Cross Layer Design Approaches for the Enhancement of QoS in MANETs

THESIS

Submitted in partial fulfillment

of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

GAWAS MAHADEV ANANT

Under the Supervision of

Dr. LUCY J. GUDINO

and

Co-supervision of

Prof. K. R. ANUPAMA

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE,

PILANI

2018

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

Certificate

This is to certify that the thesis entitled **'Cross Layer Design Approaches for the Enhancement of QoS in MANETs'** and submitted by **GAWAS MAHADEV ANANT**, ID.No. **2010PHXF0020G** for award of Ph.D. of the Institute embodies original work done by him under our supervision.

Signature of the Supervisor :

Name in capital letters : **Dr. LUCY J. GUDINO**

Designation : **ASSISTANT PROFESSOR**

Date : 14/3/18

Signature of the Co-supervisor : K.R. Anupama

Name in capital letters : **Prof. K. R. ANUPAMA**

Designation : **PROFESSOR**

Date : 14/3/2018

Acknowledgements

Completion of this doctoral thesis was possible with the support of several people. I would like to express my sincere gratitude to all of them

Firstly, I would like to express my deep and profound gratitude to my supervisor Dr. Lucy J. Gudino for the continuous support during my Ph.D. study and related research, for her patience, motivation, and immense knowledge. Her guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better supervisor and mentor for my Ph.D. study.

I take this opportunity to express my sincere gratitude to my co-supervisor, Prof. K. R. Anupama for her insightful supervision and encouragement. I very much appreciate all her assistance and constructive comments.

Besides my supervisors, I would like to thank my Doctoral Advisory Committee, Prof. Bharat Deshpande, and Prof. Neena Goveas, for offering many valuable suggestions on my thesis improvement, and also for the hard question which incited me to widen my research from various perspectives.

I express my thanks to Prof. Souvik Bhattacharyya, Vice-Chancellor, BITS-Pilani, Prof. G Raghurama, Director, BITS-Pilani, K. K. Birla Goa Campus, Prof. S. K. Verma, Dean, Academic Research Division, BITS-Pilani, Prof. P. K. Das, Associate Dean, Academic Research Division, BITS-Pilani, K. K. Birla Goa Campus and members of the Doctoral Research Committee of Dept. of CS & IS, BITS-Pilani, K. K. Birla Goa Campus for providing administrative support and facilities to carry out my research at the Institute.

I am thankful to my colleagues in Dept. of CS & IS, BITS-Pilani, K. K. Birla Goa Campus for their valuable guidance, and support throughout my research work.

My special sincere thanks have to go to my parents, Anant Gawas and Anuradha Gawas, for everything they have been giving me. I salute them all for the selfless love, care, pain and sacrifice they did to shape my life. Although they hardly understood what I researched on, they were willing to support any decision I made. I would never be able to pay back the love and affection showered upon by my parents. I would also like to extend my thanks to my sisters Archana and Nilima for their support and valuable prayers.

I owe thanks to a very special person, my dear wife, Mridini Gawas, for her continued and unfailing love, support and encouraging when I was depressed and frustrated. She taught me how to be patient and persistent while I was facing difficulties. I appreciate my baby, my little girl Riyanshi, for abiding my ignorance and the patience she showed during my thesis writing. Words would never say how grateful I am to both of them. Yet with all humbleness, I thank them for their noble gesture and splendid support.

My heart felt regard goes to my father in law, mother in law for their love and moral support. I consider myself the luckiest in the world to have such a lovely and caring family, standing beside me with their love and unconditional support.

I thank the Almighty for granting me the wisdom, health and strength to undertake this research task and enabling me to its completion by showering his blessings upon me.

GAWAS MAHADEV ANANT

Contents

List of Tables

List of Figures

Abstract

Mobile Ad hoc Networks (MANETs) are gaining increasing popularity in recent years because of the flexibility and convenience in setting up a network. MANETs are distributed, dynamic, and self-configurable wireless network without infrastructure support. MANETs offers unique benefits in certain application such as, isolated sites of natural disasters where existing communication networks are destroyed, communications on the battlefields where military units may move constantly, etc. Future applications on the network are expected to be capable of carrying a multitude of real-time multimedia application data such as voice and video as well as best effort data. Such applications require guarantees on the Quality of Service (QoS) such as high bandwidth, delay, jitter, and tolerance to packet losses. Delivering multimedia traffic on the Internet is challenging, as best effort based Internet Protocol will not be able to meet stringent QoS requirements. High mobility, limited bandwidth and congestion in the network makes the problem even worse as it leads to packet losses and increased packet delay. Therefore, the real-time multimedia applications in MANETs require QoS-aware mechanisms for channel access, identification of proper forwarding nodes, and congestion prevention as well as management. In order to address the above issues in resource constrained MANETs four protocols are proposed in the thesis.

Cross Layer Best effort QoS aware routing (CLBQ) Protocol : The large amount of real-time traffic tends to be in bursts, which is bandwidth intensive and liable to congestion. Congestion in MANETs leads to transmission delays, packet losses, wastage of network resources and network lifetime reduction due to nodes energy depletion. The CLBQ protocol utilizes cross layer interactions between MAC and Network layer to obtain multi-rate support from 802.11b. The protocol discovers multiple node disjoint multipath routes which are constructed through less congested nodes rather than the shortest routes. The protocol estimates the one-hop delay based on the link quality. Based on the estimated delay, the protocol dynamically eases the congestion by selecting appropriate data rate and optimal route for admitting a flow with a requested delay. The simulation result shows that the protocol significantly reduces both the packet drop ratio and the end-to-end delay without much impact on control overhead.

Adaptive Best Effort Traffic scheduler for IEEE 802.11e EDCA (ABET-EDCA) protocol : To improve QoS features of IEEE 802.11 for wireless LAN, IEEE approved

an amendment, named IEEE 802.11e. However, in the IEEE 802.11e, MAC parameters are statically set without taking into account the volatile nature of the network. Also, IEEE 802.11e does not provide any measures to support fairness for TCP traffic in the presence of multimedia traffic. Therefore, in the process of providing QoS enhancements to delay sensitive multimedia traffic, TCP traffic is left to starve for a long time leading to performance degradation. The ABET-EDCA protocol aims to maintain QoS for multimedia traffic without sacrificing TCP performance and network utilization. The protocol prioritizes the TCP traffic during congestion by adaptively tuning the MAC parameters based on traffic load conditions to improve the network performance and to maximize the system throughput. The simulation results show that the performance of low priority TCP traffic is enhanced while having minimal impact on high priority delay sensitive UDP-based multimedia traffic.

Multi Objective Cross-Layer Optimization (MOCLO) protocol: This protocol aims to improve the performance of both multimedia and TCP traffic in terms of packet delay, and throughput. The MOCLO protocol proposes a modification to IEEE 802.11e MAC layer to prioritize and protect the important packets within each class of traffic flow. It uses cross layer interaction between PHY-MAC-Network layers to find multiple optimal route to support different flows in access categories based on priority thereby reducing the load on a particular route. The performance of the network with MOCLO protocol is thoroughly evaluated through the extensive simulations, which highlight the different metrics performance, namely packet delay, throughput and packet loss under different running conditions.

Adaptive Multi QoS Cross-Layer Cooperative Routing in MANETs (AMCCR) protocol: The AMCCR protocol is an energy aware, end-to-end routing scheme which uses both cooperative communication and rate adaptation techniques to enable the nodes to adapt their data rates to match the channel conditions and node mobility thereby improving network throughput. It uses cross layer communication between PHY, MAC, and network layers. The protocol guarantees multiple QoS requirements and optimizes the trade-off between end-to-end delay and energy of the system. AMCCR uses an adaptive mechanism to select the transmission mode, i.e., direct or relay mode so that a low data rate host can be assisted by high data rate relay nodes. AMCCR protocol also supports dual-hop half-duplex communication via selected relay node by coding technique. The extensive simulation conducted shows that in cooperative mode, AMCCR significantly shows better performance in terms of throughput, delay and network lifetime.

Chapter 1

Introduction

The wireless Network has been experiencing exponential evolution in the past two decades, driven by the rapid advances in wireless technologies, network infrastructures, growing availability of wireless applications and popularity of portable computing devices, all getting more powerful in their capabilities. The wireless networks offer low cost, convenient, scalable, and integrated solution that helps to set up a network effortlessly. The current trend in wireless networks is towards catering pervasive and ubiquitous computing to both roaming and fixed users, creating a global connectivity. The aim is to integrate the existing wireless communication systems with the evolving Internet to enable communications between physical and virtual things.

The wireless network architecture is categorized into two types: infrastructure based networks and ad hoc networks. Cellular networks and wireless local area networks (WLAN) are examples of infrastructure based networks where users are connected via base stations (access points) and backbone networks [1]. The current infrastructure based networks have greatly improved the data transmission speed, which enables a variety of high-speed mobile data services and usually applications require single hop connectivity to the wired network. Setting up a fixed infrastructure network might be infeasible for situations like, emergency search-and-rescue or military operations, where a temporary communication network needs to be deployed immediately. A self-organizing network of autonomous wireless devices that communicate without any infrastructure or centralized administration over a shared wireless channel is referred as an Ad hoc Network [2]. Communicating in an ad hoc fashion brings interesting solutions to guarantee ubiquitous connectivity for the "Internet of future". Ad hoc network can be exploited in creating smart and self-aware environments. Ad hoc networks can be built around any wireless technology, including infrared, radio frequency (RF), global positioning system (GPS), and so on.

Mobile Ad hoc network (MANET) is an autonomous, infrastructure less, distributed, self-organizing and self-administering network where certain nodes move in and out of the network and the remaining nodes automatically reconfigure to adjust to the new development [3]. MANET supports a dynamic network topology. Due to limited transmission ranges, direct communication between the nodes is not possible. Hence, multi hop communication is usually supported where each node acts not only as transmitter and receiver but also as a relay [4].

With recent advances in technology, communication nodes have shrunk in size, require lesser energy for operation and incorporate more advanced functions. This has motivated many innovations in the field of MANET to carry out data intensive real time multimedia applications such as voice over IP, audio and video streaming etc. which require the network to provide guarantees on the Quality of Service (QoS) [5]. QoS is defined as a set of service requirements that need to be met by the network while transporting a packet stream from source to destination. The three important QoS requirements for multimedia applications are high bandwidth guarantee, delay guarantee, and tolerance to packet losses. The routing plays an important role in QoS provisioning. The routing protocol designed should find a route which should satisfy end to end QoS guarantees and should be of low cost, stable and energy efficient.

In this thesis, an attempt is made to improve the performance of MANETs to support real time multimedia application by designing an energy efficient, congestion adaptive, multi-path complying multi-QoS routing protocol in association with MAC and physical layers.

1.1 Overview of Mobile Ad Hoc Networks

Ad hoc is a Latin word which means ”for this” or “for this only”. Ad hoc network is a decentralized type of wireless network which does not rely on infrastructure. Packet radio networks (PRNET) was the first ad hoc network system developed in early 1970’s by the DARPA (Defense Advanced Research Projects Agency) [6] to improve tactical network communications in the battlefield. Initially, PRNET used ALOHA [7] and later switched to CSMA to support multi-hop multiple-access, which covered wide communication area. The term multi-hop means that the data from the source travels through several other intermediate nodes to reach the destination. The main advantage of using CSMA in PRNET is dynamic sharing of the radio resources. In PRNET, the network comprised of mobile radio repeaters, wireless terminals, and mobile stations. The installation of the system was simple, self-initializing and self-organizing. The packets were passed on from one repeater to the other until data reached its destination.

Based on the applications, ad hoc networks can be classified into three categories; Mobile Ad hoc Networks (MANETs), Wireless Mesh Networks (WMNs), Wireless Sensor Networks (WSN) [6] [8]. MANET is an autonomous, infrastructure less, distributed, self-organizing and self-administering network where certain nodes move in and out of the network and the remaining nodes automatically reconfigure to adjust to the new development. MANETs supports a dynamic network topology. Due to limited transmission ranges, direct communication between the nodes may not be possible all the time. Hence, multi hop communication is usually supported where each node acts not only as transmitter and receiver but also as a relay. MANETs inherit some of the common characteristics found in wireless networks, and add certain characteristics specific to ad hoc networking which are described as:

1. Wireless: Nodes communicate wirelessly and share the transmission medium.
2. Ad hoc: MANET is a provisional network formed dynamically in a random manner by a collection of nodes as need arises.
3. Decentralised : The network should detect the presence of new nodes automatically and include them seamlessly. Also, if any node moves out of the

network, the remaining nodes should automatically reconfigure themselves to adjust to the new scenario.

4. Autonomous: MANETs does not depend on any fixed infrastructure or centralized supervision. Each node operates in a distributed peer-to-peer mode, generates independent data, and acts as an independent router.

5. Multi hop routing: Each of the nodes in the network acts as a router and forwards each others' packets to enable information sharing between mobile hosts.

6. Mobility: Each node moves freely and randomly while communicating with other nodes. Thus, the topology is dynamic in nature due to constant movement of the nodes, causing the intercommunication among nodes to change continuously.

The dynamic, decentralized, self-organizing and scalability feature of MANET makes them suitable in scenarios where rapid, infrastructure less and low maintenance deployment is required. MANETs support wide variety of applications ranging from small scale static network to large scale mobile and highly dynamic networks. Some of the applications are listed below:

1. Military Tactical Operations: In battlefields, it is difficult to have an infrastructure oriented network. In such areas, fast and short term establishment of communication can be achieved via an MANET for a proper tactical coordination amongst the soldiers.

2. Disaster Relief Operations: During natural calamity, it is possible that existing communication infrastructure might have been completely destroyed and restoring communication quickly is vital. By deploying a MANET, a communication infrastructure can be set up seamlessly.

3. Emergency Services: In several applications such as search and rescue operations, policing and firefighting etc. requires unplanned and spontaneous interpersonal communications. Faster deployment of MANETs can support such applications.

4. Vehicular Communication: Wireless ad hoc network applications with mobility is extended in vehicular technology and are called Vehicular Ad hoc wireless

NETworks (VANETs). In VANETs, vehicles transmit traffic and safety-related information, including traffic congestion-avoidance messages, accident warnings, and accident reports, etc., which assist drivers in making the proper decisions to avoid vehicle collisions and congestion.

5. Other applications: MANETs also finds its use in commercial and civilian applications such as e-commerce, home and enterprise networking, sports, education, entertainment etc.

Over the last few years, several routing protocols and algorithms have been proposed and their performance under various network environments and traffic conditions were studied and compared [9]. The primary objective of an ad-hoc network routing protocol is to construct an efficient route between a pair of nodes so that messages may be delivered reliably and in a timely manner. Important criteria to be considered while designing the new routing protocols include:

1. Simplicity and ease of implementation.
2. Optimal routes, and possibly, multiple routes should be available between each pair of nodes to increase robustness.
3. The routing protocol should be distributed in nature and should quickly adapt to changes in topology and traffic pattern, resulting from mobility and route failure.
4. The routing protocol should result in optimized bandwidth and power utilization with minimal control overhead.
5. The routing protocol should be scalable, secure and reliable.
6. The routing protocol should support QoS requirements of the application under consideration.

1.2 Research Motivation

Multimedia services play a central role in many social and entertainment applications. These applications often have stringent and reliable QoS requirements, which the network must cater. Based on applications, QoS in MANETs is classified into two categories: Hard QoS applications and Soft QoS applications. Applications such as air traffic control systems, nuclear reactor control systems require strict QoS guarantees whereas applications such as video conferencing can tolerate degradation in guaranteed QoS to some extent. Providing hard QoS guarantees is very difficult due to the following reasons:

1. Dynamically varying network topology :The route with admitted QoS will not be able to sustain QoS guarantees due to frequent path breaks caused by random movements of nodes. This requires route to be reestablished and delay due to this may cause some packets to miss their delay targets.

2. Imprecise state information : Routing decision mainly depends on link state information and flow state information. Usually the nodes in network maintains link specific state information such as bandwidth, delay, distance values, loss rate, error rate etc. and flow specific information such as session ID, source and destination addresses, and QoS requirements of the flow. Due to dynamic nature of the channel and frequent topological changes, it is difficult to obtain accurate routing decisions.

3. Lack of central coordination : QoS provisioning in MANETs is complicated due to lack of central coordination. Nodes in the network has to coordinate with each other to have proper QoS provisioning.

4. Error prone shared channel: Bit error rate of wireless media is very high compared to wired media due to channel impairments such as attenuation, multipath propagation, fading and interference.

5. Hidden terminal problem: Due to hidden terminal problem, packets collide at the receiver node leading to retransmission of packets. Delay due to retransmission is not acceptable for flows having strict QoS requirements.

6. Limited resource availability: Performance of the QoS provisioning is also affected by the limited resources such as bandwidth, battery life, processing power, and storage space. Proper resource management is one of the requirements for QoS provisioning.

7. Variation in node capabilities: The mobile node may be equipped with one or more radio interfaces that have varying transmission/receiving capabilities and operate across different frequency bands. This result in asymmetric links. Designing the protocols and algorithms for this heterogeneous network can be complex, requiring dynamic adaptation to the changing node power and channel conditions.

For proper operation of multimedia services in MANETs, the QoS routing is essential instead of best-effort routing. Different QoS metrics can be considered to satisfy QoS requirements in route selection. For example, the minimum required throughput, maximum tolerable delay, maximum tolerable delay jitter, maximum tolerable packet loss ratio etc.

To guarantee the QoS provisioning, the route selection should take into consideration multiple QoS metrics instead of a single metric. To accomplish this, efficient routing protocol design strategies are necessary. Thus, for any routing protocol design, some of the following QoS metrics must be taken into consideration during the route discovery and route selection phase:

1. Available network resources such as bandwidth and node energy.

2. Channel information such as bit error rate and signal strength.

3. Transport layer parameters such as throughput and packet drop rate.

4. Mac layer parameters such as Queue buffer utilization and channel contention time.

5. Application requirement such as end to end delay, and user priority.
6. Link states, link lifetime, link stability.
7. Node Location and position information node coordinates, mobile speed, and neighbor density.

Unfortunately, multi QoS aware routing protocols are difficult to design using the layered OSI architecture. The layered network architecture is remarkably successful for networks made up of wired links, where the key assumptions and abstraction boundaries work well. The strict layering approach reveals to be suboptimal in many application domains of MANETs [10]. The main drawback of the layered architecture is the lack of cooperation among non-adjacent layers i.e., each layer works in isolation with little information about the network. Also, the strict modularity does not allow to design joint solutions to maximize the overall network performance.

Several routing protocols have been developed to support QoS in MANETs such as CEDAR [11], QoS-AODV [12], AQOR [13], (AQA-AODV) [14]. However, these algorithms do not emphasize to exploit the information across the multiple layers. The estimation of QoS metrics becomes imprecise because different QoS metric information is available at the different layers. Therefore, it is important to design QoS routing protocol using accurate estimation techniques to measure QoS metrics by utilizing the network information across the multiple layers.

1.3 Research Gap

For real time multimedia traffic, the data rate and delay are the crucial QoS factors. Therefore, to satisfy these QoS requirements, each route in the network should provide a correct estimate of the available data rate and delay. Nowadays, many wireless standards, such as 802.11a/b/g [15], can be operated at various data rates to facilitate the efficient use of the limited resources of MANETs. The multi-rate enhancements make it more difficult to estimate the required delays for multimedia transmission.

