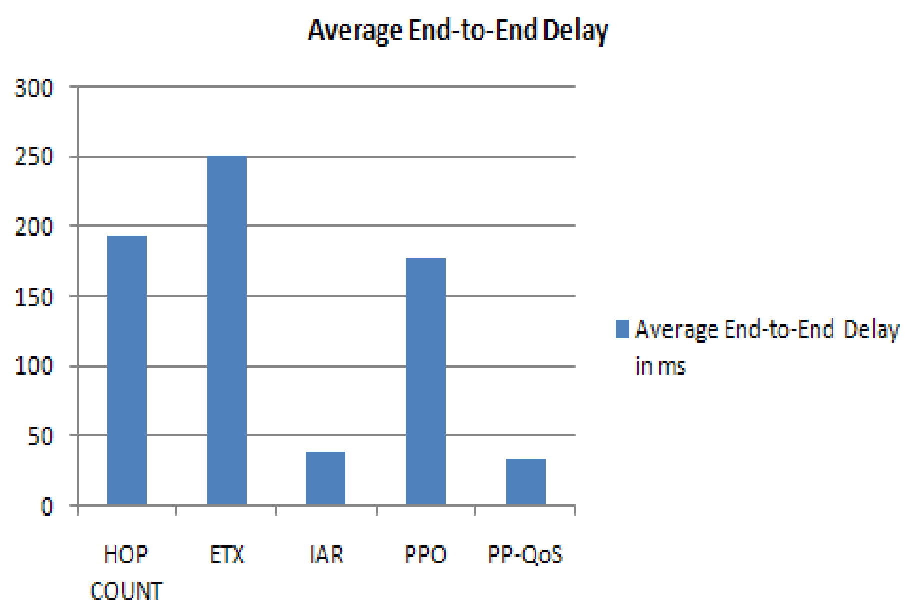


Figure 7.14: Average End-to-End Delay of lightly-loaded WMNs with random topology

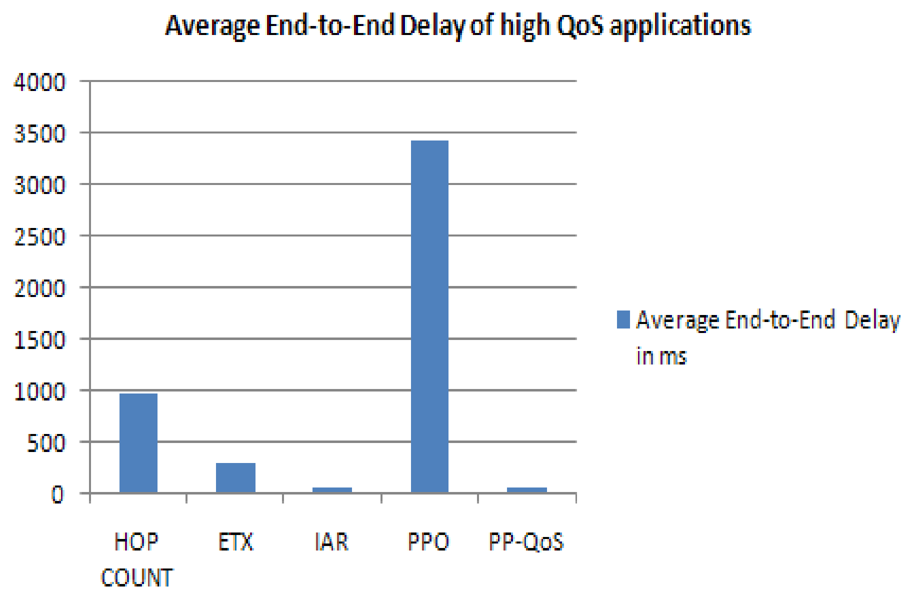


Figure 7.15 Average End-to-End Delay of high-demand applications in lightly-loaded WMNs with random topology

e) PP-QoS versus IAR

The PP-QoS ranks second among all the routing metrics in lightly-loaded networks. However, as shown in a) and b), PP-QoS achieves the best performance in almost every experiment in heavily loaded networks. Our experiments indicate that the performance of PP-QoS is improved in the busy network situations. Notice that, PP-QoS consists of two parts, the real-time priority and IAR. While the real-time priority part is devoted for selecting the suitable paths based on the application priority, the IAR part assists in evaluating the channel busyness level of nodes. When more concurrency transmissions coexist, there are more nodes that act as forwarders for real-time applications with the aim of enabling PP-QoS to select the best available path by using the current sum of real-time priority value on each node. On the contrary, the fewer nodes relay packets for real-time applications in a lightly-loaded network. In this case, PP-QoS almost selects a path based solely on the value of IAR part, resulting on the selection of non-optimal paths.

This subsection compares PP-QoS with IAR. The above arguments are supported by the results of illustrated in Figures 7.16, 7.17 and 7.18.

Simulations are conducted with the CBR flows varied between 5 and 30 to examine the behaviour of routing metrics under difference traffic flows. As shown in Figure 7.16, PP-QoS performs better than IAR as the number of CBR flows increases in average throughput. Figure 7.17 shows, in average end-to-end delay, although the initial results of PP-QoS on lightly-loaded scenarios are not as good as that of IAR, the performance of PP-QoS approximates to IAR as the number of CBR flows increases. Obviously, Figure 7.18 confirms that PP-QoS maintains much lower average end-to-end delay for high-demand applications than IAR as the number of CBR flows grows. In other words, PP-QoS offers lower latency for real-time applications than IAR. Thus, this further demonstrates that PP-QoS is QoS-aware routing metric, capable of handling delay sensitive applications such as real-time applications with high communication demand.