There are several QoS routing protocols [12] [14] [16], which focus only on satisfying one of the QoS requirements by using a single base rate. For example, Ad hoc Qos On-demand Routing (AQOR) protocol proposes a model which uses the end-to-end delay estimate for admission control while discovering the QoS-satisfied routes whereas Core-Extraction Distributed Ad hoc Routing (CEDAR) selects the routes constituted by stable links with high available bandwidth.

In literature, various rate adaptation schemes have been proposed for WLAN. Auto rate fallback (ARF) [17] and Collision Aware Rate Adaptation(CARA) [18] are a few notable examples. However, these schemes are not applicable, in their current format, to MANETs as they are more suitable for the centralized control unit. The performance of these protocols is degraded during congestion as the transmission rate switches to lower rate [19]. This rate switch increases the channel occupation time, thereby compounding the congestion and reduces overall network throughput and capacity. The lack of techniques to reliably identify and characterize the congestion in wireless networks has prevented development of rate adaptation solutions that incorporate congestion information in their decision framework.

Congestion is one of the most crucial aspects of wireless ad-hoc networks. Congestion leads to problems like long delay, high overhead, and low throughput. To overcome these problems, congestion aware routing protocols are proposed such as CARM [20], CADV [21] and DLOAR [22]. However, these protocols do not suggest a mechanism, to adapt to the congestion if takes place. In the above schemes, the congestion factor is considered only when establishing a new route and remains the same until node mobility or link failure results in route disconnection.

Many protocols in the literature work on optimizing individual objectives such as either provide congestion aware routing ignoring rate adaptation mechanism, or focus only on rate adaptation ignoring congestion around the network leading to sub-optimal solution. Hence, there is a need to propose a solution which can consider both rate adaptation as well as congestion adaptation.

The need to support QoS with proper service differentiation for the delay-sensitive multimedia applications has led to the emergence of IEEE 802.11e, which is an extension to existing 802.11 for WLAN [23]. The IEEE 802.11e introduces two medium access methods: Enhanced Distributed Channel Access (EDCA) and HCF Controlled Channel Access (HCCA). The standard has defined the new enhanced distributed coordination function (EDCF). The main idea of EDCF is to apply different MAC parameters for different traffic flows. The IEEE 802.11e protocol has been extensively studied in the literature over the past decade. The studies have majorly focused on evaluating the performance of EDCF [24]. Though the EDCF provides service differentiation among the different traffic flows, it does not specify any mechanism to provide differentiation among the same traffic flow.

Today's Internet traffic is still dominated by a TCP based applications such as email, FTP, HTTP. IEEE 802.11e has not adequately addressed the issue of handling best effort TCP traffic containing TCP data and TCP ACK packets in the presence of high delay sensitive UDP traffic. In the process of providing QoS enhancements to delay sensitive multimedia traffic, best effort traffic is left to starve for a longer time. Although it is necessary to use effective QoS protocols to provide QoS guarantee to real time traffic, it is also important to provide QoS for TCP traffic so that all users can get reasonable services. Some of the notable work has been proposed in [25] [26][27] to prioritize the TCP ACK packets to improve the TCP performance.

However, these protocols suffer from three drawbacks. Firstly these protocols are designed to suite for WLAN with centralized access point control unit. Secondly, these protocols do not provide any mechanism to identify the network event due to which the TCP packets are lost. Thirdly, the above studies attempted to address the problem of low priority traffic in general, their effectiveness is not guaranteed in networks highly loaded with high priority traffic.

It is well known fact that nodes usually spend most of their energy in communication [28]. Past several years, energy efficiency problem in wireless communication has received lot of attention of several researchers. This problem can be approached in two ways: energy efficient route selection algorithms at the network layer or efficient communication schemes at the physical layer. Cooperative communication (CC) [29] is a promising technique for conserving the energy consumption in MANETs. The broadcast nature of the wireless medium is exploited in cooperative fashion. Research has verified that cooperative communication can achieve tremendous improvements in network capacity, delay, and energy consumption. Early studies of cooperative communication were performed on physical layer cooperation [30][31]; studies were later extended to the Media Access Control (MAC) layer [32][33]. A combined cross layer approach involving physical layer, MAC layer and network layer can be explored. The cooperative communication protocol at network layer can exploit the rate adaptation techniques at MAC layer so that the nodes to adapt their data rates to match the channel conditions and node mobility. Cross layer techniques can also provide a substantial improvement in network lifetime of multi-hop network. However, so far no standard on cooperative routing design has been achieved, and hence leave it an open research topic.

1.4 Aim and Objectives

The aim of this research is to devise a cross layer technique to provide an efficient routing function for MANETs while satisfying the application's QoS requirements. The objectives of the research are as follows:

1. To design a distributed congestion adaptive routing protocol which can provide multiple congestion aware node disjoint paths based on the Qos metrics such as data-rate, packet forwarding delay, and buffer queuing delay.

2. To propose a cross layer route discovery framework for QoS route selection for MANETs.

3. To design a protocol to provide QoS for TCP traffic during congestion, so as to maximize the system throughput while having minimal negative impact on high priority delay sensitive UDP-based real time traffic.

4. To improve the performance of both real time multimedia and TCP traffic by prioritizing and protecting the important packets within each class of traffic flow through cross layer optimization between PHY-MAC-Network layers.

5. To exploit the spatial diversity and multi rate support from physical layer by using adaptive MAC transmission technique to design energy efficient cooperative routing.

1.5 Organization of Thesis

The work embodied in this thesis is arranged in seven chapters.

- Chapter 1 presents overview of MANETs, their characteristics and applications. It also presents research motivation and objectives of this research.
- Chapter 2 focuses on cross layer design for MANETs, by analyzing the benefits and challenges of the cross layer approach and by providing an extensive review of the cross layer solutions and architectures.
- Chapter 3 reviews various rate adaptation techniques and the protocols that uses rate adaptation. This chapter presents a detailed design of adaptive and reliable congestion control, multi-rate adaptation routing protocol "Cross Layer Best effort QoS aware routing protocol (CLBQ)". The protocol uses the cross layer interactions between MAC and Network layer to utilize multi-rate supports from 802.11b network. The protocol consists of rate adaptation and congestion aware optimization to improve the overall performance in terms of throughput, packet delivery, and latency in MANETs.
- Chapter 4 presents a protocol called, Adaptive Best Effort Traffic Scheduler for IEEE 802.11e EDCA (ABET-EDCA) which is a modification to EDCF. The ABET-EDCA uses a novel cross layer based packet scheduling scheme. The protocol prioritizes the TCP ACK packets during congestion in the presence of UDP traffic by adaptively tuning the MAC parameters based on traffic load conditions thereby improving best effort traffic performance.
- Chapter 5 extends the work in chapter 4 and proposes a Multi Objective Cross Layer Optimization (MOCLO) protocol to improve the performance of both real

time and best effort traffic in terms of packet delay, and throughput. The work proposed in MOCLO protocol proposes a modification to IEEE 802.11e MAC layer to prioritize and protect the important packets within each class of traffic flow. It uses cross layer interaction between PHY-MAC-Network layers to find multiple optimal route to support different flows in AC based on priority thereby reducing the load on a particular route.

- Chapter 6 presents the Adaptive Multi-QoS Cross layer Cooperative Routing protocol (AMCCR) protocol that uses cooperative communication and rate adaptation to guarantee multiple QoS requirements and optimizes the trade-off between end-to-end delay and energy of the system. The protocol uses cross layer communication between physical, MAC, and network layers.

- Chapter 7 concludes the research findings of the thesis and presents the future work to be carried out in connection with the research presented in this thesis.

Chapter 2

Background and Related Work

For the past decade, the field of MANETs has been accepted as a genuine area of research. MANETs provides a spontaneous and robust wireless communication system. Initially, MANETs researchers were focused mainly on designing distributed and dynamic communication protocols for shared channel and for route discovery and offered best-effort protocols to ensure optimum network operation in an unpredictable wireless environment.

Quality of Service is a set of service requirements that work on a network while transmitting high-priority data from source to destination under limited network capacity [34]. Several applications such as multimedia, real time systems require strict QoS support such as delay, bandwidth, reliability, energy consumption, jitter, throughput, etc. IP networks usually offer a best-effort service for data transmissions which is not adequate to support stringent QoS requirement of real time applications. A lot of work has been done to support QoS in the wireless network, but unfortunately, it is not fully exploited in MANETs due to resource limitation and highly dynamic nature of nodes.

2.1 QoS in MANETs

The current research on QoS support in MANETs includes QoS models, QoS resource reservation signaling, QoS MAC, and QoS routing. It is important that all these components coordinate with each other to provide QoS in MANETs.

2.1.1 QoS Models

The QoS model specifies an architecture in which some kind of services could be provided in MANETs. The QoS models should consider the various challenges in MANETs like dynamic topology and time varying link capacity. The QoS models in MANETs must also consider the existing QoS architecture on Internet in order to seamlessly connect with the Internet. The existing QoS model in Internet architecture are InterServ [35] and DiffServe [36]. H. Xiao et al., proposed a new QoS model for MANETs called A Flexible QoS Model for MANETs (FQMM) [37]. FQMM considers the characteristics of MANETs and tries to take advantage of both the per-flow service granularity in InterServ and the service differentiation in DiffServe.

2.1.2 QoS Signaling

QoS signaling is used to reserve and release the resources, and to set up, tear down, renegotiate flows in the network. It acts as a control center in a QoS based system [38]. Two important mechanisms are included in the QoS signaling system.

- Firstly, the QoS signaling information must be reliably shared between the nodes. To achieve this, either in-band signaling, or out-of-band signaling is used [39]. In in-band signaling, the control information is carried along with data packets, whereas in out-of-band signaling, the explicit control packets are used.
- Secondly, the QoS signaling information must be correctly interpreted and accordingly corrective measure should be enabled. QoS signaling coordinates the behavior of QoS MAC, QoS routing and other components such as

admission control and scheduling. The functionality of QoS signaling is determined by the QoS models.

2.1.3 QoS Routing

QoS routing searches for the path with sufficient resources for fulfilling QOS requirements, but does not reserve the resources. It is the responsibility of QoS signaling to reserve the resources along the path determined by QoS routing. The Qos routing should work together with the resource management to establish the paths through the network that meets end-to-end QoS requirements.

2.1.4 QoS MAC protocol

QoS supporting components at the upper layers, such as QoS signaling and QoS routing assumes the existence of a QoS MAC protocol, which solves the problem of medium contention, support reliable unicast communication, and provides resource reservation for real time traffic in a wireless environment. Y. Xiao et al., proposed IEEE 802.11e [40] which is a standard based on IEEE 802.11 to support QoS in WLANs. The QoS is provided by setting up different Access Categories (ACs). Each AC has different contention parameters such as Arbitration Inter-frame Space (AIFS) values, maximum and minimum contention window size, and TXOP. Using these contention parameters, virtual competition among traffic categories is created for accessing the channel.

2.2 Classification of QoS Solutions

The classification of QoS solutions based on the five approaches is as shown in Figure 2.1.

1. Based on the interaction between routing protocol and the QoS provisioning mechanism, the QoS protocols can be classified into two types:

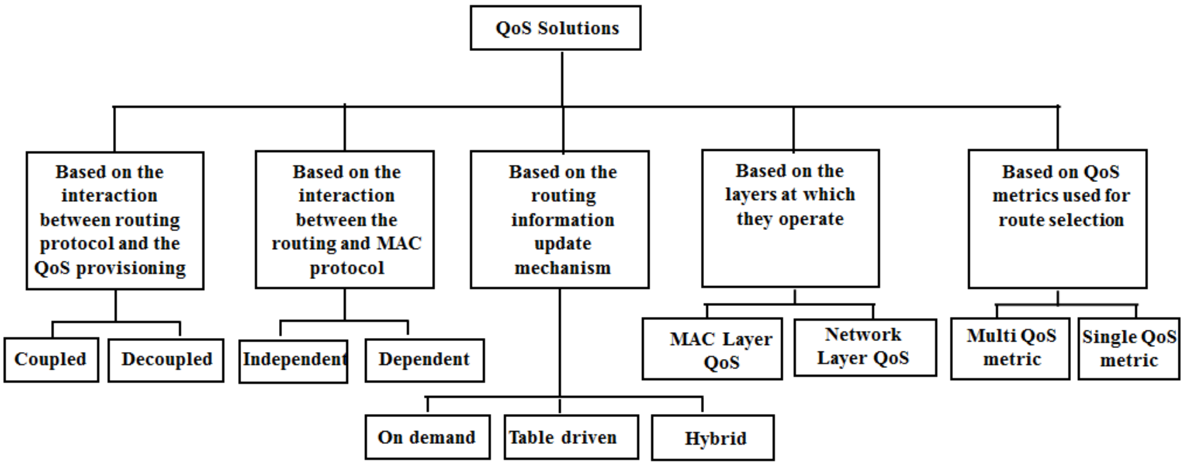


FIGURE 2.1: Classification of QoS solutions

(a) Coupled: In this approach, there will be close interaction between the routing algorithm and the QoS provisioning mechanism for providing QoS guarantees. Some of the popular protocols are: Ticket-Based Probing (TBP) [41], Predictive Location-Based QoS Routing Protocol (PLBQR) [42], Quality of Service Ad Hoc On-Demand Distance Vector (QoS-AODV) [12] and Core-Extraction Distributed Ad Hoc Routing (CEDAR) [11].

(b) Decoupled: In this approach, QoS mechanism is independent of routing protocols. Examples are In-band signaling system for supporting quality of service (INSIGNIA) [43], stateless wireless ad hoc network (SWAN) [44].

2. Based on the interaction between the routing protocols and MAC protocols, the QoS protocols can be classified into two types:

(a) Independent: In this approach, the network-layer protocols are not dependent on MAC. Examples: PLBQR [42], QoS-AODV [12], and SWAN [44].

(b) Dependent: In this approach, the network-layer protocols are dependent on the MAC layer. Examples: Congestion Aware Routing protocol for Mobile ad hoc networks (CARM) [20], and CEDAR [11].

3. Based on the routing information update mechanism employed, the QoS protocols can be classified into three types:

(a) Proactive routing protocols: These protocols are designed to maintain consistency, up-to-date routing information between each pair of nodes in the network by transmitting route updates at a fixed time intervals. Since the resulting routing information is usually maintained in tables, the protocols are sometimes referred to as table-driven protocols. Some of the existing proactive protocols include Destination-Sequenced Distance-Vector (DSDV) protocol [45], Wireless Routing Protocol (WRP) [46], Global State Routing (GSR) [47], Optimized Link State Routing Protocol (OLSR) [48]. The main advantage of proactive protocols is that the nodes can obtain route information quickly but the overhead depends on the number of nodes in the network. Hence, network scaling is an issue.

(b) Reactive or on-demand routing protocols: These protocols establishes a route to a destination only when there is a demand for it and is usually initiated by the source node through the route discovery process within the network. Once a route has been established, it is maintained by the node until either the destination becomes inaccessible along every path from the source or until the route is no longer used or has expired. Reactive routing protocols are very useful in resource constrained environment as it reduces overhead compared to proactive protocols. Some of the reactive routing protocols include Dynamic Source Routing (DSR) [49], Ad-hoc on-Demand Distance Vector (AODV) [50], Temporally Ordered Routing Algorithm (TORA) [51]. Reactive protocols have become the mainstream for MANETs routing, they still have the following main disadvantages:

- High latency in route finding.
- Excessive flooding can lead to network congestion.

Some of the well known existing QoS aware reactive protocols RTLB-DSR [52], MM-DSR [53], QoS-AODV routing protocol [12], RT-TORA protocol [54].

(c) Hybrid protocols: Hybrid protocols uses best functionalities from proactive and reactive routing approaches. The Zone-Based Hierarchical Link-State Routing Protocol (ZRP) [55] is an example of a hybrid protocol that combines both proactive and reactive approaches, thus trying to bring together the advantages of the two approaches. ZRP defines around each node a zone that contains the neighbors within a given number of hops from the node. Proactive and reactive algorithms are used by the node to route packets within and outside the zone, respectively.

4. There are two types of QoS protocols based on the layers at which they operate.

   (a) MAC Layer QoS Solutions: These protocols aims to address the issue of hidden and exposed terminal problems as well as they provide the resource reservation and QoS guarantees for real-time traffic. Some of the well known QoS based MAC protocols are Adaptive QoS MAC Protocol Based on IEEE802.11 in Ad Hoc Networks (AQMP) [56] and Multi-channel MAC protocol (MMAC)[57].

   (b) Network Layer QoS Solutions: These routing protocols work together with resource management to establish paths through the network that meet end-to-end QoS requirements. Some of the QoS based network protocols are Link-stAbility and Energy aware Routing protocol (LAER) [58] and QoS-AODV [12].

5. QoS protocols can also be classified based on QoS metrics used for route selection.

   (a) Single metric constrained: These protocols selects the route based on single QoS metric constrained. For example, the metric could be either delay, bandwidth, or energy level in the particular route. Some of the well known protocols are CEDAR and On-demand QoS Routing (OQR) [59].

   (b) Multi metric constrained: These protocols selects a feasible path that satisfies multiple constraints simultaneously. AQOR and QoS-AODV are typical examples of multi constrained protocols.

2.3 QoS Routing Protocols for MANETs

QoS routing is a key MANET function for the transmission and distribution of multimedia services. It has two main objectives; (1) finding routes that satisfy QoS constraints, and (2) making efficient use of limited resources. The complexity involved in designing QoS based protocol for the routing decision process is mainly due to the requirement of considering multiple objectives at the same time [60]. Also, the QoS routing schemes must provide solutions for distribution of QoS metrics. Generally, QoS routing protocols search for routes with sufficient resources in order to satisfy the QoS requirements of a flow. The information regarding the availability of resources for a QoS feasible paths is managed by a resource management function. The QoS routing protocol should find paths that consume minimum resources according to the relevant QoS metrics.

However, there is no universal QoS routing protocol standard suitable for MANETs [61]. In MANETs, frequency of path break is quite high compared to wired networks. Therefore, QoS routing protocol should respond quickly in case of path breaks and recompute the available resources on the weak path or bypass the broken link without degrading the level of QoS.

2.3.1 Metrics to be Considered for Route Selection

QoS metrics differ from application to application in MANETs. For example, for multimedia applications, the data rate and delay are the key factors. In sensor networks applications, battery life and energy conservation would be the prime QoS metrics, whereas, in military application, security and reliability become more important. For efficient QoS routing implementation, QoS requirements need to be specified in the routing protocol. Consequently, they are used as constraints on route discovery and selection. An application may typically request a particular QoS at a specific layer of the OSI stack by specifying its requirements in terms of one or more metrics. In this section, we list some of the metrics used in routing protocols for evaluating and selecting the path, in order to improve overall QoS to meet the specific requirements of the application.

2.3.1.1 Network Layer Metrics

1. Throughput: It is defined as the acceptable data throughput of a path or node. The acceptable throughput is often referred to as the available bandwidth.

2. Average End-to-End Delay: It is the delay incurred in successfully transmitting the packets from the source to the destination. It includes the delay caused due to packet buffering in the queue, channel contention delay, transmission delay, retransmission delay, and propagation delay. It is considered to be one of the important metric to analyze the QoS routing performance.

3. Node buffer size: It is the number of packets in a node's transmission buffer. Node buffer size determines the amount of delay a packet experiences while traveling through a node.

4. Packet Delivery Ratio (PDR): It is the ratio of the successfully packets delivered to the destination node to the total packets generated at the source node. It is a vital metric to measure the packet losses which directly affects the network throughput.

5. Network Lifetime (NL): It is statistically calculated as expected lifetime of a network.

6. Normalized Routing Load (NRL): It is the number of routing packets exchanged per data packets delivered to the destination node. NRL metric measures the efficiency of routing protocol over a low bandwidth and congested wireless networks.

2.3.1.2 MAC Layer Metrics

1. MAC delay: It is the time taken to transmit a packet between two nodes in a contention-based system. It includes the total time required to transmit the packet

from a node till it receives the acknowledgement. It provides an indication of packet traffic.

2. Link stability: It is described as the predicted lifetime of a link connection. It indicates the length of time during which node pairs are connected.

3. MAC energy efficiency : It is the ratio of energy used for sending data to that of total energy expended for data along with MAC headers and control frames.

2.3.1.3 Physical Layer Metrics

1. Signal-to-interference ratio (SIR): It is the received carrier signal-to-interference ratio at a destination node and can be used as a routing metric indicating link quality.

2. Bit error rate (BER): It is defined as the rate at which errors occur in a transmission system or number of retransmissions required over a connection.

3. Node relative mobility: It is measured as the ratio of the number of neighbours that change over a fixed period to the number that remain the same.

2.4 Cross Layer Design

In this section, two architecture design, layered and cross layered, along with motivations for adaptation of cross layer design and existing cross layer designs are discussed.

2.4.1 Layered Architecture v/s Cross Layered Architecture

In traditional packet-based network architecture, the communication modules are organized into protocol layers which are nested with multiple level of abstraction. In each protocol layer, the metadata which controls the packet delivery are organized into protocol headers [62]. The network functionalities and services are commonly classified and modeled through the well-known Open System Interconnection (OSI) network model [63]. The OSI model represents a 7-layer protocol stack. Each of the layers is modular and defines the specifications for a particular network aspect and provides services to the upper layer. Inside a protocol stack, exchange of control and data information take place only between adjacent protocol layers and is supported by the concept of a service access point (SAP). A SAP provides access to a selected subset of protocol functionality via a precisely defined set of primitive operations. A particular protocol layer may offer or use more than one SAP, depending on its function and the information it needs to exchange with its adjacent protocol layers. This paradigm of the OSI reference model has been the predominant design approach since the emergence of all modern networking architectures [64].

A protocol at a given layer is implemented by a software, or a hardware, which communicates with other system over the network by Protocol Data Units (PDUs). A PDU is constructed by payload, which consist of data generated by an entity at a higher layer and a header which consist of protocol information. The main concepts motivating layering in ISO/OSI model are as follows:

- Each layer performs a subset of the required communication functions
- Each layer relies on the next lower layer to perform more primitive functions
- Each layer provides services to the next higher layer
- Changes in one layer do not require changes in other layers

The main advantage of the layering pattern is the modularity in protocol design, which provides the interoperability and enhanced design of communication protocols. A protocol within a given layer, defines the functionalities it provides, while implementation details and internal parameters are hidden from the rest layers through encapsulation. This encapsulation prevents the layers from coordinating with each other to provide user defined or application specific services. To mitigate this side effect of the encapsulation between the abstract layers in the OSI model, a number of designs have been proposed [65] which violates the reference architecture by (i) creating of new interfaces between the levels, (ii) merging of adjacent layers, (iii) coupling without new interfaces. Thus, any attempt to violate the OSI reference model is considered a cross layer design. Hence, cross layer design justifies the cross layer interactions from the physical to the transport layer in order to allow information transfer across the boundaries of the layers. Cross layer designs do not abolish the layered structure of the protocol stack, but provide the inter-layer communication between two non-adjacent layers. The cross layer design also discloses the internal status and parameters of each layer to other layers by creating new interfaces between layers and redefining the layer boundaries. Additionally, cross layer designs may allow a layer to determine its behavior based on the data that it retrieves or receives from the other layers. Therefore, cross layer designs indicate that each layer is able to share parameters, status, and other information with other layers, without breaking the protocol stack structure. However, there is no specific reference model that specifies the functionality of each new module that must be realized in a cross layer design solution.

2.4.2 Cross Layer Motivations

Recently, as the wireless communications and networking quickly occupying the center stage in research and development activity in the area of communication networks, the effectiveness of layered protocol architecture has grabbed a lot of attention of the research community across the globe. Time after time it is argued that although layered architectures has been effective for wired networks, they are not suitable for wireless networks. The motivation behind violating the layered architecture in wireless network is listed below :

1. A possibility of opportunistic communication modalities offered by the wireless link in a wireless network. For example, the nodes can exploit the broadcast nature of the channel and cooperate with one another in forwarding the overheard packets. Such novel mechanism of communication requires violating the layered architecture.

2. Wireless has several link handling issues, that the new service demands of protocol design cannot be dealt well in the framework of the layered architectures. A typical example is, when a TCP ack packet in the wireless network, it is interpreted as a network link congestion issue. But the reason for the delayed packet, could also be link failure or channel error.

In literature, there are diversified interpretations made on cross layer design. The first possible reason is that, different research scholars entrust on optimizing the different layers of the stack. Second, there are no specific guidelines about when to adapt cross layer design and which layer should participate, with how much of capacity. And third, the negotiation between the network performance and implementation is blur. The numerous research efforts from around the globe have proposed cross layer design to address different problems that arise due to the evolution of wireless mobile communications. The provision of Internet services over mobile communication networks has been the driving force in this evolution. The three main motivations supporting the adoption of cross layer design in protocol design for MANETs:

1. protocols needs to be adaptive to dynamics of network,
2. to support the specific QOS requirement by the applications, and
3. to tackle the energy and security constraints.

There are several design challenges in MANETs namely: security, an energy issue, topology control to deal across the layers, and requires joint solutions involving multiple protocol layers. In order to serve the largest possible network of users in next-generation mobile communication networks can be satisfied through cross layer design that exploits valuable properties of the wireless channel. The cross layer design optimization solutions can provide improved QoS to the mobile terminal for its multimedia applications.

2.4.3 Cross Layer Signaling Architectures

Given that there are no particular restrictions in the way cross layer signaling takes place, various cross layer design approaches have been proposed in the literature. In the following section, the kind of architecture violation that has taken place in a particular cross layer design is described. In [65], Raisinghani et al. studied cross layer architectures proposed by several research fraternity and mentioned about the objectives of the cross layer signaling model which includes rapid prototyping, portability, and efficient implementation of the cross layer design while retaining the impact on TCP/IP modularity. The author classified the cross layer architecture into four categories:

1. Inter-layer signaling pipe: Wang et al. [66] proposed a cross layer signaling technique using inter-layer signaling pipe which propagates the message signals within the layers along with a packet data stream and can be associated with a particular packet incoming or outgoing from the protocol stack in bottom-up or top-down approach. The methods considered for signal information encapsulation and its propagation from one layer to another are achieved through either packet headers or packet structures. Figure 2.2 describes the design.

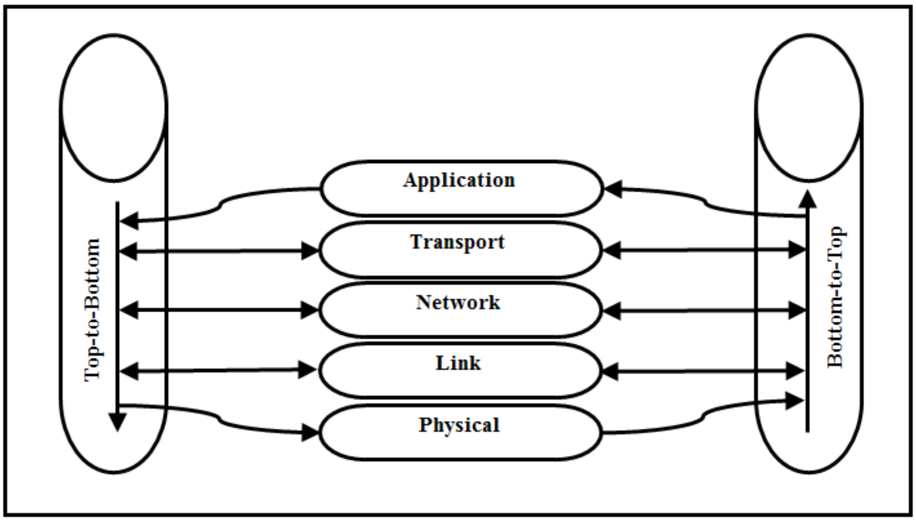


FIGURE 2.2: Inter-layer signaling pipe

- Packet headers: The packet header is used to carry an inter-layer message where the signaling information is accommodated into an optional portion of

IPv6 header. This packet header follows the processing path of a packet and can be accessed by any of the layers. The main drawback of this technique is the limitation in the direction of the signal is in line with direction of packet flow. This is not an effective method as cross layer design also requires signaling in the opposite direction too.

- Packet structures: In this technique, the signaling information is incorporated into the packet structure. Whenever a packet is generated by the subsequent layer or received from the network interface, a corresponding packet structure is allocated. The structure contains all the related information such as protocol headers and data, the protocol stack information such as network interface id, socket descriptor, configuration parameters and other. The packet structure does not disturb the existing functionality of the protocol stack layers. If the cross layer signal is not meant for certain layers, the corresponding layer will not access or modify the packet structure provided by the other layers.

2. Direct Inter-layer Communication: In order to improve the inter-layer signaling pipe scheme, Qi Wang et al.in [66] proposed Direct Inter-layer Communication scheme known as Cross Layer Signaling Shortcuts (CLASS) which is as shown in Figure 2.3. In CLASS approach, a non-neighboring layers of the protocol stack can exchange messages without processing at every adjacent layer. The CLASS messages can be used for bidirectional signaling. This results in fast signaling and apt information delivery. Internet Control Message Protocol (ICMP) is one of the direct signaling communication protocols. However, the ICMP approach is mostly limited by one to one request-response action, not considering some of the complicated event-based signaling approach. Also, signaling with ICMP messages incurs operations with extra protocol headers (IP and ICMP), checksum calculation and other processing overhead.

3. Central Cross layer Plane: Figure 2.4 shows the Central Cross layer Plane which is the most widely used cross layer signaling architecture implemented in parallel to the protocol stack. In [10], Chen et al. proposed a common database that can be accessed by all layers. The common database is like a new layer,

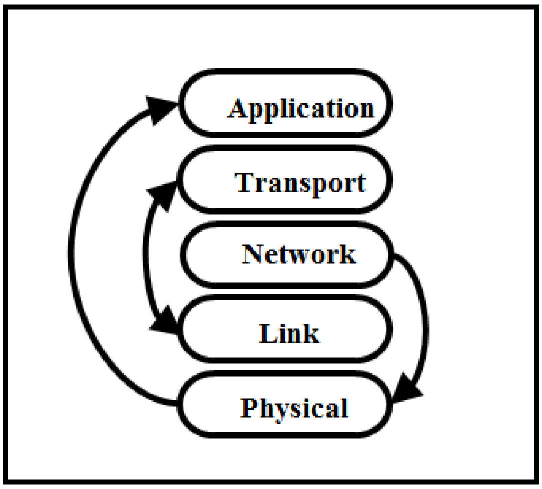


FIGURE 2.3: Direct inter-layer communication

providing the service of storage/retrieval of information to all the layers. The shared database approach is particularly well suited to vertical calibrations across layers. The vertical calibration refers to adjusting the parameters that span across layers. The main motivation is that, the performance seen at the level of the application is a function of the parameters at all the layers below it. Thus, joint tuning can help to achieve better performance than individual settings of parameters independently. Consider an example of vertical calibration where the delay requirement dictates the persistence of link-layer automatic repeat request (ARQ) [67], which in turn becomes an input for deciding the rate selection through a channel-adaptive modulation scheme. Vertical calibration can be done in a static manner, which means setting parameters across the layers at design time with the optimization of some metric predetermined. It can also be done dynamically at runtime. Raisinghani et al. in [65] proposed a model called ÉCLAIR to implement the cross layer feedback to be communicated between any pair of layers in the protocol stack. ÉCLAIR comprises of the two modules, tuning layer and optimizing subsystem. In ÉCLAIR, the central system plane communicates with the protocol stack by a cross layer interfaces called tuning layers. Each tuning layer utilizes a set of API functions to read or write to the internal protocol control information and data structures. The optimizing subsystem (OSS) activates optimization algorithms. The OSS collects control information from the tuning layer through the protocol optimizers and adapts the

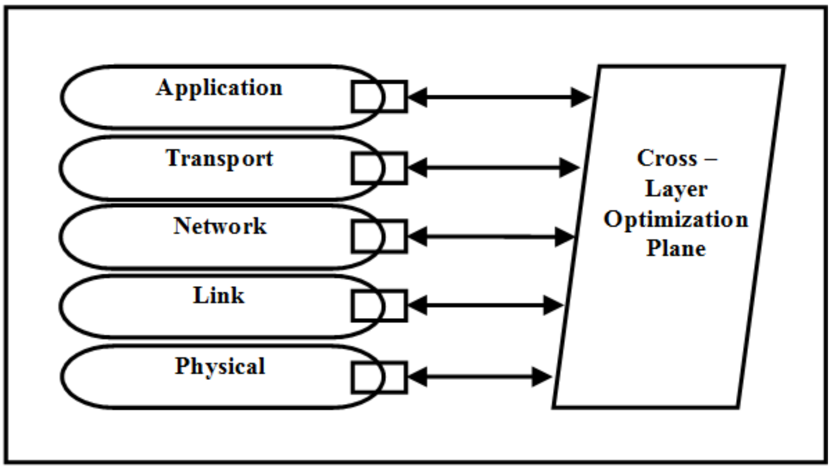


FIGURE 2.4: Central cross-layer plane

protocol's behavior during runtime.

4. Distributed Cross Layer Signaling: Most of the above discussed techniques aim at defining cross layer signaling between different layers belonging to the protocol stack of a single node. Many of the recent optimization proposals perform cross layer optimization based on the information obtained at different protocol layers of distributed network nodes. This adds another degree of freedom in how cross layer signaling can be performed in a distributed manner.

   However, several optimization proposals exist which perform cross layer optimization based on the information obtained at different protocol layers of distributed network nodes. This corresponds to network-wide propagation of cross layer signaling information, which adds another degree of freedom in how cross layer signaling can be performed. Figure 2.5 explains the distributed cross layer signaling mechanism.

2.5 Related Work

In QoS-aware routing protocols, since the path selection is based on the desired QoS metrics, the routing protocol can be termed as QoS-aware. As the path selection is based on the desired QoS metrics, the routing protocol is termed as QoS-aware routing protocols. In the literature, numerous QoS based/aware routing protocols have been

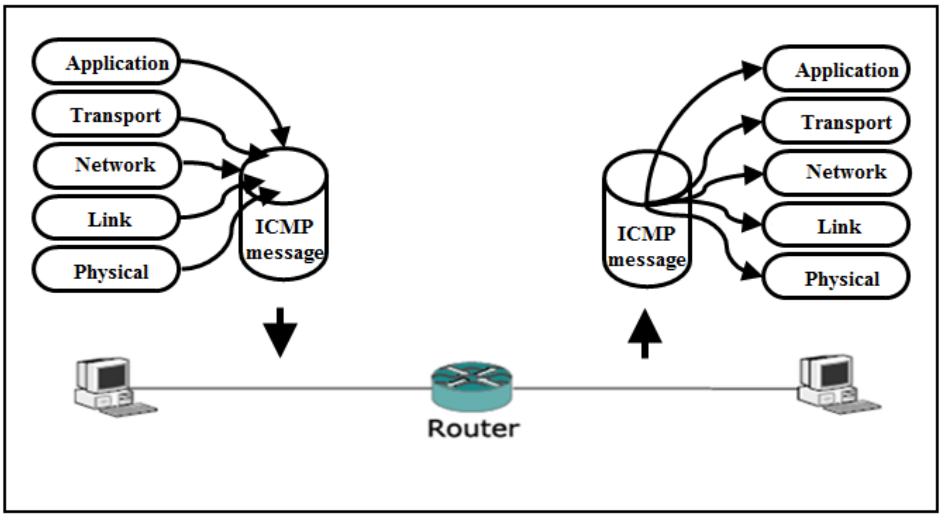


FIGURE 2.5: Distributed cross-layer signaling

proposed for finding paths. In the following section, some of the popular QoS routing protocols are described.

The first major work on MANET QoS was the INSIGNIA framework proposed by Lee et al. [43] where resources are reserved in an end-to-end manner through a Resource Reservation Protocol (RSVP). This QoS framework is designed to support adaptive services as a primary goal in ad hoc networks. It allows packets of audio, video and real-time data applications to specify their maximum and minimum bandwidth needs. This framework also plays a central role in resource allocation, restoration control and session adaptation between communicating mobile hosts. Based on the availability of end-to-end bandwidth, QoS mechanisms attempt to provide assurances in support of adaptive services.

Core-Extraction Distributed Ad Hoc Routing (CEDAR) [11] is a reactive routing protocol consists of three key components: core extraction, link state propagation and route computation. During core extraction phase, a self-organizing routing infrastructure known as "core" is formed which consists of a subset of nodes in the network that are distributed and dynamically elected by using only local state of the core. During Link state propagation phase, bandwidth availability information of stable high bandwidth links is shared with other core nodes, whereas information about dynamic links or low bandwidth links is kept local. In route computation, a path from the domain of the source to the domain of the destination is established by using the directional information provided by the core path. Setting up a core path requires one round trip delay, and the QoS route computation requires another round trip delay. Therefore, the total time required to discover a route is two round-trip delays. Also, if the core node is moving out of the selected route, rerouting is very costly.

To guarantee bandwidth for real-time applications, the On-demand QoS Routing (OQR) protocol was proposed by C.R. Lin et al. [59]. Since routing is on-demand in nature, there is no need to exchange control information periodically and maintain routing tables at each node. The OQR protocol uses TDMA scheme for channel access and bandwidth as the key QoS parameter. The OQR protocol uses an on-demand resource reservation scheme and hence achieves lower control overhead.

Conditional Max-Min Battery Capacity Routing (CMMBCR) algorithm is proposed in [59]. This algorithm chooses the route if all the nodes in the route have remaining battery capacities higher than a threshold. This method considers both the total transmission energy utilization of routes and the residual power of nodes. When all nodes in some probable routes have enough residual battery capacity, a route with minimum total transmission power among these routes is selected.

Shengming Jiang et al. [68] have proposed a prediction based link availability estimation model for MANET. This model predicts the probability of an active link between two nodes being continuously available for a predicted period based on the movement of current nodes. They used the exponential distribution model for prediction of link availability. In this model authors considered the temporary changes in the node movement, but did not consider the frequently occurring changes which may affect the link prediction.