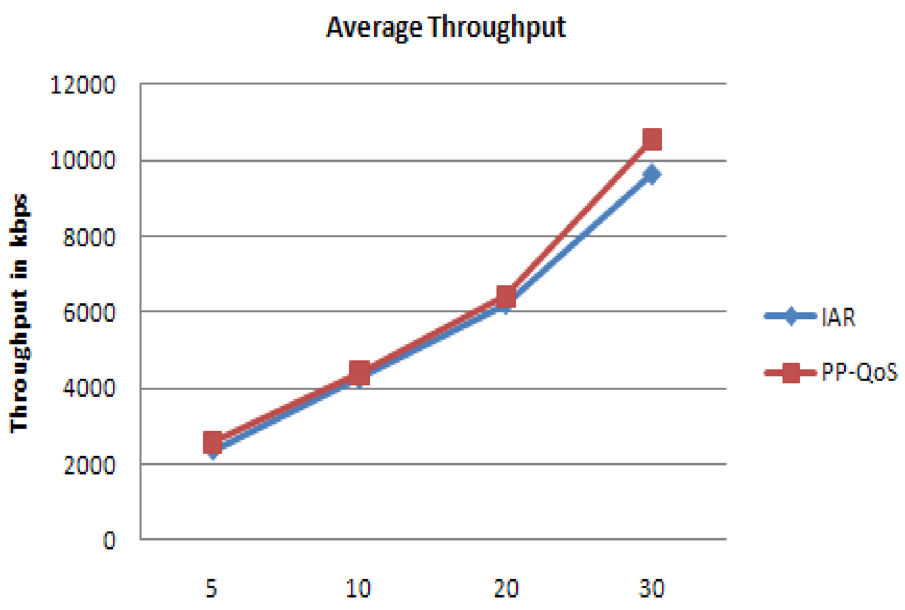


Figure 7.16: PP-QoS V.S IAR in average throughput

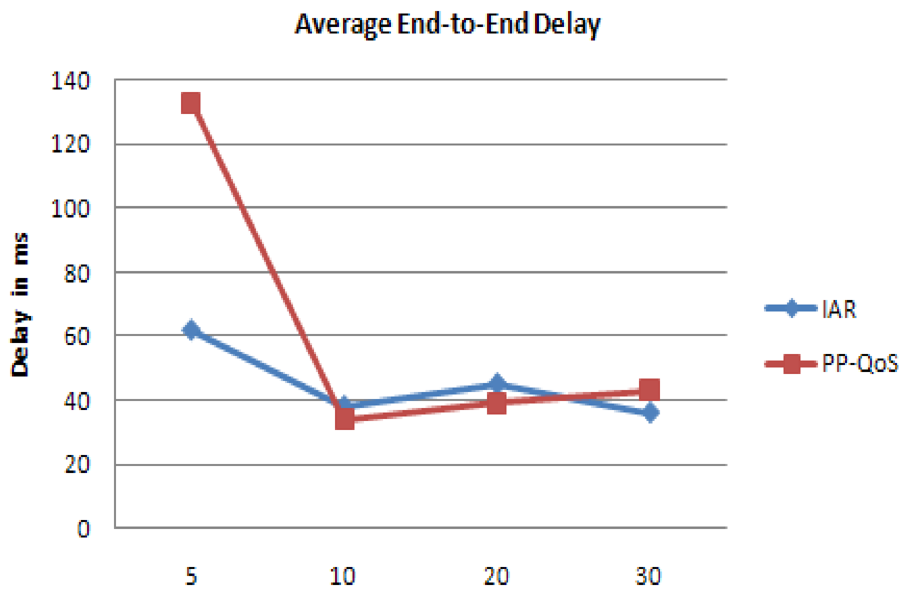


Figure 7.17: PP-QoS V.S IAR in overall average end-to-end delay of the network

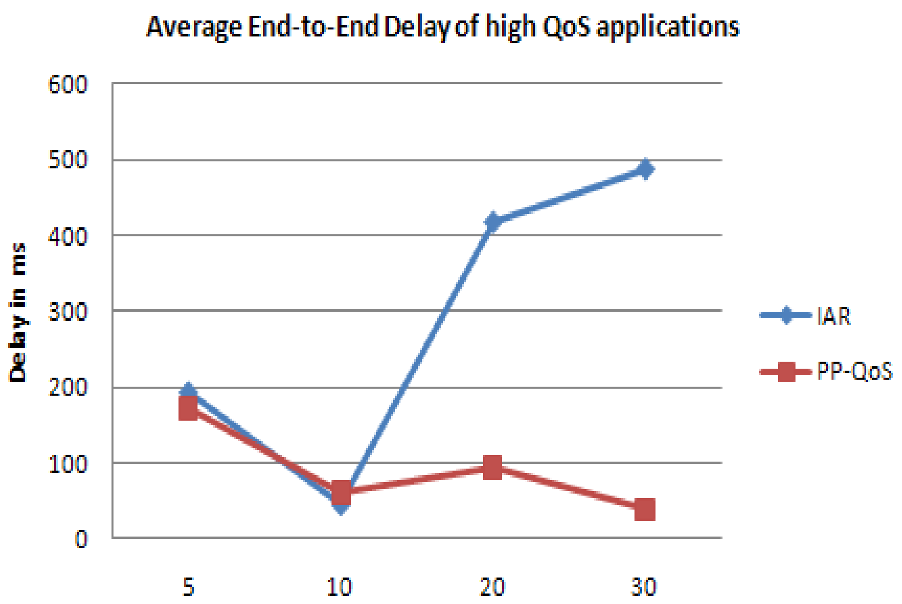


Figure 7.18: PP-QoS V.S IAR in average end-to-end delay of high-demand applications