Perkins et al. proposed a QoS-AODV routing protocol [12] with added features to his earlier work [69]. This enhanced protocol helps in choosing the optimal path from source to destination using hop count and load rate as a basis along with delay and available bandwidth. To support this, extra fields are added in RREQ message. The route table entry corresponding to each destination also contained additional fields such as maximum delay, minimum available bandwidth as well as a list of sources requesting delay and bandwidth guarantees. QoS-AODV minimizes the transmission of RREQ packets by not forwarding them to the nodes which are noncompliant to the QoS requirements and hence save the overhead in the routing establishment process. However, the protocol does not take into account the dynamics of topology changes in MANETs due to nodes moving out of range or due to a node failure causing link failure. This results in inaccurate delay estimates.

Adhoc QoS-aware routing protocol, AQOR [13] is an on-demand routing protocol based on metrics like available bandwidth and end-to-end delay. When a route is needed, the source node initiates a route request, in which the bandwidth and delay requirements are specified. To estimate the available bandwidth, each node puts its reserved bandwidth in periodic Hello messages and are sent to their neighbors. With the knowledge of available bandwidth and end-to-end delay, the sufficient bandwidth path with minimum delay is chosen as the QoS route. A drawback of AQOR's bandwidth estimation method is that it assumes the interference range is same as the transmission range, which is not true in general. Thus, AQOR's bandwidth estimation method will not correctly incorporate the bandwidth being used by neighbors in the interference range of the node.

SQ-AODV [70] is a Stability-based QoS-capable Ad-hoc On-demand Distance Vector protocol, which is an enhancement of the AODV protocol focusing on, how residual node energy can be used for route selection and maintenance and also how protocol can quickly adapt to network conditions. The uniqueness of this scheme is that it uses only local information, and requires no additional communication or cooperation between nodes. It possesses a make-before-break capability i.e., it finds alternate route before the existing route fails to forward the packets due to energy depletion of the node. Make-before-break capability of the protocol minimizes packet drops and the protocol is compatible with the basic AODV data formats and operation, making it easy to adopt.

Reward-based routing protocol (RBRP) is another protocol based on AODV. They extend the route discovery process using the Q-learning strategy to select a stable route to enhance the network performance. This technique improves performance achieved with AODV through an enhanced route selection based on hop count, bandwidth, battery power and speed of mobile nodes. However, it does not take into consideration some important constraints inherent in the mobile ad hoc networks such as mutual interference of the nodes. This lead to inaccurate estimation of the available bandwidth.

A QoS-aware routing protocol with an adaptive feedback scheme called (AQA-AODV) [14] have been proposed for video streaming in MANETs. AQA-AODV introduces a mechanism to estimate link and path available bandwidth. It uses an adaptive scheme that can provide feedback to the source node about the current network state, to allow the application to appropriately adjust the transmission rate. The scheme proposes a route recovery approach, which provides a mechanism to detect the link failures in a route and re-establish the connections taking into account the conditions of QoS that have been established during the previous route discovery phase.

Delay Aware AODV Multi-path (DAAM) routing protocol [71] has been proposed which combines the features of QoS-AODV [12] and AODV-Multipath [72]. This protocol enables the computation of multiple node-disjoint paths without incurring the overhead generated by link-state routing methods. The cumulative delay during the route discovery process from the source node to destination node is recorded by each node. The disadvantage of DAAM is that route delay information might not always be up to date as mentioned in this study. The efficiency and functionality of DAAM did not compare to/with other QoS protocols and the authors recommended this as a future work.

R. Asokan et al. proposed Energy and delay-aware AODV (ED-AODV) protocol [73] which is a modified version of AODV with additional delay and energy extensions is proposed. The two parameters minimum energy and maximum delay are added to the AODV routing table entry. A source node transmits a route request RREQ packet containing the QoS energy and delay extensions. On receiving the RREQ packet, an intermediate node compares its available energy and time delay to reach the node to the energy and delay field indicated in the RREQ Packet. RREQ packet are discarded if QoS requirements are not fulfilled.

A Quality-of-Service aware Source initiated Ad-hoc Routing (QuaSAR) protocol [74] provides quality control mechanism in a network that is subjected to frequent topological changes due to mobility. The protocol gathers the information related to battery power, signal strength, and latency/delay during the route discovery phase and uses this information to select the best possible route. The main drawback of QuaSAR is that, the RREQ packets must carry all the four QoS parameters. It also suffers from poor scalability, since the source routing can cause the length of route request and route reply messages to become lengthy in larger networks.

Next section presents some of the well known cross layer solutions involving Physical (PHY), Medium Access Control (MAC), Network (NET) and Transport (TRA) layers.

L. Doss et al. [75] demonstrated that when physical layer properties such as path loss and shadowing are considered for route construction, network performance is enhanced when compared with the results provided by a simple free propagation model. Hence, the authors concluded that the hop-count may not be an optimal metrics for the routing process and that the routing metrics for MANETs should take into account the current state of the channel as well as the quality of each link. This protocol proposes a cross layer framework between the physical layer and network layer for designing efficient routing protocol.

Rajashekhar Biradar et al. [76] proposed a mesh based multicast routing scheme that finds stable multicast path from source to receiver based on cross layer interaction between physical and network layer. The multicast mesh is constructed by using route request and route reply packets with the help of multicast routing information cache and link stability database maintained at every node. The stable paths are found based on selection of stable forwarding nodes that have high stability of link connectivity. The link stability is computed by using the parameters such as received power, distance between neighboring nodes and the link quality that is assessed using bit errors in a packet.

K. Chen et al. [10] proposed a cross layer optimization approach for multimedia applications running over ad hoc networks that involves information sharing between the routing and middle-ware layers. This approach aims to provide end-to-end QoS guarantees in resource-limited mobile ad hoc networks. In Chen's model, the application layer generates and shares multimedia data with other users in the network. The middle-ware layer is responsible for locating, accessing, and replicating data to applications through the data accessibility service. At the network layer, this approach adopts a predictive and proactive location-based routing protocol. The framework includes a system profiles component through which the routing layer and middleware layer share information.

Girici et al. [77] proposed a cross layer approach based on joint routing and link scheduling for ad hoc networks. The approach aims at making routing and link activation decisions that are aware of three QoS factors: energy, delay, and network lifetime. To enable QoS-aware decisions, the nodes store the state of neighboring nodes, including residual battery energy, available number of transceivers, and transmission power requirements of each neighbor. The purpose of this approach is to enable a node to select the best next hop for each packet transmission according to energy, delay, and network lifetime considerations. The protocol introduces a new link cost metric that considers three terms: transmission energy per packet (physical layer QoS), expected volume on a node's links throughout the network lifetime (data link layer QoS), and the queue length along the directed link (network layer QoS). This approach adopts a slotted time structure with a separate time-slotted control channel for nodes to reserve the slots for data delivery.

Zouhair El-Bazzal et al. proposed Turbo AODV (TAODV) protocol [78] which is a cross layer design for routing in mobile ad hoc networks. The design allows the physical layer sharing the received signal strength (RSS), and the medium access control (MAC) layer sharing the remaining energy, as well as the remaining queue length information with the network layer. Based on this information, the network layer computes a weight value that forms the routing metric used for route selection in the enhanced protocol. The routing packets headers are changed by adding new fields, routing packets forwarding algorithms are modified and the route selection algorithm at the destination is completely enhanced.

Jian Chen et al. proposed a routing protocol called QoS-AOMDV [79]. The QoS-AOMDV is based on cross layer design, which cooperate in sharing network-status information in different layers while maintaining the layers separation to optimize overall network performance. The information in network layer and data link layer are combined to make a general criterion, so that it could satisfy the dynamics of ad hoc network. The protocol exploits a cross layer design to share the energy information in network and the queue length of buffer in data link layer to prolong the network lifetime and reduces the network congestion. Protocols use the state information flowing throughout the stack to adapt their behavior accordingly. The QoS-AOMDV proposes multiple routes between source and destination. In the route selection phase, the destination nodes do not reply route request (RREQ) immediately. Route reply (RREP) is carried out based on the general cost criterion. In transmission phase, data is transmitted in multiple paths one by one to balance the energy and the traffic loads in multiple paths.

2.6 Conclusion

Wireless ad hoc networks have several characteristics that distinguish it from wired networks. The strict layer design approach is not suitable and does not function efficiently in wireless networks. As a solution, there has been currently a rise in the use of cross layer design. The wide range of cross layer proposals discussed in the chapter, demonstrates the popularity of the cross layer approach in the research community. However there are also some open research problems limiting the development of systematic techniques for cross layer design in wireless ad hoc networks. From the protocol design point of view, existing studies on cross layer design utilizes cross layer information to enhance the quality of the discovered routes, but each mechanism adapts either a single routing metric or mostly focused on solutions involving at most two protocol layers. Moreover, most of the discussed protocols, do not provide much of the QoS provisioning for the real time multimedia traffic.

Chapter 3

Cross Layer Best Effort QoS Aware Routing Protocol

3.1 Introduction

In recent times, there has been an immense growth in demand for support of multimedia applications in MANETs. Most of the real-time multimedia traffic tends to be in bursts, bandwidth demanding and is responsible for the congestion. Congestion leads to packet losses, retransmissions, bandwidth degradation and also incur additional time and energy depletion during congestion recovery. For real time multimedia traffic, the data rate and delay are the crucial QoS factors [80]. Therefore, to satisfy these QoS requirements, each route in the network should provide a correct estimate of the available data rate and end to end delay.

In this chapter, a detailed design of adaptive and reliable congestion control, multi-rate adaptation routing protocol “Cross Layer Best effort QoS aware routing protocol (CLBQ)” is presented. The protocol implements cross layer interaction between PHY, MAC and network layer to utilize multi-rate supports from 802.11b. To provide better network efficiency in delay constrained applications, the protocol considers link quality, data rates and MAC delay as QoS parameters. The optimal data rates between the links are chosen based on the estimated delay to admit a flow with certain delay requirement. The MAC layer delay and instantaneous network queue load are considered while evaluating retransmission attempt due to collisions. The protocol chooses less congested routes rather than the shortest routes.

3.2 A Literature Review

The main design goal of network design is to allocate network resources effectively and fairly among the nodes in the network. The most commonly shared resources are link bandwidth and the node queues, which are limited in capacity. While the node is contending for the channel access, packets get buffered in these node's queues. When multiple neighbor nodes are contending for the same wireless channel, the buffer overflows and packets have to be dropped. When such drops become frequent, the network is said to be congested.

In traditional ad hoc routing protocols, it is assumed that all packet losses are due to congestion. This could be true in a static network. But in case of MANETs, packet losses may be due to broken routes as well as congestion [81]. In case of route failure, the route discovery process is to be reinitiated to find a new route to the destination. Performing route discovery in an already congested network may further increase the congestion. Therefore, in a resource constrained MANETs, it is necessary to use proper congestion control mechanisms to improve the network performance [82].

Wireless standard IEEE 802.11a/b/g support rate adaptation to accommodate time-varying channels. More often, the IEEE 802.11 DCF uses low data rate next hop node for transmission under heavy traffic conditions which results in decrease in the throughput. In multi-rate networks, if lower data-rate link followed by the higher data-rate link, packets will build up at the node heading the lower data-rate link, leading to long queueing delays [20]. This further leads to packet drop and causes congestion due to packet retransmission. This is more evident in a traffic with intensive data such as multimedia and has a negative influence on the QoS. Unlike well-established networks such as the Internet, in a dynamic network like MANETs, it is expensive in terms of time and overhead to recover from congestion. Some of the recent routing protocols which provide the solution for the above problems are described below.

3.2.1 Congestion Aware Protocols

X. Chen et al. proposed Congestion Aware Routing protocol for MANETs (CARM) [20], which is an improvement of DSR protocol. It gathers information from the MAC layer to find congestion free routes. CARM uses a link data-rate categorization approach to prevent routes with mismatched link data-rates. It employs Weighted Channel Delay (WCD) parameter using which a cost to each link in the network is assigned based on the MAC overhead, and buffer queuing delay. WCD metric measures the congestion level and adopts a route Effective Link Data- rate Category (ELDC) to avoid the Mismatched Data-Rate Route (MDRR) problem. However, CARM does not provide any mechanism to adjust the data rates when congestion occurs.

Yi Lu et al. proposed a Congestion-Aware Distance Vector (CADV) [83] protocol based on proactive protocol DSDV, which maintains the route information of all the nodes in the network. The routing decision is made based on the hop count to the destination node and the estimated delay which is a measure of congestion at the next hop. The CADV gives higher priority to the route with lower estimated delay. However, estimated delay is a weaker metric to measure the congestion condition. Network scalability is limited due to excessive control overhead. CADV is not feasible for MANETs where frequent topological changes are inevitable and incurs high overhead for maintaining the routing table updates.

Joo-Han Song et al. proposed Delay-based Load-Aware On-demand Routing (D-LAOR) [22] protocol, which determines the optimal path based on the estimated total path delay and the hop count. D-LAOR protocol is an extension of the AODV protocol. It provides a mechanism for the source node to select the less congested path through which Route Request (RREQ) packets have traversed. In D-LAOR, the RREQ packets at congested node is dropped thereby preventing the congested node from becoming an intermediate node in the path. D-LAOR determines the congested node by comparing the estimated total node delay and the number of packets being queued in the interface of a node in the RREQ packet-forwarding path. In this protocol, the data rate for packet transmission is assumed to be static and moreover it does not provide a mechanism to overcome the congestion.

Duc A. Tran et al. [84] proposed a Congestion-adaptive Routing Protocol (CRP) which prevents congestion from occurring in the first place, rather than dealing with it reactively. The important aspect in CRP design was the bypass concept. A bypass is an alternate path connecting a node and the next non-congested node. If a node is aware of a potential congestion ahead, it finds a bypass path that can be used in case the congestion actually occurs or is about to occur. Part of the incoming traffic is split through the bypass path and the congested path, thus minimizing the traffic coming towards the potentially congested node. Since CRP adapts to the congestion, the queuing delay is minimized. The main drawback of this mechanism is, packets are transmitted at a static data rate. The protocol uses single congestion metric, which is the ratio between the number of packets currently buffered to the buffer size, and this metric is not sufficient to predict the congestion.

The above routing protocols are congestion-aware protocols which discover the route that are less congested. These protocols do not suggest a mechanism to adapt to the congestion if takes place. In the above schemes, the congestion is taken into consideration only when establishing a new route and remains the same until node mobility or link failure results in route disconnection.

3.2.2 Rate Adaptation Protocols

The IEEE 802.11 family of standards for WLAN is illustrated in Figure 3.1. The IEEE 802.11 MAC layer uses Distributed Coordination Function (DCF) as the fundamental MAC scheme which is based on the Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) technique. IEEE 802.11 a/b/g support multirate capability which has been studied at length by several researchers to exploit the possibility of improving the network performance [85]. The IEEE 802.11a standard is designed to work in the 5 GHz band and provides 52- subcarrier orthogonal frequency division multiplexing with discrete data rates up to 54 Mbps. The IEEE 802.11b standard is designed to work in the 2.4 GHz band, with four transmission rates, 1, 2, 5.5 and 11 Mbps. IEEE 802.11g is an improvement over IEEE 802.11 with rates up to 54 Mbps using the same frequency band 2.4 Ghz as IEEE 802.11b. These high transmission rates are possible through efficient modulation schemes that are optimized for the channel conditions. Use of high data rates provided by IEEE standard has some

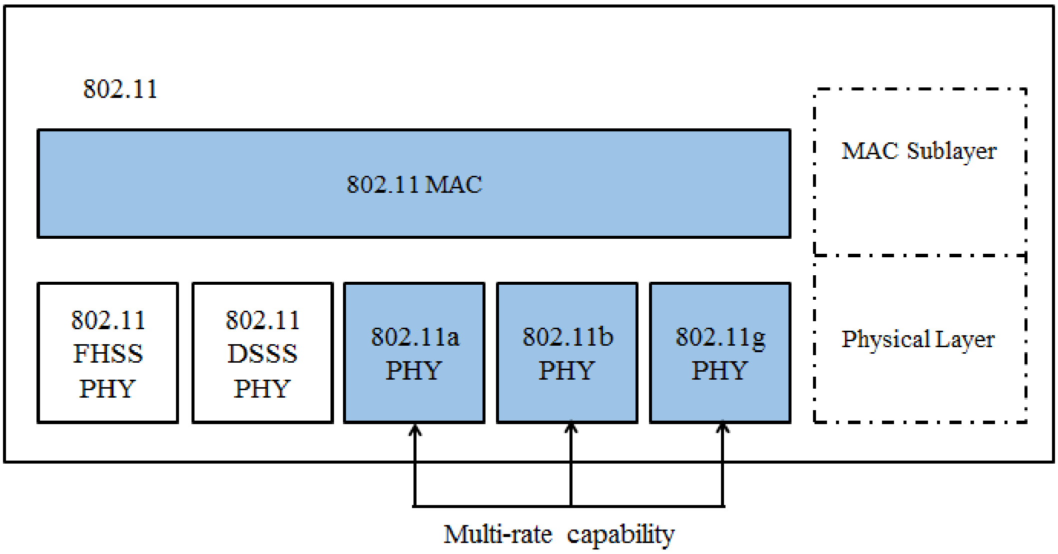


FIGURE 3.1: IEEE 802.11 family of standards

trade-offs. High data rate requires a high signal to interference noise ratio (SINR), which is achieved through an efficient modulation scheme. The Physical layer of the IEEE 802.11 protocol stack uses several modulation and coding schemes, namely, DSSS, FHSS, and OFDM. Due to dynamic nature of MANET and volatile nature of channel, the current state of the channel should be accessed to decide about the modulation scheme to be applied to support appropriate IEEE 802.11 protocols [86].

The range of the link depends on the data rate, i.e., higher the data rate, shorter will be the range. Achieving high data rate and long range simultaneously is practically infeasible [87]. High data rates can cause frequent link failures in case of mobility. This leads to increase in routing overhead due to rediscovery of the routes. Therefore, to support higher data rates, the number of hops between the source and the destination node has to be increased.

Many researchers have used rate adaptation technique which uses variable data rate to cope up with the deteriorating channel condition. With rate adaptation, the link can be used for a longer period of time, hence, eliminating the need for re-route discovery [85]. For example, by adapting the data rate to be lower on long-range links, nodes far away can be offered a limited service instead of being disconnected from the network. Overall, the ability to adapt the data rate offers possibilities to increase capacity in ad hoc networks [19].Effectiveness of the rate adaptation scheme depends on the relative fairness among the competing nodes and optimal throughput. A good rate adaptation scheme should handle transmission failures which may occur due to channel errors or packet collisions. Some of the well-known rate adaptation schemes are described below.

The Automatic Rate Fallback (ARF) [17] is a simple and widely used sender based rate adaptation scheme. In ARF, if two consecutive ACKs are not correctly received by the sender node, it then sends the next data transmission at a lower rate. If a source node receives 10 consecutive successful ACKs, then the next data transmission takes place at the next higher rate. The main drawback of ARF is decrease in the throughput when the source node decreases its rate even though transmission failures are caused by collisions. Moreover, the ARF does not ensure fairness between the competing nodes. ARF is suitable for WLAN with the centralized control unit.

M. Lacage et al. proposed an Adaptive Auto Rate Fallback(AARF) [88] protocol to optimize the ARF performance in stable channel conditions. As similar to ARF, AARF switches to previous lower rate when the transmission packets fails but doubles the number of consecutive successful transmissions needed to switch to a higher rate. This helps to reduce unnecessary attempts in stable conditions. In a single user environment AARF outperforms ARF scheme, but performance degraded under multi-user environment.