7.4 Conclusion

There is an explosive growth of a wide range of applications carried out by the emerging wireless networks. These applications are associated with a huge demand for QoS provision in these networks, including WMNs. However, the current cutting edge standards of WMNs are not capable enough of handling the QoS-aware applications such as real-time applications efficiently. In addition, existing work such as [50] [52] [53] [54] [56] [57] [58] [59] [60] [61] [98] has not considered the prominent characteristics of WMNs to provide QoS-aware applications within the design and context of routing metrics. Hence, in this chapter, the characteristics of WMNs such as gateway involvement, QoS guarantee, and interoperability are amplified for the design of routing metric. Two routing metrics PPO and PP-QoS are then proposed to enhance the routing efficiency in WMNs to satisfy the demands of QoS-aware applications as well as providing fair communications to the other types of applications. In the design of PPO, the real-time priority (R) is introduced to quantify the communication demand level of application which is distributed by the Internet gateway. Then in the design of PP-QoS, PPO is combined with IAR to produce a cross-layer routing metric, crossing the application layer, the network layer and the MAC layer. To fully utilise the characteristics of WMNs, the proposed routing metrics enhance the gateway capability of orchestrating the communication phase within the challenging applications.

In the next chapter, the thesis is concluded and the future work is also outlined.

Chapter 8

Conclusions and future directions

8.1 Introduction

Wireless connectivity is generally considered as the most feasible way of providing cordless Internet access required by many applications found in science, engineering and a number of other fields [10] [13] [23] [26] [50] [71] [77] [87] [101]. In addition to the traditional wireless networks [4] [5] [73] [75] [115] [116], there has been substantial interest recently in the construction of Wireless Mesh Networks (WMNs) due to the distinguished characteristics [13] such as low-cost, easy-deployed etc. More importantly, WMN has shown a promising contribution to the next generation energy-efficient networking architectures. WMNs have been already deployed in many situations and there are many initiatives to deploy this technology in many cities and rural areas worldwide in the foreseeable future. For instance, the WMN technology has been deployed in the metropolitan mesh networks in San Francisco, Philadelphia and Taipei etc, and community mesh networks, in Chaska, Minnesota, Rio Rancho, and New Mexico etc [117], and rural mesh networks in South Africa [26].

The performance of WMNs highly depends on the efficiency of the underlying routing algorithms. Besides the differentiated radio power of wireless devices, the adaptability of routing algorithms becomes a major challenging issue. Routing algorithms for WMNs have been studied widely by both academia and industry due to the key role in providing high quality communication [47] [51] [57] [59] [60] [61] [67] [78] [80] [92] [94] [99] [96] [98] [103]. However, the design of these algorithms does not fully utilise the uniqueness and capability of WMN technology. Thus, these algorithms can limit the network performance and severely affect the communication quality. This already resulted in inefficient network performance when handling the challenging and new emerging real-time applications using these existing algorithmic/routing solutions.

Unicast and multicast communication patterns are among the most important communication operations to provide one-to-one communication and simultaneous transmission from a source to a set of destination hosts, respectively. Both, unicast and multicast are used in many applications such as file transfer, web browser, video multicasting, online gaming and so forth.

Nevertheless, the main point that is apparent in the literature is that there is a current lack of efficient unicast and multicast communication. Efficient communication supports the operation of real-time and QoS sensitive applications, to satisfy both sides service providers and users in WMNs. Supporting QoS-aware communications to handle real-time and QoS sensitive applications is foreseen to be vital for the success of the next generation efficient WMNs. Unfortunately, existing solutions supporting unicast and multicast based applications in WMNs are not perfectly equipped to cater to this task as these standards come with an inherent complexity and suffer from innate problems with respect to QoS provisioning as discussed in the previous chapters. More importantly, multicast communication has not been addressed so far within the context of emerging challenging high-bandwidth applications, such as real-time multimedia applications. Indeed, these efforts have not resulted in high QoS provisioning for real-time applications. Thus, there is a growing need for devising and analysing innovative multicast communication algorithms for WMNs which are capable of facilitating efficient multicast based communication applications. Therefore, this study has proposed novel routing and algorithmic solutions to handle real-time multimedia applications, in different operating conditions within various communication constraints. Indeed, this study has a significant impact and will open new scientific technological horizons towards optimal communication environment for tomorrow's communities and connectivity.

Routing metrics can be integrated in the routing algorithm including both unicast and multicast to assist the path selection preference and, thus, enhance the overall communication efficiency. However, optimising performance requirements under heterogeneous constraints is still an unexplored issue in WMNs. In other words, existing routing metrics are not capable enough to capture all constraints and performance parameters in WMNs. Thus, there is a growing need to integrate multiple cross-layer routing metrics, considering the sensitivity of the applications to the QoS and the key role of gateway nodes in WMNs, an issue that is not addressed well by existing work. Thus, this study has considered this important issue and presents new cross-layer routing metrics that have shown high efficiency when handling QoS-aware applications.

In conclusion, the main aim of this thesis is to significantly contribute to the efficient optimal operation of unicast routing, multicast routing and cross-layer design based communication in the next generation WMNs. In what follows, the main findings and contributions made by this thesis are outlined.

8.2 Summary of the Results

The major focus of this thesis has been the design of new unicast, multicast algorithms and cross-layer routing metrics for WMNs. The main contributions of this research are summarised as follows.