Y. Xi. et al. proposed Adaptive Multi-rate Auto Rate Fallback (AMARF) [89] scheme for IEEE 802.11 WLANs. In AMARF, each data rate is assigned a unique success threshold, which is a criterion to switch from one data rate to the next higher data rate. The success threshold is dynamically changed in an adaptive manner according to the current network conditions, such as packet length and channel parameters. The

AMARF protocol is implemented without any modification to the existing IEEE 802.11 standards. Although, AMARF outperforms the ARF scheme, it does not take into account the fairness for competing nodes.

Seongkwan Kim et al. proposed Collision-Aware Rate Adaptation algorithm (CARA) [18]. Unlike ARF and AARF, CARA differentiates between frame collisions and frame transmission failures caused by channel error. It uses the RTS/CTS mechanism to estimate the quality of the channel. A sender node decrements its rate only in case of consecutive channel errors, but not in case of collisions. The RTS/CTS mechanism is enabled only when the number of transmission failures reaches a certain degree. When hidden terminals exist, it suffers from the drawback of RTS oscillation. The CARA performs better in terms of throughput than ARF. However, CARA scheme is well adapted for WLANs, yet not suitable for MANETs and does not support system fairness.

The first receiver based algorithm is proposed by Gavin Holland et al. is Receiver-Based Auto Rate (RBAR) [19] where rate adaptation mechanism is initiated by the receiver instead of the sender. RBAR uses the RTS-CTS mechanism to estimate SNR (signal-to-noise ratio). It then uses pre-defined SNR thresholds to evaluate the best possible rate for the channel. An RBAR sender always sends an RTS packet before transmitting the data packet. The receiver selects the data rate based on signal strength of received RTS packet and responds to the RTS packet of the sender with a CTS packet which contains the suggestion to alter the transmit power for data transmission. Sender switches to new data rate and alters the transmit power if CTS packet contains a message about altering transmit power. The receiver adjusts its data rate for sending ACK or NACK after the data transmission is completed.

Several researchers have shown that due to the channel dynamics for MANETs, significant time is spent only on retransmissions. Hence, rate adaptation schemes need to undergo further improvement to minimize losses due to retransmissions and optimally utilize network resources. It is imperative to have extensive, real and correct measurements made available to enable better study and understanding of the traffic and channel models. Rate adaptation schemes require these measurements to build, evaluate and optimize their strategies.

Thus, the probable solution in multi-hop networks could be, to employ a routing protocol which can support rate adaptation and can give priority to higher data rate links while building a route. Also, the routing protocol is required to provide acceptable level of QoS to delay sensitive application, so that the traffic is routed through less congested routes in the network. The solution should provide a mechanism to overcome the congestion effectively and efficiently with minimum overhead.

3.2.3 Multipath Routing Protocols

Multipath routing is the routing technique of using multiple alternative paths through a network for data communication. The main benefits of multipath routing is fault tolerance, increased bandwidth, or improved security. AODV [50] is considered to be one of the most popular routing protocols in MANET. Many variants of AODV have been proposed in literature. Some of the multipath extension of AODV are discussed below.

M. K. Marina et al. proposed Ad hoc On-demand Multipath Distance Vector (AOMDV) [90] as an extension to the AODV protocol for computing multiple loop-free, link-disjoint paths. Loop-free is guaranteed by using a mechanism called “advertised hop count”. Link-disjointness of multiple paths is achieved by using a flooding mechanism. AOMDV improves the fault-tolerance by selecting disjoint paths. Ad hoc On-demand Distance Vector Multipath Routing (AODVM) [91] is another protocol based on AODV for finding multiple node disjoint paths. The main limitation of both these protocol is that, they do not provide any solution for routing control overhead due to packet flooding.

L. Reddeppa et al. proposed Scalable Multipath On-demand Routing protocol (SMORT) [92] which is another variant of AODV. Its main objective is to minimize the routing overhead. It uses the fail-safe multiple paths instead of node-disjoint and link-disjoint paths. A source node searches for a route by flooding a RREQ packets. After receiving RREQ packet, an intermediate node sends route reply (RREP) message to source, if it has a path to the destination. Otherwise it forwards the RREQ packet.

Many multipath routing approaches have been published in literature as variants of DSR protocol. Some of the protocols are discussed below.

S. J. Lee et al. proposed Split Multipath Routing protocol (SMR) [93], which is an on-demand multipath source routing protocol, that builds multiple routes using a request reply cycle. The SMR protocol finds an alternative route that is maximally disjoint from the source to the destination. It uses the packet flooding mechanism. The SMR uses the source routing approach, where the information of the nodes in the route is included in the RREQ packet. In SMR, the data traffic is split into multiple routes to prevent congestion.

Lei Wang et al. proposed Multipath Source Routing (MSR) [94], which is an extension of the on-demand DSR protocol and uses weighted round robin packet distribution to improve the delay and throughput. It proposed a scheme to distribute traffic among multiple routes in a network. MSR uses the same route discovery process as DSR with the exception that multiple paths can be returned, instead of only one. The limitation of the protocol is that, it does not propose any solution for routing control overhead.

Although these protocols can build multiple loop-free and link-disjoint multiple routes, all of them encounter a broadcast storm of routing packets in the process of looking for multiple disjoint routing paths. In order to ensure that the destination can select disjoint paths, all the above multipath routing protocols do not discard duplicate RREQs at intermediate nodes. Because bandwidth in MANETs is limited, reducing the routing overhead is considered to be major challenge while designing a routing protocol. None of the current multipath routing protocols have taken measures to minimize routing overhead due to flooding. Moreover, the QoS protocols discussed in the literature such as CEDAR [11], QoS-AODV [69], AQA-AODV [14] do not consider the integrated solution for congestion aware multipath routing with multi data rate support for multimedia traffic in MANETs.

3.3 Cross Layer Best Effort QoS Aware Routing Protocol

Different applications have different QoS requirements. To guarantee these QoS requirements, optimization of cross layer functionalities where higher layer functioning is improved based on the information available at the lower layer, are necessary. In this section, an adaptive and distributed congestion-aware cross layer interaction protocol called Cross Layer Best effort QoS aware routing protocol (CLBQ) which consists of rate adaptation and congestion aware optimization to improve the overall performance in terms of throughput, packet delivery, and latency for the MANETs is presented. The CLBQ discovers a less congested, high throughput routes based on the QoS metrics data-rate, packet forwarding delay, and buffer queuing delay. In this approach, each node takes the advantage of sharing the inter-layer information, such as packet forwarding delay from MAC layer and the queue length from the network layer as the congestion metric.

The proposed protocol focuses on packet forwarding delay which includes MAC delay, queuing delay, transmission delay and delay due to packet collision while discovering the congestion free route for delay sensitive application. Calculating packet forwarding delay is a crucial factor in determining the backoff factor, and becomes more significant when the network is large.The proposed cross layer architecture in CLBQ protocol is as shown in Figure 3.2.

3.3.1 Issue with Mismatched Link Data Rate Route

In a network with multi data rate transmission capability, links in a chosen path may have different data rates depending on the channel condition. Consider an example network shown in the Figure 3.3 with path S-A-B-D where S is source node and D is destination node. Assume that the links between S to D adapted to different data rates. For example, link S-A has a rate of 11 Mbps followed by a link A-B with a rate of 2 Mbps, which may result in packet accumulation at node A's buffer, creating possibilities

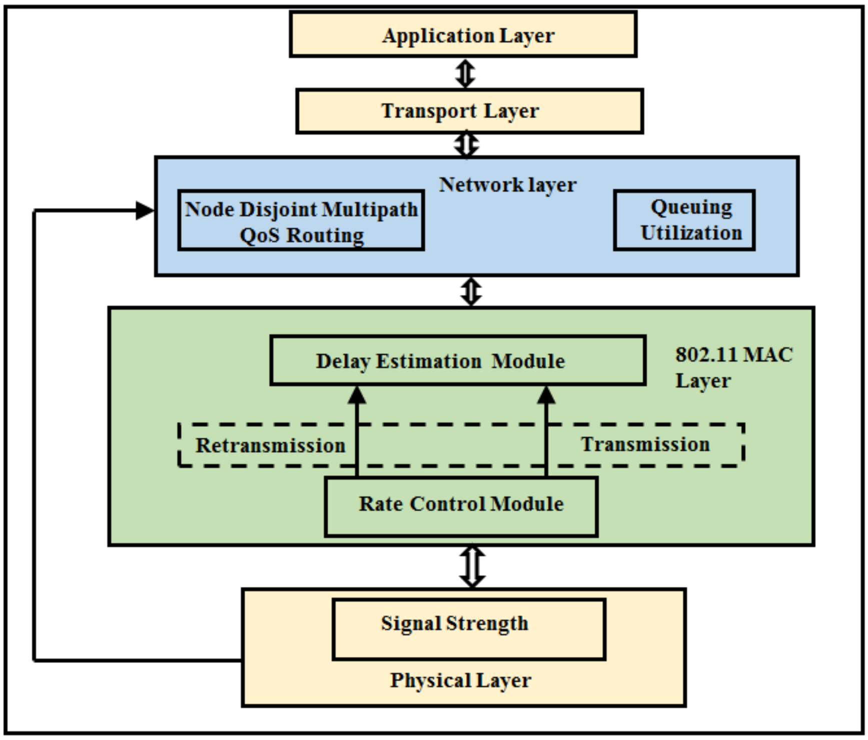


FIGURE 3.2: Proposed cross layer architecture in CLBQ

of bottleneck. Therefore, if lower data-rate links follow higher data-rate links, leading to long queuing delays at node's queue and may result in packet loss. Increased packet losses will lead to more congestion due to packet retransmissions, further degrading the network performance with extended end-to-end delay.

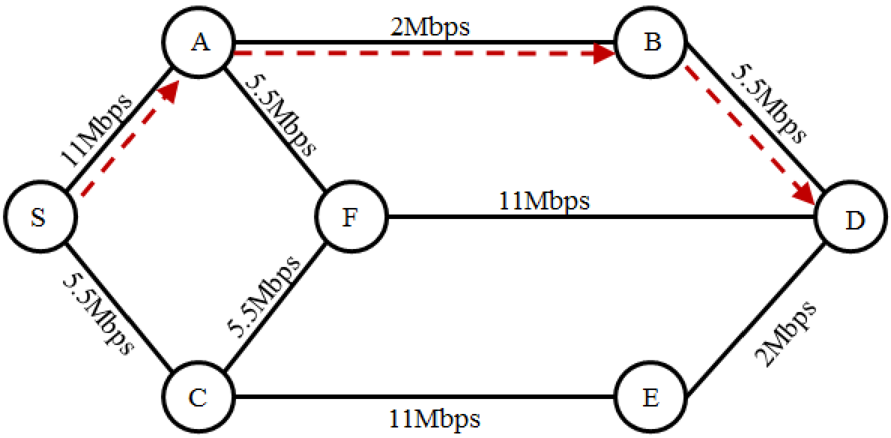


FIGURE 3.3: Link data rate mismatch in a route

3.3.2 Estimation of Queue Utilization Factor

The queue load of a node indicates the number of packets in the queue at any given time t. At the sender side, when congestion occurs, the queue load increases and when it crosses the threshold, sender should quickly send out the packets by increasing the data rate. But at the receiver side, data rate cannot be increased to support the data rate increase at the sender if the queue load of the receiver is already at threshold level. Therefore, to offer a congestion control, considering queue load only at the sender is not enough. The data rate between the sender and receiver must be balanced by knowing the queue loads at both sender and receiver to avoid high level congestion. Hence, CLBQ approach uses adaptive feedback mechanism that helps to select data rate based on queue loads of both sender and receiver.

The metric, average queue load indicates the prolonged congestion in the network due to traffic variation. The average estimated queue load of the node over time interval Δt is computed using the formula,

$$Q_{avgload}(t) = \delta \times Q_{currentload}(t) + (1 - \delta) \times Q_{avgload}(t - 1) \tag{3.1}$$

where,

$Q_{avgload}(t)$ denotes the estimated average queue load at time t,

$Q_{currentload}(t)$ denotes the current queue load,

δ can be any number selected from the range [0, 1],

The δ act as a weight parameter to regulate network congestion. If δ is selected to be too small, then it would not reflect the prolonged network congestion. If δ selected is too large, the average queue length follows the current queue length, which also degrades the congestion estimation technique, thus resulting in less effective metric. In the proposed CLBQ protocol, the congestion status of a node is based on queue utilization. Let Q_{size} be the size of the queue in a node. The average queue utilization Q_{util} for the node N_i is calculated as,

$$Q_{util} = \frac{Q_{avgload}(t)}{Q_{size}} \tag{3.2}$$

3.3.3 Congestion Evaluation

In the proposed CLBQ protocol, the congestion status is indicated by three levels i.e., Forward level, Alert level and Drop level. When the congestion status is in Forward level, i.e., $Q_{util} < Q_{thresh}$, the packets are forwarded to the next node. In Alert level, when $Q_{util} = Q_{thresh}$, the queue load balancing procedure is invoked. In case of Drop level, where $Q_{util} > Q_{thresh}$, packets are dropped due to high congestion level. Q_{thresh} is defined in the range of 80-85% of Q_{size}.

3.3.4 Estimation of Packet Forwarding Delay

The IEEE 802.11 protocol is a carrier sense multiple access with collision avoidance (CSMA/CA) MAC protocol using binary exponential backoff. The fundamental technique to access the medium is called distributed coordination function (DCF). DCF uses the RTS/CTS scheme to eliminate collision, and resolves hidden stations problem. Packet forwarding process in contention based DCF is as shown in Figure 3.4. To support asynchronous data communication, the DCF uses two kinds of frame spaces: DCF Inter Frame Spaces (DIFS) and Short Inter Frame Spaces(SIFS). Each node wanting to transmit the data packets senses the channel to be idle for at least DIFS time interval, while SIFS is used to guarantee the higher priority for control packets ahead of data packets. Hence, SIFS interval is smaller than DIFS interval.

Whenever an i^{th} mobile node N_i, wants to send packets either generated by itself or received from neighbor nodes to node N_{i+1}, it senses channel for DIFS period. If the channel remains idle for a DIFS period, then node N_i delays the transmission for the time duration during which random backoff time counter reaches zero to avoid any attempt by other nodes. The random backoff counter is decreased for each subsequent slot time if the medium remains idle. Slot time is the time unit in the backoff process. If during the backoff process, the medium becomes busy, the backoff counter is paused and resumed only when the medium becomes idle. Let $P_{idle}(t)$ be the probability that no other nodes are transmitting data and the channel is sensed idle. The probability that the channel is busy and node N_i enters backoff state is given as 1- $P_{idle}(DIFS)$

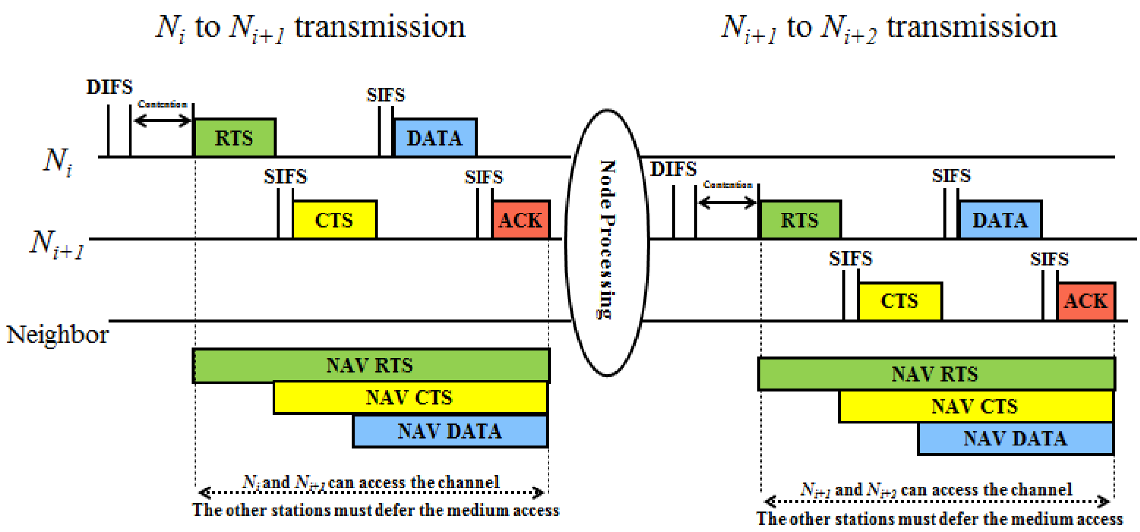


FIGURE 3.4: Forwarding process in contention based DCF

and consumes a delay of DIFS+Bf, where the Bf is the backoff time which will incur additional delay of RTS + SIFS + CTS + SIFS + DATA + SIFS + ACK before its next attempt.

When backoff counter reaches zero, the node N_i attempts to transmit the control packet RTS to the next node N_{i+1} and waits for SIFS time to receive CTS packet from the next node. The probability that node N_i senses the channel idle is given as $P_{idle}(DIFS)$. The node N_i enters again into backoff state, if it fails to receive CTS during time slot. The probability that the node N_i receives CTS is given as $P_{idle}(slot)$ and incurs delay of RTS + SIFS + CTS + SIFS time period to make sure the channel reservation is success. The probability of failing to receive CTS is given as 1- $P_{idle}(slot)$ and incurs delay of RTS +2 x SIFS. The packet forwarding delay at node N_i, which includes MAC contention and transmission delay is calculated using equation [95],

$$\begin{aligned} D^i_{delay} &= P^i_{idle}(DIFS) * (DIFS + avg_{bt} + DA(i)) \\ &\quad + (1 - P^i_{idle}(DIFS)) * (SIFS + DB(i)) + \left(\frac{L}{R}\right) \end{aligned} \tag{3.3}$$

where,

P^i_{idle}(t) is the probability that a given node N_i will not encounter any other neighbor node transmitting data during time interval t,

avg_{bt} is a mean random backoff time interval before transmission,

$DA(i)$ is the predictable delay by node incurred during the packet forwarding state,

$DB(i)$ is predictable delay incurred during backoff mechanism,

L and R are packet length and data rate respectively.

The probability P^i_{idle}(t) is given as,

$$P^i_{idle}(t) = e^{-\lambda t} \tag{3.4}$$

where,

λ is the cumulative packet arrival rate (including neighbor nodes) at node N_i.

The predictable delay $DA(i)$ is computed as,

$$\begin{aligned} DA(i) &= P^i_{idle}(slot) * (RTS + 2 * SIFS + CTS) \\ &\quad + (1 - P^i_{idle}(slot)) * (RTS + 2 * SIFS + DB(i)) \end{aligned} \tag{3.5}$$

The predictable delay $DB(i)$ is given by,

$$\begin{aligned} DB(i) = \left[\frac{1}{\left\{P^i_{idle}(DIFS) * P^i_{idle}(slot)\right\}}\right] * \Big[P^i_{idle}(DIFS) * (DIFS + avg_{bt} \\ +RTS + 2 * SIFS + P^i_{idle}(slot) * CTS) + (1 - P^i_{idle}(DIFS)) * X\Big] \end{aligned} \tag{3.6}$$

where,

X = RTS+3$*$ SIFS + CTS + L + ACK, and ACK is length of acknowledgement packet.

The proposed protocol assumes contention window as CW_{min}=32 and CW_{max}=1024. According to the binary exponential backoff algorithm in CSMA/CA protocol, the backoff delay avg_{bt} of the n^{th} re-transmission ($0 \le n \le 5$) is given by:

$$\begin{aligned} avg_{bt} &= \sum_{n=0}^{4} P^i_{idle}(slot) * (1 - P^i_{idle}(slot)^n * 2^{n-1} * CW) \\ &\quad + (1 - P^i_{idle}(slot))^5 * 2^4 * CW \end{aligned} \tag{3.7}$$

In ad hoc networks, the propagation delay is influenced by the distance between node N_i and node N_{i+1}. Since the value of the propagation delay is quite small compared to the other delays, the effect of the propagation delay is ignored.

3.3.5 Route Discovery

This section proposes a distributed algorithm that can select the data rates and determine a delay-efficient route from source to destination for admitting a new flow. A route should satisfy the delay requirement of the requesting service i.e., the end-to-end delays for data packets should be smaller than the required delay.

When a source has a data to communicate with the destination node, it checks its routing table for a valid route to the destination. If found, it sends the packet to the next hop node in the route towards the destination. However, if valid route is not found in routing table, the source initiates the route discovery process. During the route discovery process, the source creates a Route Request (RREQ) packet. Compared to traditional RREQ packet, the proposed protocol uses two additional fields $Delay_{thresh}$ and $Delay_{Remaining}$ as shown in Figure 3.5. The $Delay_{thresh}$ field contains the end-to-end threshold time requirements to be satisfied for the packets. $Delay_{Remaining}$ is the remaining time to reach the destination from the current node and is updated at every intermediate node. At the source node, both fields are initialized to the $Delay_{thresh}$ value. Source address, destination address, and sequence number of the modified RREQ packet header uniquely identifies a RREQ packet. It is important to use an appropriate data rate while broadcasting a RREQ packet. For example, if RREQ packets are forwarded using the lowest rate i.e. 2 Mbps, the discovered routing path will consist of long-range links over which higher data rate communication will not be successful. In other words, the data rate of RREQ packets essentially limits the maximum feasible data rate for the links of the discovered routing paths. Thus, the source node selects highest data rate DRi for RREQ packet, which it had used during last communication with neighbor hop. If the node forwarding the RREQ packet does not receive the response within the threshold time from neighbor nodes, it drops the rate to next lower data rate $DRi - 1$ to resend RREQ. The neighbor node on receiving

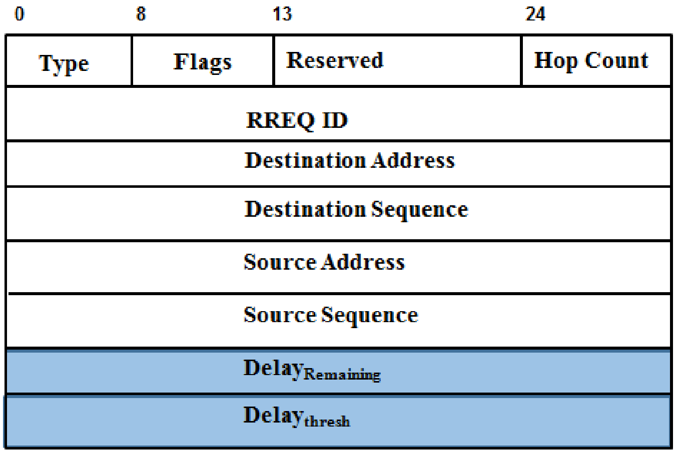


FIGURE 3.5: Modified RREQ packet format

the RREQ packet checks if it is the destination node. If not, before inserting the packet in its interface queue, it analyses the queue utilization as per Equation 3.2. If the Q_{util} of a node is in Drop level, the node drops the packet because forwarding RREQ packet to other nodes will intensify the congestion in two aspects. Firstly, transmission of this RREQ packet increases the medium contention around the congested area, leading to packet drop, which results in retransmission. Secondly, route discovery across the congested area, results in additional transmission burden. If the Q_{util} of node is in Forward, or Alert level, the node pushes the packet in interface queue to be forwarded to next neighboring nodes. Subsequently, the node computes the packet forwarding delay D^{i}_{delay} given by Equation 3.3. The node, then performs the forwarding eligibility test as per following equation,

$$Delay_{Remaining} = (Delay_{thresh} - D^{i}_{delay}) \tag{3.8}$$

If $Delay_{Remaining} \leq 0$, the intermediate node drops the packet. Otherwise node updates the $Delay_{Remaining}$ field in RREQ packet and starts the timer T_{wait} to receive multiple RREQ packets traversing through disjoint path. After timer T_{wait} expires, the intermediate node selects the RREQ packet with maximum $Delay_{Remaining}$ time. The intermediate node updates the routing table with additional information of $Delay_{thresh}$ corresponding to the route entry. When a MAC layer passes the received packet to the network layer, it informs the data rate at which it is received. It also updates the link status of its communicating neighbors and their data rates in the routing table. The intermediate nodes then broadcasts the RREQ packet to their neighbors until it reaches the destination. The route request process of CLBQ protocol is explained in Algorithm 1.

Algorithm 1 Route Request Algorithm for CLBQ

1: $N \leftarrow$ Node
2: $S \leftarrow$ Source Node
3: $D \leftarrow$ Destination Node
4: $I \leftarrow$ Intermediate Node
5: **if** $N \equiv S$ **then**
6:     **if** S has data packet to send **then**
7:         **if** S has path to D **then**
8:             Start $data\_transfer()$
9:         **else**
10:             Create RREQ Packet
11:             Select appropriate DRi
12:             Initiate $RREQ\_flooding()$
13:             **if** No RREP received within threshold time **then**
14:                 Decrease DRi to $DRi - 1$
15:                 Repeat step in 11.
16:             **end if**
17:         **end if**
18:     **end if**
19: **end if**
20: **if** $N \equiv I$ **then**
21:     **if** $Q_{util} = Drop\quad level$ **then**
22:         Drop the packet
23:     **end if**
24:     **while** $T_{wait} > 0$ **do**
25:         Hold the RREQ packet in *received_rreq_packet* table
26:         Compute $Delay_{Remaining}$ as given in Equation 3.8
27:         **if** $Delay_{Remaining} < 0$ **then**
28:             $discard\_RREQ()$
29:         **end if**
30:     **end while**
31:     Select RREQ packet with min $\{Delay_{Remaining}\}$
32:     $relay\_RREQ()$
33: **end if**
34: **if** $N \equiv D$ **then**
35:     Receive first RREQ packet start a timer T_{wait}
36:     **while** $T_{wait} > 0$ **do**
37:         Hold the RREQ packet in *received_rreq_packet* table
38:     **end while**
39: **end if**

3.3.6 Route Reply

Route reply (RREP) packets are generated for the corresponding RREQ packets received, only by the destination node and are unicast to the source node. The intermediate nodes are prohibited from generating RREP packets to route request to avoid stale route information. In the proposed protocol, traditional RREP packet is modified with additional field $Max_{queuedelay}$ as shown in Figure 3.6, which holds the maximum value of queuing delay among the intermediate nodes on the downstream route. When the destination node prepares the route reply for all RREQs received, it initializes $Max_{queuedelay}$ field to zero. Let $Q_{delay}(t)$ be the weighted moving average

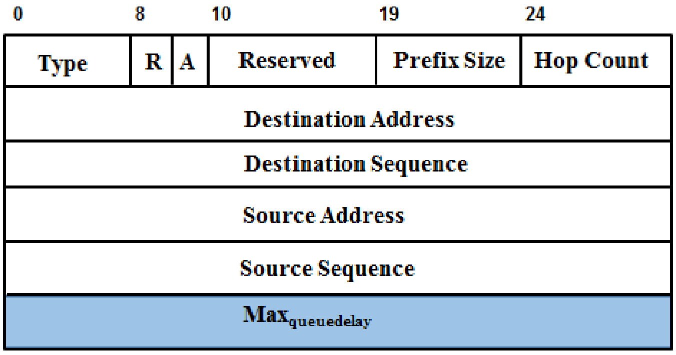


FIGURE 3.6: Modified RREP packet format

queuing delay of any node for time interval Δt and is given by,

$$Q_{delay}(t) = \eta \times Q_{delay_{current}}(t) + (1 - \eta) \times Q_{delay}(t - 1) \tag{3.9}$$

$$\eta = \frac{Q_{size} - Q_{load}}{Q_{size}} \tag{3.10}$$

where,

η acts as a weight parameter to regulate network congestion,

$Q_{delay_{current}}$(t) is the queue delay at current time t.

When the RREP packet reaches the intermediate node, it compares $Q_{delay}(t)$ with $Max_{queuedelay}$ specified in the packet. If $Q_{delay}(t)$ is greater than $Max_{queuedelay}$, then $Max_{queuedelay}$ is replaced by $Q_{delay}(t)$ value and the RREP is forwarded to downstream node of the route. The source node may receive several RREP packets from different routes. On receiving the first RREP packet, source node waits for T_{wait} duration. At the end of T_{wait} duration, the route table of the source node may have multiple route entries to the destination. The source node then chooses the primary route that has minimum $Max_{queuedelay}$ value. The route reply process of CLBQ protocol is explained in Algorithm 2.

Algorithm 2 Route Reply Algorithm of CLBQ

$Max_{queuedelay} \leftarrow 0$
if $N \equiv D$ **then**
  Prepare RREP for each RREQ packet received
  relay_RREP()
end if
if $N \equiv I$ **then**
  Compute Q_{delay} as in Equation 3.9
  if $Q_{delay}(t) > Max_{queuedelay}$ **then**
    $Max_{queuedelay} = Q_{delay}(t)$
  end if
  relay_RREP() to downstream node along the path
end if
if $N \equiv S$ **then**
  Receive first RREP and start a timer T_{wait}
  while $T_{wait} > 0$ **do**
    Hold the RREQ packet in *received_rrep_packet* table
  end while
  Choose the route with minimum $Max_{queuedelay}$
  send_data()
end if

3.3.7 Route Maintenance

Due to random movement of mobile nodes, the network topology rapidly changes thus causing frequent link interruption in MANETs. In the proposed protocol, any node which detects either a QoS violation in terms of delay or a link failure, informs the source node by sending a route error packet (RERR). During the data transmission, each node computes the packet forwarding delay according to the Equation 3.3. The intermediate node, then performs the forwarding eligibility test as per Equation 3.8. If $Delay_{Remaining} \leq 0$, the intermediate node notices the QoS violation, and it sends RERR packet to the source node. If forwarding eligibility test is successful, the intermediate node updates the $Delay_{Remaining}$ in the data packet and forwards it to the next node in the route. When the primary path is broken, the source node searches the valid secondary path, if available in route table for the same destination. If the source node find the valid secondary path, transmission of data is continued using the secondary path. To keep the secondary path active while using the primary path, the lifetime of each active secondary route is increased by a fixed amount of time. When the secondary paths are also broken, the source starts a new route discovery process. In this way, the routing overhead is minimized caused in finding and maintaining multiple paths. In a case if the source node itself move away from the neighbor nodes, it reinitiate the route discovery process to find route to the destination node. If an intermediate node in the route moves away, upstream node in the route sends a link failure notification message to each of its active upstream neighbors through RERR until it reaches the source node.

3.3.8 Data Transmission and Rate Adaptation

When a data packet is ready at a node N_i, it searches for the active primary route in the route table. If route is found in routing table with data rate DRi, node uses it to send the packet at data rate DRi. In order to improve network performance, nodes in the network should adapt to various data rates dynamically, as per the following:

1. On successful transmission of data packets for time interval T_i, the route can be assumed to be stable enough to support a higher data rate. If the current data rate is not the highest, then the transmitting node initiates an Incremental Probe (IP) packet to the next hop with increased data rate $DRi + 1$. If the node receives ACK within two attempts, transmitting node adapts to rate DRi + 1, and subsequently updates are made in the routing table.

2. During the data transmission, on not receiving two consecutive ACKs, routing layer might incorrectly conclude it as a link failure event. To overcome this, the proposed protocol runs a network event failure test to detect and differentiate a possible network events, including route disconnections, and buffer overflow, that may cause packet loss. A physical layer information is acquired to initiate corrective action according to the type of event. To detect the link failure, signal

powers of the received packets is used. Every node in the network maintains a window of signal powers of received packets i.e., received i^{th} packet P_i and corresponding signal power S_i. The node calculates the difference in signal power of the consecutive received packets $DF_{i,i-1}$. Exponential weighted moving average is used to nullify the effect due to channel fading and is given by,

$$ES_{i,i-1} = \sigma \times DF_{i,i-1} + (1 - \sigma) \times ES_{i-1,i-2} \tag{3.11}$$

where,

σ is a constant smoothing factor between 0 and 1.

The $ES_{i,i-1}$ is evaluated as follows:

(a) If $ES_{i,i-1}$ is less than zero, implies nodes are moving away from each other, and signal strength is fading away. Hence, it is a case of link failure and routing protocol initiates the route recovery procedure.

(b) If $ES_{i,i-1}$ is greater than zero implies that the nodes are moving towards each other. Hence, the receiving node is not able to process the data at the rate at which it is arriving. Therefore, the protocol concludes that the ACK loss is not due to link failure, but it is due to congestion which lead to packet loss. Hence, the transmitting node drops the data rate to next lower rate $DRi-1$, if the current rate is not the minimum. If the current rate is already minimum, then the node sends a RERR packet through downstream nodes to the source node so that the source node can initiate route discovery process.

3. If the Q_{util} value of the intermediate node N_i moves in Alert level, the possibility of packet drop may increase. This will consequently leads to re-transmission of packets, thus affecting the throughput. Hence, to avoid this, the node takes the following initiative, to reduce its Q_{util}:

   (a) The node forwards the Increment Probe (IP) packet to the next node N_{i+1} at rate DRi+1.

   (b) If node N_i receives ACK packet within the time slot, it adapts to data rate $DRi + 1$ and updates it in its routing table.

(c) If node N_i does not receive ACK within time slot, it sends a reverse Decremental Probe (DP) packet, requesting the downstream node N_{i-1} in the route to reduce the data rate to $DRi - 1$ thus ensuring less packet inflow in congested route.

   i. If the downstream node N_{i-1} is an intermediate node, it updates the data rate of forward link to $DRi-1$ in route table, and forwards the data with this new data rate.

   ii. If the downstream node N_{i-1} is a source node, on receiving the DP packet, checks in route table if alternate route is available with data rate greater than $DRi - 1$. If alternate route is available with higher data rate to the destination node, it forwards the data packet through alternate route. Otherwise source node send data packet with $DRi - 1$.

   iii. If the data rate cannot be dropped further, and no other alternate route is available, then the source node will initiate route discovery process.

3.4 Simulation Results

The performance of proposed CLBQ protocol is evaluated using simulation experiment carried out using NS2.35 [96] simulation tool. In the simulation, 802.11b network is assumed at the MAC layer, and configured with 80 mobile nodes uniformly distributed over an area of 1500 * 1500 m. The random waypoint model is used to model the node's mobility. Twenty nodes were randomly selected as constant bit rate (CBR) real-time sources, generating 512 bytes of data packets to be sent to the randomly chosen destination nodes. The MAC layer is based on IEEE 802.11 DCF. The interface queue at the MAC layer can hold maximum 50 packets. The duration of each simulation is 900 seconds and each data points are calculated as an average of 10 simulation runs. The key performance metrics evaluated are packet delivery ratio, average end-to-end delay, normalized routing load and average data rate used by the nodes to transmit data packets. The simulation parameters assumed at the MAC layer is mentioned in Table 3.1.

TABLE 3.1: Simulation parameters in the MAC and Physical layer.

Parameters	Value
MAC header	52 bytes
PHY header	28 bytes
Rate for MAC/PHY header	1Mbps
RTS	44 Bytes
CTS	38 Bytes
ACK	38 Bytes
Slot time	20 μ s
SIFS	10 μ s
DIFS	50 μ s
CW_{min} / CW_{max}	32/1024

TABLE 3.2: Scenario variations

Varying parameters	Constant parameters
Traffic load = 10-80 packets/s	Pause time = 5 s Speed = 5 m/s CBR Source = 20
Speed = 2 - 20 m/s	Pause time = 5 s Packet rate: 10 packets/s CBR Source = 20

3.4.1 Performance Evaluation

The performance of proposed CLBQ protocol is compared with CRP protocol. For the purpose of fair comparison, identical traffic and mobility scenarios are used for CLBQ, and CRP protocol. In the simulation, two protocols are compared and evaluated by varying two parameters as shown in Table 3.2. The key performance metrics used for evaluating CLBQ are packet delivery ratio (PDR), average end-to-end delay and normalized routing load (NRL).

3.4.1.1 Effect of Varying Packet Rate

In this simulation experiment, effects of traffic load on congestion in MANETs is analyzed by varying packet rate. The packet rate of CBR source node is varied from 10 packets per second to 80 packets per second. The number of source connections are fixed to 20. The mobility of all the nodes is fixed to 5 m/s. The performance of the proposed protocol is compared with CRP protocol.

Figure 3.7 shows the variation of the packet delivery ratio as a function of packet rate. It can be noticed from the figure that, when the traffic load is less, there is not much difference in the performance of the CLBQ and CRP protocols. As the traffic load scales up, route links face a higher probability of congestion, and the packet drop rate increases due to collisions or buffer overload. This results in re-transmitting the packets more than once. However, the CLBQ outperforms the CRP protocol. This is primarily due to two reasons. Firstly, queue load balancing mechanism through rate adaptation is executed at node during the congestion. And secondly, due to the availability of multiple routes which were discovered during initial route discovery phase so that in case the primary route is congested and does not satisfy QoS, it can switch to the alternate route without much delay. The CLBQ shows upto 30% in delivery ratio when the packet rate is 80 packets/s, which is significantly better than CRP.

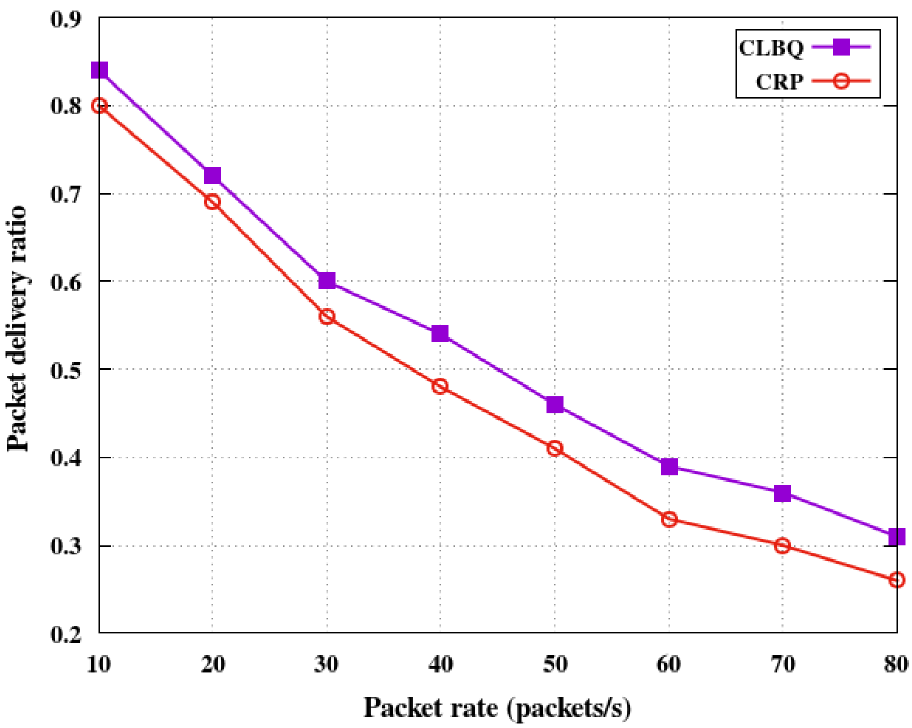


FIGURE 3.7: Packet delivery ratio v/s Packet rate

Figure 3.8 shows the plot of average end-to-end delay versus the packet rate for CLBQ and CRP. It can be noticed from figure that, in both the cases the delay increases with the increase in packet rate. It is evident from the figure that, CLBQ outperforms CRP. This is mainly because CLBQ uses the route which satisfies delay requirement for the data packet. Also, in case of congestion, it uses variable data rate to balance the traffic load on the route to reduce the packet drop and re-transmissions. In case of CRP, on congestion detection, it uses a bypass route from the previous non congested node, to divert the traffic to the destination node. But this does not guarantee the data delivery within required time period. If the primary route fails, CRP starts route discovery process all over again, which incurs additional delay. From the figure it can be noticed that, as the packet rate increases from 30 to 80 packets per second, CLBQ shows significant difference in improvement over CRP.