- In Chapter 1, we discuss the star-of-the-art of routing algorithm solutions relating to the unicast and multicast based applications within WMNs. The motivation and research problem have been outlined as well.
- In Chapter 2, we first review existing routing algorithms (including unicast and multicast) [37] [39] [44] [45] [46] [47] [48] [49] [50] [51] [69] [70] [71] [87] [88] [93] [99] [101]. These routing algorithms can be classified into three routing schemes, *reactive routing, proactive routing* and *hybrid routing*. In reactive routing, nodes broadcast the routing request over the network to discover the best path, but generate high routing overhead. Therefore, reactive routing such as AODV, DSR, GSR, GSR-PN, and AODV-CGA etc, is more suitable for highly dynamic networks i.e. the networks with high mobility nodes. In contrast, proactive routing such as DSDV, calculates paths before using them by exchanging the link changes among all the nodes. This routing is suitable for small networks with static nodes. Compared to other two routing schemes, in hybrid routing, the gateway can keep the network topology proactively maintained for each period of time and handle the reactive requests from the mesh nodes with different mobility levels. Therefore, hybrid routing is found as the most suitable routing scheme for mesh networks. Unfortunately, we observe that existing routing proposals [37] [39] [44] [45] [46] [47] [48] [49] [50] [51] [55] [61] [69] [70] [71] [87] [88] [93] [95] [99] [100] [101] including unicast, multicast and cross-layer routing metrics, do not consider the new characteristics of WMNs. In particular, these existing proposals omit the capability of the gateway to

orchestrate the routing and communication in WMNs. This is mainly because most of these proposals have been transferred from MANETs, where there are no gateways. In addition, load balancing has been lightly considered in existing work within WMNs despite the fact that it is a key element in improving the communication quality in WMNs. Therefore, this chapter explores hybrid routing algorithms within the QoS constraints.

- In Chapter 3, Gateway Centralized Multi-hop Routing algorithm (GCMR) is proposed as the first gateway based unicast algorithm for WMNs. GCMR involves the gateway as a routing orchestrating node to handle better the routing tasks. In GCMR, the leaf-to-gateway update mechanism is applied to report topology changes to the gateway. To reduce the number of update messages, the traffic prediction method and the TTL setup mechanism are also introduced. The performance results show that GCMR achieves outstanding performance compared to AODV and HWMP especially in reducing latency and jitter. In particular, GCMR with the traffic prediction method provides the lowest delay among the four routing algorithms. Furthermore, GCMR reduces the average jitter at least 40% lower than AODV-HOP and HWMP. The experimental results show this algorithm can handle efficient real-time communication within a wide range of applications such as online video, file transfer application and so forth.
- In Chapter 4, Gateway-cluster based Load Balancing Multicast algorithm (GLBM) is proposed to optimise the load balancing during the multicast sessions. A new Load-count routing metric is applied to find the lowest load edges among source, receiver and gateway during the construction of the multicast tree. We have conducted extensive simulation experiments using the well-known NS-2 event driven simulation environment. This chapter compares GLBM with well-known existing algorithms,

namely, MAODV and ODMRP to examine the performance in terms of a number of metrics such as throughput, jitter and overhead. The results show that GLBM improves the average throughput of 27% better than that of MAODV and 14% better than that of ODMRP. This is an important gain, especially for the high-bandwidth communication and instant communication applications such as Internet TV, video conference, distance learning and multicast based gaming.

- In Chapter 5, as a multicast version of GCMR, Multicast Gateway-Centralized Multi-hop Routing protocol (MGCMR) is proposed, in which the gateway calculates the best-metric multicast tree based on the information obtained from the instant update messages. The simulation is modelled by setting up both 50-nodes WMNs and larger 100-nodes WMNs to compare MGCMR with MAODV, ODMRP and GLBM. In 50-nodes WMNs, MGCMR provides up to 50% less of the average jitter time than MAODV, ODMRP and GLBM. In WMNs with 100-nodes, MGCMR also outperforms its counterparts in terms of the jitter time and reduces the delay at least 50%. The results also show that MGCMR exhibits higher packet delivery rate and higher throughput. This is important to enable WMNs to offer high-bandwidth demanded applications such as multicast based video applications.
- In Chapter 6, Neighbourhood Load Routing metric (NLR) is proposed as a new cross-layer routing metric with the aim of minimising the cumulative sum of traffic load within each forwarder's neighbourhood on the selected path. By applying NLR, the interference of selected path to other non-forwarder nodes is reduced. Our experiments confirm that NLR improves the communication quality, compared to HOP COUNT, ETX, Load-count and SPP metrics. In WMNs with grid topology, NLR outperforms its counterparts by reducing at least 33.3% lower of the average end-to-end delay and reducing at least 29.3% of the average jitter. The results show

that NLR provides better and efficient communication in all kinds of mesh networks compared to existing routing metrics.

- In Chapter 7, two QoS-aware routing metrics are proposed. Packet Priority-Oriented routing metric (PPO) is designed to assign priorities to different applications to provide differentiated services. Packet Priority QoS-aware routing metric (PP-QoS) is designed by combining PPO with IAR [59]. The experiments compare PPO and PP-QoS with Hop-count, ETX and IAR. The results show that PP-QoS performs well in general. In particular, PP-QoS achieves the lowest latency for high-demand applications with saving 22% less delay than other metrics in heavily-loaded WMNs. In addition, PP-QoS also achieves high efficiency for applications with real-time communication requirements. In applying PP-QoS, a wide range of services, with different sensitivity levels to the QoS are provided with the aim of satisfying the need of users and providers. High-demand applications such as video streaming receive high-bandwidth with low latency and low error rate, and low-demand applications such as web browsing receive lower bandwidth.

To be concluded, WMN is a promising technology to provide low-cost, easy-deployed networking connectivity. In addition, a community mesh network is a type of WMN to enable the connections among homes and enterprises. With the new algorithms of this research, the communication efficiency is improved significantly to support real-time and high-bandwidth communication within the sizable community mesh networks. Thus, the benefited applications include real-time applications such as video application, audio application, online gaming etc, as well as low-demand applications such as web browsing.