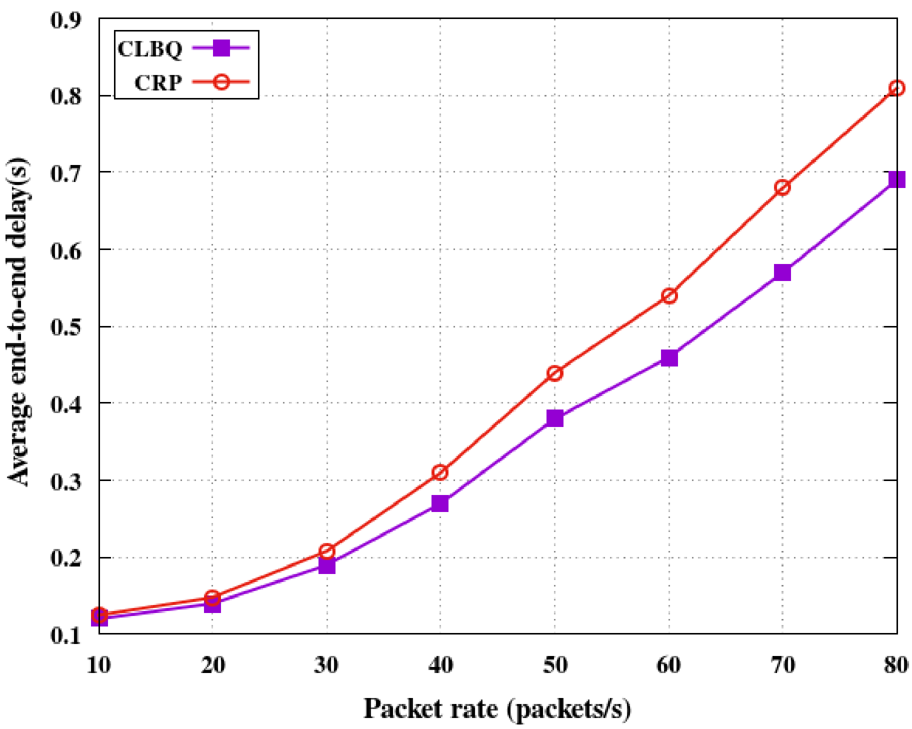


FIGURE 3.8: Average end-to-end delay v/s Packet rate

Figure 3.9 shows the normalized control overhead v/s paket rate, for CLBQ and CRP protocols. It can be noticed from the figure that, CLBQ shows an improved performance over CRP. Increase in traffic load does not affect the performance of CLBQ severely, as it predicts the congestion at the node and initiate load balancing mechanism by adapting to variable data rate. As the protocol discovers multiple node-disjoint route paths during route discovery process and supports rate adaptation, number of route rediscovery is significantly reduced. In case the congestion or link failure is predicted in the primary route, CLBQ switches the traffic through the secondary route without much of the packet drop and delay. As CRP does not have alternate backup route, rediscovery of routes are very frequent which in turn degrades the performance.

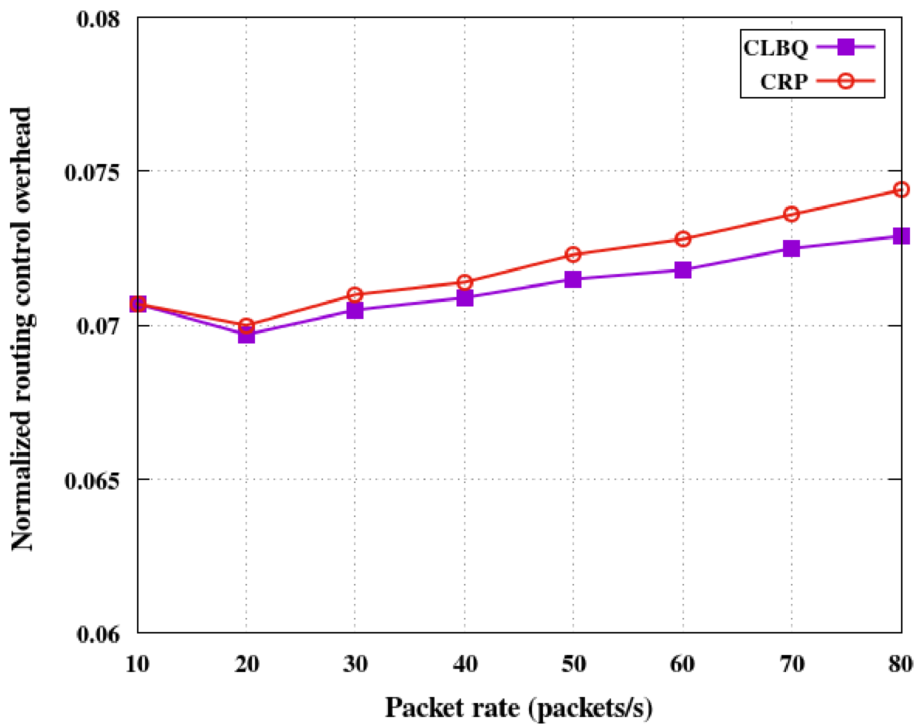


FIGURE 3.9: Normalized routing control overhead v/s Packet rate

Figure 3.10 illustrates the variation of the average rate during the entire simulation time with reference to varying packet rate. In case of CRP, the rate remains constant, equal to 2 Mbps, due to the lack of rate adaptation mechanism. Rate adaptation mechanism in CLBQ helps to vary the data rate from 2 Mbps to 11 Mbps to support varying channel condition and to minimizes the congestion. As the traffic load increases, the queue load balancing mechanism in CLBQ help to adopt suitable data rate in the event of a critical situation without much of the packets drop. Due to this node experiences more consecutive success, than consecutive failures.

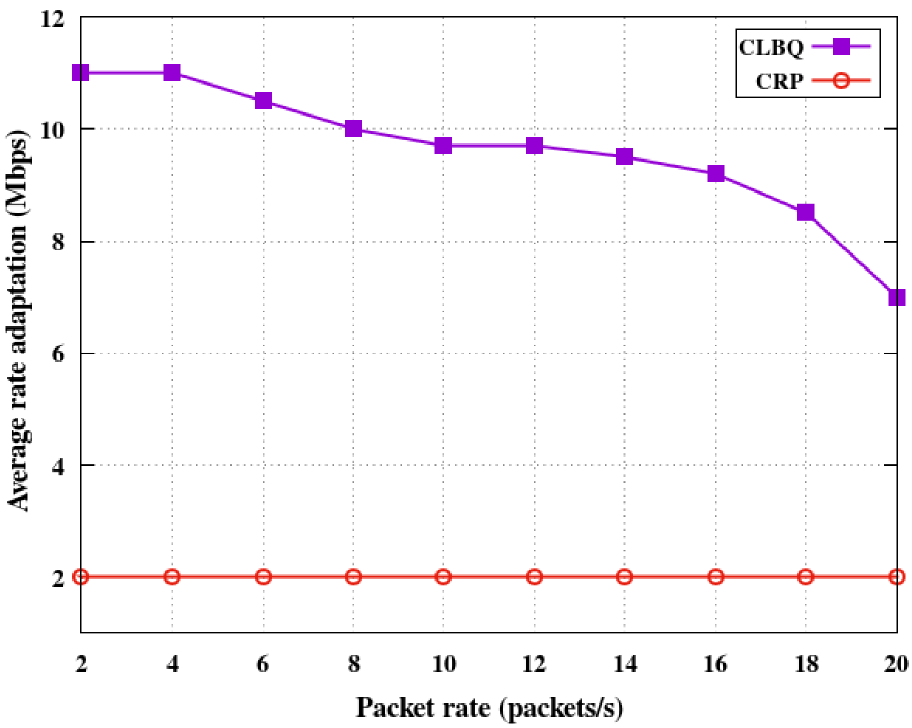


FIGURE 3.10: Average data rate adaptation v/s Packet rate

3.4.1.2 Effect of Varying Node Mobility

In the second set of experiments, the effect of mobility of nodes on the performance CLBQ is analyzed. The number of source connections are fixed to 20. The packet sending rate is set to 10 packets /s. Experiments are conducted by setting nodes speed to 2, 4, 6, 8, 10, 12, 14, 16, 18, and 20 m/s.

Figure 3.11 demonstrates the variation of the packet delivery ratio as a function of node's mobility. When the node's mobility is minimal, both CLBQ and CRP performs almost similar, as there is less probability of link failure. This results in reduction of buffering time in queue, the probability of congestion and collisions. When nodes are highly mobile, link failure probability increases and hence, the packet delivery ratio decreases. The CLBQ shows better delivery ratio compared to CRP when node's speed increases and this is mainly due to node-disjointness in multiple routes discovered during route discovery phase. When an active route is about to break due to mobility of nodes, the intermediate node predicts the link failure and notifies the source node. The source node at once invalidates the current route in its route table and selects another valid node-disjoint route from its routing table to continue with the communication between source and destination without any interrupt.

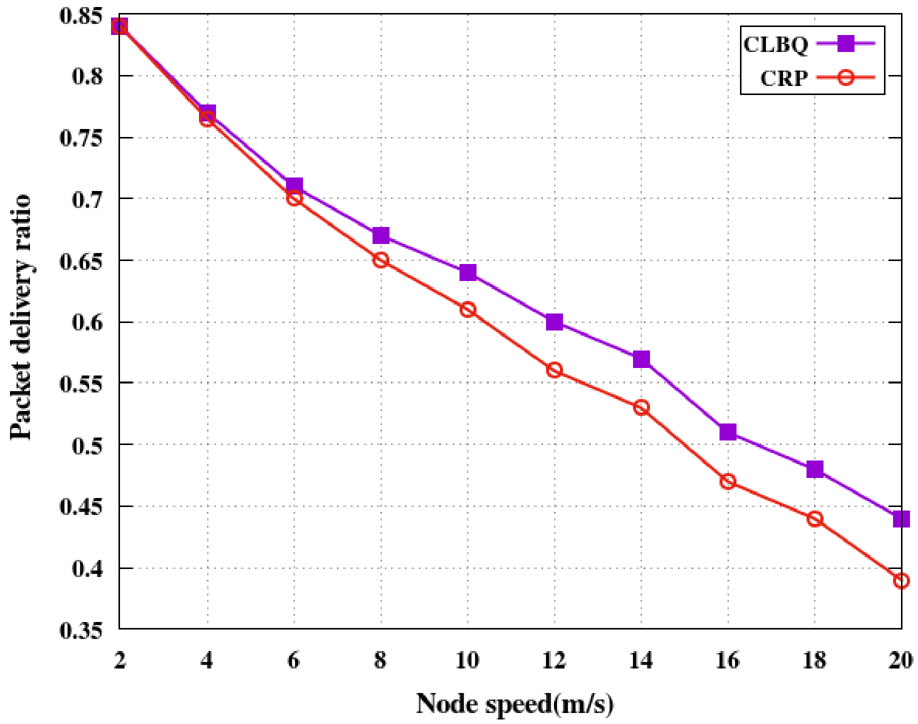


FIGURE 3.11: Packet delivery ratio v/s Node speed

The variation of the average end-to-end delay as a function of nodes mobility is as shown in Figure 3.12. From the figure, it can be observed that both the delay curves increases with the increase in speed of nodes. This is because of high mobility of nodes results in an increased probability of link failure that in turn causes an increase in the number of routing rediscovery processes. This makes data packets wait for more time in its queue until a new routing path is found. When mobility of nodes is less, both the protocols gives the same performance. The proposed CLBQ shows substantial improvement in the average delay compared to CRP protocol. This is because CLBQ guarantees the minimum delay route for delivery of packets and switching to alternate route in case of link failure. Whereas CRP does not guarantee the packet delivery within the required time limit to the destination node. In CRP, bypass mechanism does not work efficiently, when primary route suffers multiple link failure.

Figure 3.13 shows the normalized control overhead with increasing nodes mobility. Results clearly shows that CLBQ performs better than CRP. This is mainly because

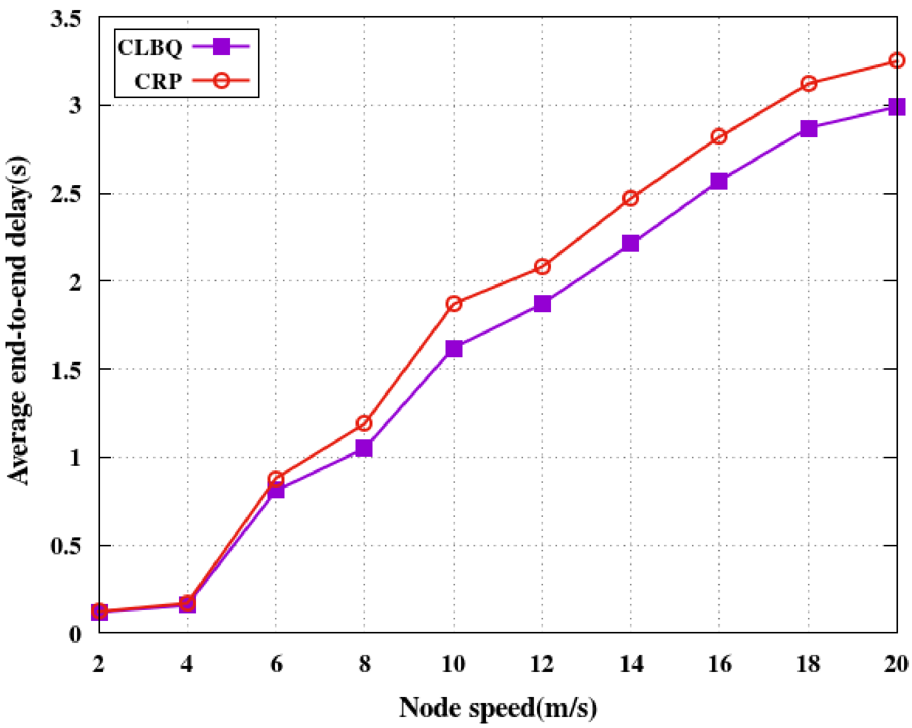


FIGURE 3.12: Average end-to-end delay v/s Node speed

CLBQ uses higher data rates for packet transmission whenever possible due to which more data packets are transmitted compared to control packets. CLBQ shows better improvement in the routing overhead proving it to be a lightweight protocol.

Figure 3.14 illustrates the variation of the average rate nodes adapted during the entire simulation time with varying mobility. It can be notice that, the rates adopted by nodes in case of CLBQ are better than CRP due to the rate adaptation mechanism. The rate varies from 2 Mbps to 11 Mbps. However, in case of CRP, the rate remains constant, equal to 2 Mbps, due to the lack of rate adaptation mechanisms. In addition, these results show that the nodes have more consecutive successes than consecutive failures, so their rates can rapidly get stabilized. As the traffic load increases, the queue load balancing mechanism in CLBQ help to adopt suitable data rate in the event of a critical situation without much of the packets drop.

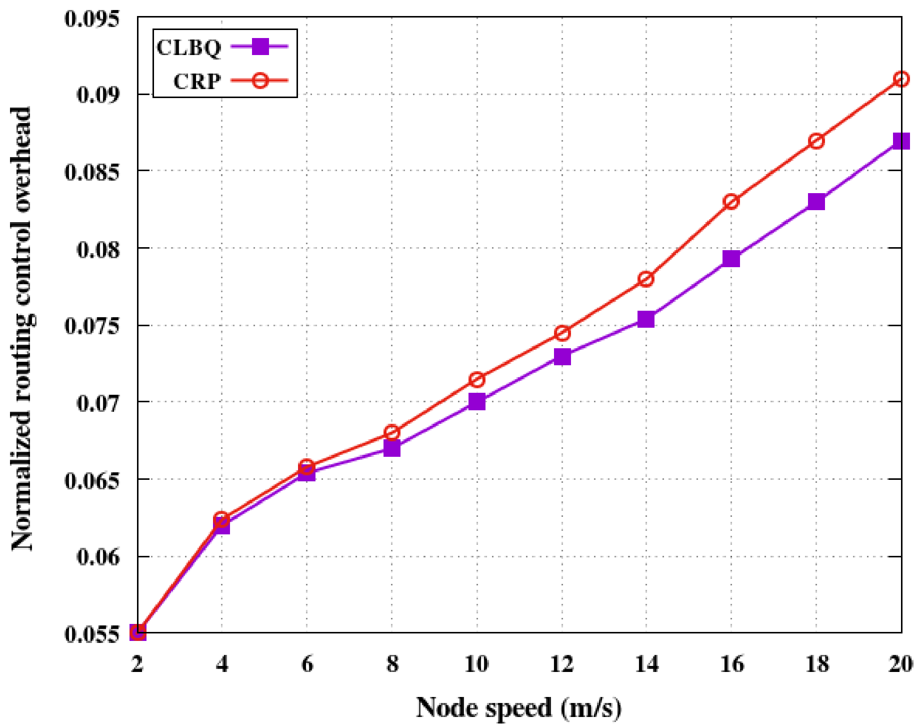


FIGURE 3.13: Normalized routing control overhead v/s Node speed

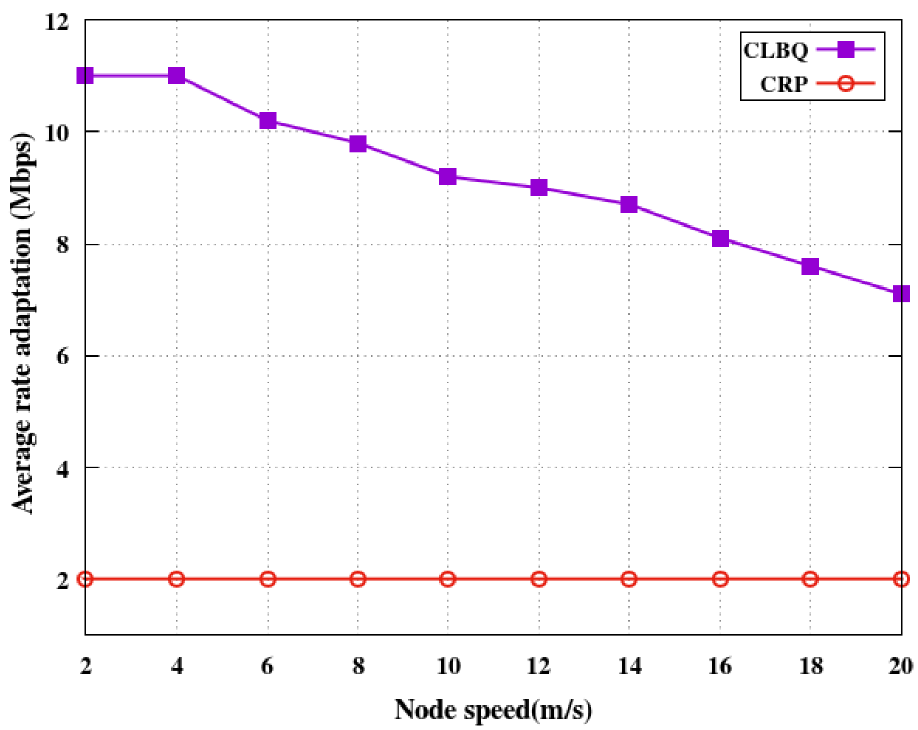


FIGURE 3.14: Average data rate adaptation v/s Node speed

3.5 Conclusion

In this chapter, a Cross Layer Best effort QoS aware routing protocol is presented. The proposed protocol is an adaptive, distributed, reliable and congestion aware based on multi-rate adaptation. It utilizes wireless resources efficiently and provides better QoS support for delay-sensitive communications. The proposed protocol discovers multiple node-disjoint routes that satisfy QoS requirements. The CLBQ is a congestion-adaptive routing protocol, where it not only avoids the congested node while constructing a route, but also deals with it reactively. A detailed simulation study has been carried out to evaluate the performance of the proposed CLBQ protocol with the CRP protocol. The results show that the CLBQ protocol significantly reduces both the packet drop ratio and the end-to-end delay without much impact on control overhead. CLBQ protocol provides better throughput than CRP due to its variable data rate adaptation technique. During the route discovery process, CLBQ protocol finds multiple congestion aware routes and in case of link failure, or congestion detected in the primary route, it switches to alternate route in minimal time, thereby reducing the delay due to route re-discovery, This in turn helps in guaranteeing the minimal delay in packet delivery.