8.3 Directions for the Future Work

For future research, we anticipate that our proposed solutions can be implemented in multi-different challenging communication patterns, which could provide a concrete basis for a number of interesting extensions as summarised below.

- We plan to enable GCMR to apply PP-QoS in order to provide the QoS provisioning to one-to-one communication in WMNs. MGCMR would be examined to provide QoS provisioning for multicast communication. The QoS-aware MGCMR can be achieved through designing the QoS-ware gateway based control mechanism. Network coding based multicasting will be investigated in our future work, especially within the context of QoS provisioning.
- NLR is proved to work efficiently in grid topology WMNs. In future, NLR would be re-designed to improve its adaptability in random topology WMNs. Besides, NLR would also be converted to a multi-channel routing metric to fit into the multiple channel characteristic of WMNs. In addition, PP-QoS would also be converted to a multi-channel routing metric.
- We also plan to consider, in our future work, other key communication patterns such as broadcast, gossip and many-to-many communications. In addition, investigating WMNs within the context of next generation energy efficient Internet is a promising research direction towards green communication.

References

[1] Sony PlayStation Portable homepage UK, [Online]. Available: http://uk.playstation.com/psp/.

[2] Sony PlayStation 3 homepage UK, [Online]. Available: http://uk.playstation.com/ps3/.

[3] General packet radio service, [Online]. Available: http://about.qkport.com/g/general_packet_radio_service.

[4] J. E. Flood, Chapter 12, in *Telecommunication Networks*. Institution of Electrical Engineers, London, UK, 1997.

[5] B. H. Walke, Chapter 2, in *Mobile Radio Networks: Networking, protocols and traffic performance,* John Wiley and Sons, LTD, West Sussex England, 2002.

[6] C. Smith and D. Collins, *3G Wireless Networks*, McGraw-Hill, Inc, New York, NY, USA, pp.136, 2000.

[7] Mobile Telephony, [Online]. Available: http://en.kioskea.net/contents/telephonie-mobile/reseaux-mobiles.php3.

[8] J. W. Lechleider, "High Bit Rate Digital Subscriber Lines: A Review of HDSL Progress," *IEEE Journal on Selected Areas in Communications 9*, pp.769–784, 1991.

[9] B. Lee, J. Cioffi, S. Jagannathan, M. Mohseni, "Gigabit DSL," *IEEE Transactions on Communication*, pp.1689–1692, 2007.

[10] Wi-Fi Alliance, [Online]. Available: http://www.wi-fi.org/.

[11] GSM Association, [Online]. Available: http://www.gsmworld.com/.

[12] What is 4G Cellular? [Online]. Available: http://www.mobilein.com/what_is_4GCellular.htm.

[13] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: a survey", in *Computer Networks Journal*, vol.47, pp.445-487, 2005.

[14] P. Racherla and C. Hu, "Understanding community-based wireless networks in tourism: Towards building a knowledge city," *in Proc. of the 14th International Conference on Information Technology and Travel & Tourism*, 2006.

[15] "TRICALCAR: Weaving Community-Based Wireless Networks in Latin America and the Caribbean," [Online]. Available: http://www.idrc.ca/en/ev-111133-201-1-DO_TOPIC.html.

[16] E-Europe homepage, [Online]. Available: http://ec.europa.eu/

[17] R. Morrow, *Bluetooth Operation and Use*, McGraw-Hill Professional, 2002.

[18] Speednet Scotland Homepage, [Online]. Available: http://www.speednetscotland.net/.

[19] Locustworld homepage, [Online]. Available: http://locustworld.com/.

[20] The Southampton Open Wireless Network, [Online]. Available: http://www.sown.org.uk/.

[21] BWN lab wireless mesh networks research project, [Online]. Available: http://www.ece.gatech.edu/research/ labs/bwn/mesh/.

[22] Seattle wireless, [Online]. Available: http://www.seattlewireless.net/.

[23] Microsoft Mesh Networks. [Online]. Available: http:// research.microsoft.com/mesh/.

[24] MIT Roofnet homepage, [Online]. Available: http://pdos.csail.mit.edu/roofnet/doku.php.

[25] The IEEE 802.16 Working Group on Broadband Wireless Access Standards, [Online]. Available: http://wirelessman.org/.

[26] Wireless Africa Home Page, [Online]. Available: http://wirelessafrica.meraka.org.za/wiki/index.php/Wireless_Africa_Home_Page.

[27] One Laptop per Child, [Online]. Available: http://www.laptop.org/.

[28] Metro Mesh Network, [Online]. Available: http://www.netkrom.com/sol_mesh.html.

[29] SMesh Project, [Online]. Available: http://www.smesh.org/.

[30] *Voice over Internet Protocol. Definition and Overview,* International Engineering Consortium, 2007. [Online]. Available: http://www.iec.org/online/tutorials/int_tele/index.asp.

[31] Community wireless, [Online]. Available: http://www.communitywireless.org/.

[32] Military mesh networks, [Online]. Available: http://www.meshdynamics.com/military-mesh-networks.html.

[33] "Differences Between Multicast and Unicast". Microsoft, 2003.

[34] X.200: Information technology - Open Systems Interconnection - Basic Reference Model: The basic model, [Online]. Available: http://www.itu.int/rec/T-REC-X.200-199407-I/en.

[35] J. Templemore-Finlayson and S. Budkowski, "Group Communication and Multicast," in *Proc. of Networking,* vol. 2093, pp. 649-656, 2001.

[36] I. Akyildiz and X. Wang, *Wireless Mesh Networks*, John Wiley & Sons Publisher.

[37] I. Hou, V. Borkar, and P. R. Kumar, "A Theory of QoS for Wireless," in *Proc. Of INFOCOM 2009*, pp. 486-494, 2009.

[38] G. Chelius, E. Fleury, "NP-Completeness of ad hoc multicast routing problems," INRIA – Institut National des Sciences Appliquées de Lyon, 2005.

[39] J. Yuan , Z. Li , W. Yu , and B. Li, "A cross-layer optimization framework for multicast in multi-hop wireless networks," in *Proc. of WICON 05*, pp. 47 – 54, 2005.

[40] R. Ahlswede, N. Cai, S. Li, and R. Yeung, "Network Information Flow," *IEEE Transactions on Information Theory*, vol. 46, pp. 1204–1216, 2000.

[41] J. Sun, Y. Liu, H. Hu and D. Yuan, "On-demand coding-aware routing in wireless Mesh networks," *the Journal of China Universities of Posts and Telecommunications*, vol.17, iss. 5, pp. 80-86, October 2010.

[42] B. Rong, Y. Qian, K. Lu, and R. Q. Hu, "Enhanced QoS Multicast Routing in Wireless Mesh Networks," *Wireless Communications, IEEE Transactions on Wireless Communications*, vol. 7, no. 6, pp. 2119 – 2130, 2008.

[43] C. Perkins, E. Belding-Royer, S. Das, "Ad hoc On-Demand Distance Vector (AODV) Routing, " July 2003. [Online]. Available: http://www.ietf.org/rfc/rfc3561.txt.

[44] C. Perkins and Bhagwat, "Highly Dynamic Destination-Sequence Distance Vector, Routing (DSDV) for Mobile Computers," in *Proc. of* ACM SIGCOMM Computer Communication Review, pp. 234–244, 1994.

[45] E. Royer, and C. E. Perkins "Multicast operation of the ad-hoc on-demand distance vector routing protocol," in *Proc. of the 5th ACM/IEEE Annual Conference on Mobile Computing and Networking*, pp.207-218, 1999.

[46] M. Gerla, S.J. Lee, and W. Su, "On-demand multicast routing protocol (ODMRP) for ad hoc networks," Internet Draft, 2003. [Online]. Available: http://tools.ietf.org/html/draft-ietf-manet-odmrp-04

[47] G. Zeng, B. Wang, Y. Ding, L. Xiao, and M. Mutka, "Multicast Algorithms for Multi-Channel Wireless Mesh Networks," in *Proc. of 2007 International Conference on Network Protocols*, pp. 1-10, 2007.

[48] X. Zhao, C. Chou, J. Guo, S. Jha, "A Scheme for Probabilistically Reliable Multicast Routing in Wireless Mesh Networks," in Proc. Of 32nd IEEE Conference Local Computer Networks, pp. 213-214, 2007.

[49] U. T. Nguyen, "On multicast routing in wireless mesh networks," *Computer Communications*, vol. 31, pp. 1385-1399, 2008.

[50] "IEEE P802.11s/D1.08. Draft amendment to standard – IEEE 802.11: ESS Mesh Networking," IEEE, 2008.

[51] X. Zhao, C. T. Chou, J. Guo, S. Jha and A. Misra, "Probabilistically reliable on-demand multicast in wireless mesh networks," in *Proc. of IEEE WoWMoM 08*, pp. 1 - 9 2008.

[52] A. Adya, P. Bahl, J. Padhye, A. Wolman, and L. Zhou, "A Multi-Radio Unification Protocol for IEEE 802.11 Wireless Networks," in *Proc. of the First International Conference on Broadband Networks (BROADNETS'04)*, pp. 344 – 354, 2003.

[53] S. Keshav, "A Control-Theoretic Approach to Flow Control," in *Proc. of ACM Sigcomm*, vol. 21, pp. 3-15, 1991.

[54] D. Nandiraju, L. Santhanam, N. Nandiraju and D. P. Agrawal, "Achieving Load Balancing in Wireless Mesh Networks Through Multiple Gateways," in *Proc. of 2006 IEEE International Conference of Mobile Ad hoc and Sensor Systems (MASS),* pp. 807-812, 2006.

[55] E. Pourfakhar and A. Masoud Rahmani, "A hybrid QoS multicast framework-based protocol for wireless mesh networks," Computer Communications, vol. 33, iss. 17, pp. 2079-2092, 15 November 2010.

[56] D. S. J. De Couto, D. Aguayo, J. Bicket, and R. Morris, "A High-Throughput Path Metric for Multi-Hop Wireless Routing," in *Proc. of the 9th annual international conference on Mobile computing and networking*, pp. 134 – 146, 2003.

[57] S. Roy, D. Koutsonikolas, S. Das, and Y. C. Hu, "High-throughput multicast routing metrics in wireless mesh networks," in *Proc. of 26th IEEE International Conference on Distributed Computing Systems*, 2006.

[58] S. Banerjee and A. Misra, "Minimum energy paths for reliable communication in multi-hop wireless networks," in *Proc. of ACM Mobi-Hoc*, pp. 146 – 156, 2002.

[59] S. Waharte, B. Ishibashi, R. Boutaba, and D. Meddour, "Interference-aware routing metric for improved load balancing in wireless mesh networks," in *Proc. of 2008 IEEE International Conference on Communications*, pp. 2979-2983, 2008.

[60] L. Ma and M. Denko, "A Routing Metric for Load-Balancing in Wireless Mesh Networks," in *Proc. of 21st International Conference on Advanced Information Networking and Applications Workshops*, vol.2, pp.409 – 414, 2007.

[61] V. C. M. Borges, , M. Curado, and E. Monteiro, "The impact of interference-aware routing metrics on video streaming in Wireless Mesh Networks," *Ad Hoc Networks*, to be published, 2011.

[62] L. Franken, "Quality of Service Management: A Model-Based Approach," PhD thesis, Centre for Telematics and Information Technology, University of Twente in the Netherlands, 1996.

[63] M. Peuhkuri, "IP Quality of Service," Helsinki University of Technology, Laboratory of Telecommunications Technology, 1999.

[64] The network simulator, NS2 homepage, [Online]. Available: http://www.isi.edu/nsnam/ns/.

[65] "List of wireless community networks by region", [Online]. Available: http://wapedia.mobi/en/List_of_wireless_community_networks_by_region.

[66] U. Varshney, "Multicast over wireless networks," *Communications of the ACM*, vol. 45, pp. 31-37, 2002.

[67] B. Wang, S. Soltani, J. K. Shapiro, and P. Tan, "Distributed detection of selfish routing in wireless mesh networks," Technical Report MSU-CSE-06-19, Michigan State University, 2006.

[68] Q. Xue and A. Ganz, "QoS Routing for Mesh-Based Wireless LANs," *International Journal of Wireless Information Networks*, vol. 9, no. 3, pp. 179-190, 2002.

[69] L. L. Weckerle , H. Li , M. Lott , M. Weckerle , W. Zirwas , and E. Schulz, “Multihop Communications In Future Mobile Radio Networks,” in *Proc. of PIMRC*, pp. 54 – 58, 2002.

[70] H. Li and D. Yu, “Comparison of ad hoc and centralized multihop routing,” in *Proc. of The 5th International Symposium, Wireless Personal Multimedia Communications*, vol. 2, pp. 791-795, 2002.

[71] R. Ananthapadmanabha, B.S. Manoj, and C. S. R. Murthy, “Multihop cellular networks: the architecture and routing protocols,” in *Proc. of 2001IEEE PIMRC*, pp. 78-82, 2001.

[72] K. Römer, F. Mattern, “The Design Space of Wireless Sensor Networks,” in *Proc. of IEEE Wireless Communications*, pp. 54–61, 2004.

[73] T. Haenselmann, Sensornetworks. GFDL Wireless Sensor Network textbook, 2006.

[74] Open Mesh, [Online]. Available: http://www.open-mesh.com/.

[75] C. Toh, Ad Hoc Mobile Wireless Networks, Prentice Hall Publishers, 2002.

[76] Y. D. Lin and Y. C. Hsu, “Multi-hop cellular: a new architecture for wireless communications,” in *Proc. of Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 3, pp. 1273-1282, 2000.

[77] “IEEE 802.11, The Working Group Setting the Standards for Wireless LANs,” [Online]. Available: http://ieee802.org/11/.

[78] B. S. Manoj and R. R. Rao, “Load Balancing in Wireless Mesh Networks,” *Wireless Mesh Networks: Architectures, Protocols, and Standards (Ed. Y. S. Zhang et al.)*, CRC Press, 2006.

[79] S. Tajima, T. Higashino, N. Funabiki, S. Yoshida, “An Internet Gateway Access-point selection problem for Wireless Infrastructure Mesh Networks,” in *Proc. of 7th International Conference on Mobile Data Management*, pp.112 – 112, 2006.

[80] B. Aoun, R. Boutaba, Y. Iraqi, G. Kenward, “Gateway Placement Optimization in Wireless Mesh Networks with QoS Constraints,” *IEEE Journal of Selected Areas in Communications*, pp.2127-2136, 2006.

[81] X. Tao, T. Kunz and D. Falconer, “Traffic balancing in wireless MESH networks,” in *Proc. of 2005 International conference on Wireless Networks, Communications and Mobile Computing*, pp. 169 – 174, 2005.

[82] M. Saghir, T. C. Wan, and R. Budiarto, “Load Balancing QoS Multicast Routing Protocol in Mobile Ad Hoc Networks,” *Technologies for Advanced Heterogeneous Networks*, pp. 83-97, 2005.

[83] J. Gao and L. Zhang “Load-Balanced Short-Path Routing in Wireless Networks,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 17, pp. 377 – 388, 2006.

[84] D. M. Shila and T. Anjali, “Load-aware Traffic Engineering for Mesh Networks, ” in *Proc. of 16th International Conference on In Computer Communications and Networks*, pp. 1040-1045, 2007.

[85] B. Sun and L. Li, “QoS-aware multicast routing protocol for Ad hoc networks,” *Systems Engineering and Electronics*, vol. 17, pp. 417-422, 2006.

[86] Y. Drabu and H. Peyravi, “Gateway Placement with QoS Constraints in Wireless Mesh Networks,” in *Proc. of Seventh International Conference on Networking*, pp. 46-51, 2008.

[87] L. Junhai, Y. Danxia, X. Liu, and F. Mingyu, “A Survey of Multicast Routing Protocols for Mobile Ad-Hoc Networks,” *IEEE COMMUNICATIONS SURVEYS & TUTORIALS,* vol. 11, no. 1, pp. 78-91, 2009.

[88] D. B. Johnson and D. A. Maltz, “Dynamic Source Routing in Ad Hoc Wireless Networks,” *Mobile Computing*, vol. 353, pp.153-181 1996.

[89] R. Bellman, “On a Routing Problem,” *Quarterly of Applied Mathematics*, pp.87-90, 1958.

[90] L. R. Ford and D. R. Fulkerson, *Flows in Networks*, Princeton University Press, 1962.

[91] A. Damle, D. Rajan, and S. Faccin, “Hybrid Routing with Periodic Updates (HRPU) in Wireless Mesh Networks,” in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, pp 318-324, 2006.

[92] A. Le, D. Kum, and Y. Cho, “An Efficient Hybrid Routing Approach for Hybrid Wireless Mesh Networks, ” in *Proc. of the 3rd International Conference and Workshops on Advances in Information Security and Assurance*, pp. 532 – 542, 2009.

[93] M. Michalak and T. Braun, “Common gateway architecture for mobile ad-hoc networks,” in *Proc. of the Second Annual Conference on Wireless On-demand Network Systems and Services*, pp. 70–75, 2005.

[94] R. Baumann, S. Heimlicher, V. Lenders, and M. May, “Routing Packets into Wireless Mesh Networks,” in *Proc. of IEEE Conference (WiMob'07)*, 2007.

[95] K. Oh and C.Y. Lee, “Multicast Routing Protocol with Low Transmission Delay in Multi-Rate, Multi-Radio Wireless Mesh Networks,” in *Proc. of 2010 IEEE International Conference on Communications (ICC)*, pp. 1-6, May 2010.

[96] P. M. Ruiz, F. J. Galera, C. Jelger, and T. Noel, “Efficient Mulicast Routing in Wireless Mesh Networks Connected to Internet,” in *Proc. of the first international conference on Integrated internet ad hoc and sensor networks*, 2006.

[97] F. K. Hwang, D. S. Richards, and P. Winter, *The Steiner Tree Problem*. Elsevier, North-Holland, 1992.

[98] R. Draves, J. Padhye, and B. Zill, “Routing in Multi-Radio, Multi-Hop Wireless Mesh Networks,” in *Proc. of the 10th annual international conference on Mobile computing and networking*, pp. 114 – 128, 2004.

[99] R. Baumann, S. Heimlicher, V. Lenders, and M. May, "HEAT: Scalable Routing in Wireless Mesh Networks Using Temperature Fields," in *Proc. of 2007 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*, 2007.

[100] G. Xu and J. Liu, "Improvement of AODV routing protocol based on wireless mesh networks," in *Proc. of 2010 International Conference on Computer, Mechatronics, Control and Electronic Engineering (CMCE),* pp. 372 – 375, 2010.

[101] B. Manoj, R. Ananthapadmanabha, and C. Murthy, "Multi-hop Cellular Networks: Architecture and Protocols for Best-effort and Real-time Communication," *Journal of Parallel and Distributed Computing*, pp. 767-791, 2005.

[102] J. Haas, "A new routing protocol for the reconfigurable wireless networks," *in Proc. of IEEE 6th International Conference on Universal Personal Communications*, pp. 562–566, 1997.

[103] A. Pirzada, R. Wishart, M. Portmann, and J. Indulska, "ALARM: An Adaptive Load-Aware Routing Metric for Hybrid Wireless Mesh Networks," in *Proc. of Thirty-Second Australasian Computer Science Conference*, pp. 25-34, 2009.

[104] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, Second Edition, the MIT Press, September 2001.

[105] G. Wang, Z. Cao, L. Zhang, K. C. C. Can and J. Wu, "A Novel QoS Multicast Model in Mobile Ad Hoc Networks," in *Proc. of the 19th IEEE International Parallel and Distributed Processing Symposium*, 2005.

[106] E. Pourfakhar and A. Rahmani, "Optimization of Multicast Routing Based on a Reliable Effective Framework in Wireless Mesh Networks," in *Proc. of Advances in Multimedia Information Processing (PCM 2009)*, pp. 416-427, 2009.

[107] K. Han, Q. Guo, “Reducing Multicast Redundancy and Latency in Wireless Mesh Networks,” in *Proc. of First International Workshop on Education Technology and Computer Science*, vol. 1, pp. 1075 – 1079, 2009.

[108] J. Jun and M. L. Sichitiu, “MRP: wireless mesh networks routing protocol,” *Journal of Computer Communications, special issue on “resource management and routing in wireless mesh networks”*, pp. 1413–1435, 2008.

[109] Z. Ke, L. Li, Q. Sun, and N. Chen, “A QoS multicast routing algorithm for wireless mesh networks,” in *Proc. of the Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing*, vol. 1, pp. 835–840, 2007.

[110] A. Lee and P. Ward. “A Study of Routing Algorithms in Wireless Mesh Networks,” in *Proc. of Australian Telecommunication Networks and Applications Conference*, 2004.

[111] H. Tokito, M. Sasabe, G. Hasegawa and H. Nakano, “Routing Method for Gateway Load Balancing in Wireless Mesh Networks,” in *Proc. of 2009 International conference on Networks*, pp. 127 – 132, 2009.

[112] F. Li, Y. Wang, and X. Y. Li, “Gateway placement for throughput optimization in wireless mesh networks,” in *Proc. of International conference on 2007 Communications*, pp. 4955-4960, 2007.

[113] V. Mhatre, F. Baccelli, H. Lundgren, and C. Diot, “Joint MAC-aware Routing and Load Balancing in Mesh Networks,” in *Proc. of the 2007 ACM CoNEXT conference*, December 2007.

[114] R. Sollacher and M. Greiner, “Impact of Interference on the Wireless Ad-Hoc Networks Capacity and Topology,” *Wireless Networks*, vol. 12, pp. 53-61, 2006.

[115] E. Royer, C. Toh, “A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks”, *IEEE Personal Communications*, pp. 46–55, 1999.

[116] K. Ozan. Tonguz, and G. Ferrari, *Ad Hoc Wireless Networks: A Communication-Theoretic Perspective*, John Wiley & Sons, 2006.

[117] "Deployment Strategies and Equipment Forecasts for Municipal Wi-Fi and Campus Networks," [Online]. Available: http://www.abiresearch.com/research/1003414-Wireless+Mesh+Networking.