Chapter 4

Adaptive Best Effort Traffic Scheduler for IEEE 802.11e EDCA

4.1 Introduction

The wide deployment of MANET technology interconnected with the Internet and a gradual increase in UDP based multimedia applications has brought about a demand for improving the QoS in MANETs. The IEEE 802.11 standard aims to provide fairness in terms of channel contention to all users with a CSMA/CA mechanism. The IEEE 802.11 uses DCF, which is based on best effort model and does not provide any QoS support for time critical real time applications.

Many researchers have worked on improving IEEE 802.11 features and proposed several QoS enhancement schemes at MAC layer [97] and Network layer [98] to support the growing demand for multimedia applications. Some of the researchers have proposed techniques to provide service differentiation in wireless LANs [99][100]. IEEE 802.11e is an enhancement to existing IEEE 802.11 which serves as an add-on to IEEE 802.11 with QoS support [101]. The Enhanced Distributed Coordination Function (EDCF) in IEEE 802.11e is an extension to classical DCF, which provides differentiated channel access to frames of different priorities. Enhanced Distributed Channel Access (EDCA) mechanism of the IEEE 802.11e works on the principle of traffic prioritization into four access categories, namely background, best effort, voice and video, and supports 8 user priority values. The channel access probability of delay-sensitive multimedia applications are increased by tuning the contention windows (CWs), Transmission Opportunity (TXOP) and Arbitration Inter Frame Spaces (AIFS) values of respective access category queue. These differentiated settings speed up the backoff expiration time, thereby creating a higher probability to win the contention [24].

However, today's Internet traffic is still dominated by TCP-based applications such as email, FTP, HTTP. TCP is one of the popular protocols designed for wired networks, which provides best effort, reliable end-to-end delivery of data over unreliable networks [102]. On the contrary, TCP is not well-suited for MANETs due to its inability to adequately accommodate frequent link changes and delays that are common due to factors such as link layer error, wireless bandwidth fluctuation or blocked by high priority traffic [103][104].

IEEE 802.11e has not adequately addressed the issue of handling best effort TCP traffic containing TCP data and TCP ACK packets in the presence of high UDP traffic. In the process of providing QoS enhancements to delay sensitive multimedia traffic, best effort traffic is left to starve for a longer time. Due to this, there may be delays while receiving TCP ACK packets for the successfully transmitted TCP data packets which may lead to false interpretation of packet loss and trigger the furious retransmission of data packets. This in turn results in bandwidth wastage and throughput degradation. Although it is necessary to use effective QoS protocols to provide QoS guarantees to real-time traffic, it is also important to provide QoS for TCP traffic so that all users can get reasonable services. To enhance the best effort traffic performance in 802.11e with high-quality of service to maximize the system throughput, this chapter proposes a protocol called, Adaptive Best Effort Traffic scheduler for EDCA (ABET-EDCA). ABET-EDCA scheme prioritizes the TCP traffic during congestion by adaptively tuning the MAC parameters based on traffic load conditions.

4.2 A Literature Review

Wireless networking based on the IEEE 802.11 standard has gained the most popularity among the different standards over the past decade. The IEEE 802.11e, which is an extension to standard IEEE 802.11 CSMA/CA contention mechanism with QoS functionalities, is a standard adopted for WLAN and now extended to MANETs. The IEEE 802.11e offers QoS guarantee for different traffic flow by assigning priorities to individual packets and supports four independent transmission queues for access categories (AC) at each node. The service differentiation is achieved by configuring four MAC parameters CW_{min}, CW_{max}, AIFS and TXOP for each AC. The various techniques proposed by several researchers to address the effects of 802.11e QoS techniques on TCP flows in WLAN and MANET are discussed below. The research work on IEEE 802.11e which focuses on the problem of asymmetry in the forward and reverse path of best effort TCP traffic and the throughput performance of TCP compared to that of UDP is also presented.

Peng Q. et al. [25] proposed a method to improve the TCP fairness by using cross layer architecture. This research work is based on the fact that when a node receives an ACK, it confirms the receipt of the TCP data at the receiver. There is a TCP ACK agent at the AP that sends TCP ACKs to the servers on the Internet on behalf of the wireless nodes. Although this scheme shows performance improvement, it needs additional hardware like buffers to store the TCP data packets in the TCP client until the TCP ACK is received and it also needs a high speed processor at the AP to generate local TCP ACK packets when a successful ACK from the server is received. This leads to additional overhead and processing complexity. The performance of this system on a multi-hop network has not been explored.

Leith et al. [26] has discussed about achieving the symmetry between the forward and reverse TCP traffic. In this paper, creating a separate higher priority queue in the AP for TCP ACK packets has been suggested. This mechanism helps TCP ACK packets to have unrestricted access to the wireless channel. This is definitely a good alternative, considering the single-hop scenario with TCP traffic. But in real life scenario, the communication is a mix of TCP traffic and real time traffic. Hence, the main disadvantage of giving highest priority to the TCP ACKs is that, other high priority data traffic like real time audio or video have to compete with ACK packets. This in turn may result in poor performance of real time applications.

For a wireless network with AP, Stephane Lohier et al. [105] proposed a MAC-layer Loss Differentiation Algorithm (LDA) based on an dynamic adaptation of the retry limit parameter in order to restrict MAC retransmissions depending on the quality of the wireless channel. This helps to reduce performance degradation due to spurious triggering of TCP congestion control. The authors proved that the lack of interactions between MAC and TCP recovery levels in the event of 802.11 signal losses, leads to performance degradations.

Park et al. in [106] proposed a scheme for best effort traffic, which uses two high priority queues and accommodates TCP ACK packets whenever one of the queue is empty. However, the scheme is ineffective as it allows only one best effort frame at a time to provide QoS for the application. Moreover, high priority queues are mostly fully occupied. This may result in ACK packets competing with high priority packets, thus affecting the multimedia throughput performance.

All the above schemes listed were designed for WLANs. Some of the notable work done in ad hoc networks is listed below.

Altman et al. [27] proposed an adaptive scheme for delaying TCP acknowledgments known as DelAck, to improve TCP performance in multi-hop wireless ad hoc networks. DelAck is a receiver-side process for minimizing channel contention among data packets and ACKs of the same TCP connection by decreasing the number of TCP ACKs transmitted by the sink. The design allows a receiver to generate a single ACK for set of 'N' TCP packets. However, there are two problems with this approach. First, the value of N dynamically changed based on the size of the TCP congestion window which in turn depends on the available bandwidth of the channel. The second issue is how to relate the value of N to the time dependent congestion. Also, there is a sudden burst of TCP packets injected into the network every time a delayed ACK is received by a source, which can lead to temporary network congestion and congestion further escalates due to undelivered packets.

For MANETs, Liu et al. [107] proposed a solution to avoid the starvation of low priority traffic when the network is under excessive high priority traffic load. Their method attempts to avoid the conventional scheduling by piggybacking low priority traffic on higher priority messages if their next hop in the routing table is the same. But this is only possible if the channel is in a free space state, i.e., there is a residual bandwidth that can be used by low priority traffic with the same route as the high priority traffic. This method is a cross layer design in which the MAC layer has access to the routing information. This approach is also completely independent of the service differentiation scheme and the routing algorithm used.

C. Mbarushimana et al. proposed cross layer based E-TCP [108] protocol for MANETs, which uses TCP bidirectionality mechanism to avoid TCP traffic starvation and spurious retransmissions. E-TCP prioritizes the TCP acknowledgement packets by adjusting TCP retransmission timer based on the medium contention. E-TCP also provides the solution to the TCP spurious retransmissions, by a mechanism that modifies TCP retransmission timer dynamically taking into consideration additional delays due to increased medium contention. In E-TCP, the main disadvantage of giving highest priority to the TCP ACKs is that, other high priority data traffic like real time audio or video have to compete with the ACK packets. Also, only ACK packets are prioritized and not the TCP data packets, which may not be ideal for high TCP traffic.

All the above schemes discussed, has been proposed for improving the TCP performance over lossy communication links. There are still major issues that are yet to be explored and need an analysis in detail to support QoS for best effort traffic in the presence of high priority real time traffic. These issues are described below.

- The issue of delays due to the presence of high priority traffic is yet to be addressed. This problem is intensified by the introduction of traffic prioritization in IEEE 802.11e and the increase of multimedia traffic over the past few years.
- There are no explicit signaling mechanisms in the network to tell TCP source nodes, how fast to send, how much to send, or when to slow down a transmission. A source node is responsible for controlling these parameters from

implicit knowledge, it obtains from the network. Asymmetry in the forward and reverse path packet transmission is a known source of poor TCP performance in MANETs.

- In MANETs, wireless network link often experience delay spikes in TCP traffic. Delay spikes occur due to link layer error recovery, route failure, wireless bandwidth fluctuation, and blocking by high priority traffic. When these delay spikes exceeds the retransmission timeout value, TCP assumes that the outstanding segments have been lost and retransmits them accordingly. Since the data packets were delayed and not lost, this retransmission is unnecessary and the timeout is spurious. Following a timeout, TCP will exponentially increase the retransmission timeout (RTO) value. For subsequent transmissions, this leads to a long waiting time period, and for non congestion losses this result in a waste of the limited available bandwidth.

- The IEEE 802.11e is proposed with the primary objective to provide QoS for real time applications. Hence, assigning the priority to the TCP data to enhance the performance TCP applications may hinder the performance of real time delay sensitive application.

- Earlier studies have shown that the default values of the contention parameters are only good for scenarios with few higher priority ACs and under moderate traffic load [109]. In EDCA, there is no standard procedure to adjust the contention parameters under varying network conditions to achieve a certain capacity ratio between the ACs and at the same time achieve high channel utilization.

- Compared to traditional networks, MANETs, possess many unique characteristics. The node in the network can drop a packet due to several networking events other than buffer overflow such as link contention error, route changes due to topological changes or channel error. Unfortunately, the current layered network architecture makes it impossible to pass these event information from lower layers to other higher layers to take corrective measures. As a result, if a packet is lost due to reasons other than buffer overflow, TCP adversely invokes its congestion control procedure.

To address these issues, an Adaptive Best Effort Traffic scheduler for EDCA (ABET-EDCA) protocol is presented in this chapter.

4.3 Proposed ABET-EDCA Protocol

ABET-EDCA is a cross layer based packet scheduling scheme, which dynamically adapt the MAC parameters based on the queuing delay of TCP packets at the MAC layer and reassigns its priorities. The scheme is further improved by implementing the dynamic TXOP limit value, which is computed at runtime. The TCP packets are prioritized in the queue and are scheduled within their desired delay level. The proposed ABET-EDCA can detect the network event occurring at the lower layers of TCP protocol stack, like link failure and buffer overflow. This information is exploited by the network and transport layer to correctly trigger appropriate actions to enhance the TCP performance. ABET-EDCA improves the performance of low priority TCP traffic with QoS provisioning, while having minimal negative impact on high priority delay sensitive UDP-based real time traffic.

The MAC protocol IEEE 802.11e is used in wireless LAN to support growing demand for multimedia traffic with QoS capability. The direct application of IEEE 802.11e in MANETs may pose performance issues due to following reasons. Firstly, MANET is an infrastructure-less network and uses multihop communication. Secondly, the routing path usually affected by rapid and unexpected network topological changes due to node mobility as well as the addition and deletion of nodes. Hence, the proposed ABET-EDCA protocol combines the functionalities of IEEE 802.11 DCF and IEEE 802.11e EDCA at the MAC layer.

4.3.1 Prioritizing TCP ACKs

In IEEE 802.11e, EDCF scheme differentiates the traffic and maps them into separate access categories AC(i) (i = 0, 1, 2, 3) depending upon their priority. The MAC channel access parameters such as contention window range [CW_{max}, CW_{min}] and AIFS are set

TABLE 4.1: IEEE 802.11e EDCA MAC parameters

AC	CW_{min}	CW_{max}	**AIFSN**
AC(0) (Voice)	7	15	2
AC(1) (Video)	15	31	2
AC(2) (best effort)	31	1023	3
AC(3) (Background)	31	1023	7

to different values for different AC(i) and is shown in Table 4.1. In IEEE 802.11e, when a node receives the packets, either generated by its own application layer or received from its neighbor nodes, the MAC layer places the packet in corresponding AC queue. The AC queues are implemented based on first in, first out "FIFO" mechanism. There is no provision to prioritize packets within the AC according to their importance or demand. As a result, when an AC wins the channel access, the packets arrived first gets the unrestricted access to the channel, and the important packets that arrived later in the queue had to wait for a longer time. In IEEE 802.11e, all TCP data and ACK packets are mapped to AC(1) queue at the MAC layer. The ACK packets are treated like any other ordinary packet, and dequeued based on "FIFO" mechanism.

The IEEE 802.11e, with UDP real time traffic flow performs well when compared to TCP traffic in terms of throughput. Poor TCP performance is mainly because the TCP ACKs are not prioritized, leading to forward and reverse path asymmetry[110]. Delay in transmitting TCP ACKs in the presence of high priority traffic may lead to retransmissions of TCP packets. Typically, TCP compute the RTO based on the Round Trip Time (RTT) values. When an ACK packet waiting delay exceeds the retransmission timeout value, the sender node assumes that the outstanding packet has been lost and invokes congestion control mechanisms and a retransmission is scheduled. This is unnecessary, since retransmitting a packet that was successfully transmitted is a waste of the already limited bandwidth.

To improve the TCP throughput, the ABET-EDCA proposes a priority queue scheduling mechanism to prioritize the TCP ACK packets. In Priority scheduling scheme, the AC(1) queue is implemented using heap data structure [111], wherein the ACK packets are stored and sent out with highest priority. This ensures the TCP packets that managed to get successfully transmitted are acknowledged instantaneously. This allows the transport layer to control the rate at which TCP ACKs are delivered to the wireless channel instead of the MAC layer. The Figure 4.1 shows the proposed QoS scheduler for ABET-EDCA.

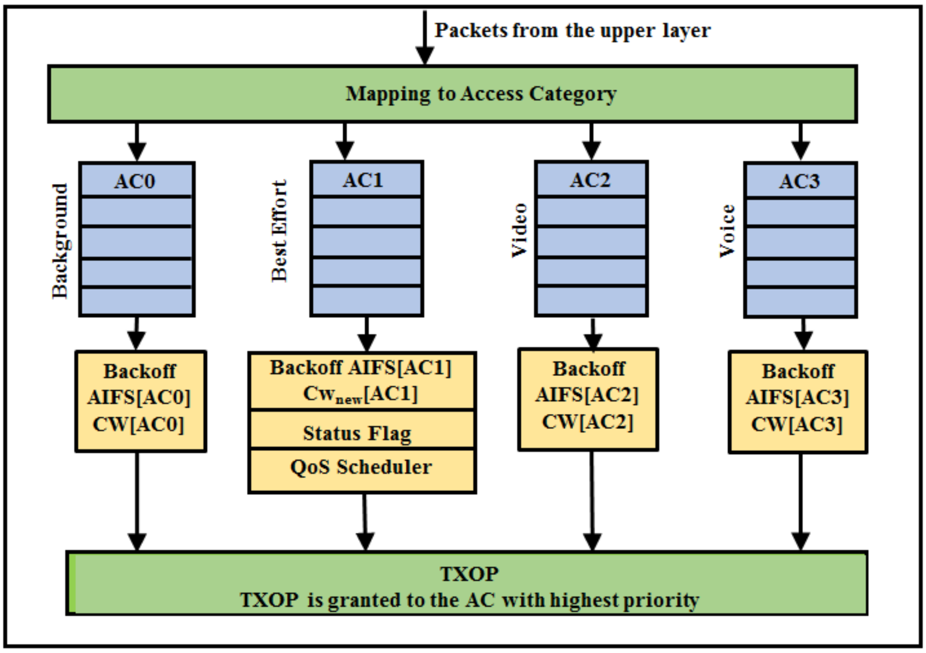


FIGURE 4.1: QoS scheduler for ABET-EDCA

4.3.2 Best Effort Traffic Congestion Detection

To maintain the semantics of the TCP protocol, a congestion control technique is implemented that requires a measurement obtained at the AC(1) queue. Queuing delay threshold $Q_{thresh}(DATA)$ and $Q_{thresh}(ACK)$ is set for the TCP data and ACK packets respectively as a benchmark to indicate the congestion. These thresholds give the estimate of how long the packets can be kept on hold in a buffer without degrading the performance. The QoS scheduler maintains the timestamp, $Last_{Data}(t)$ and $Last_{ACK}(t)$ of the last TCP data and ACK packet transmitted from AC(1) queue. The QoS scheduler computes the queuing delay $Queuing_{delay}(Data)$ and $Queuing_{delay}(ACK)$ for the TCP data and ACK packets respectively in AC(1) queue as follows:

$$\begin{aligned} Queuing_{delay}(Data) &= Node_{current}(t) - Last_{Data}(t) \\ Queuing_{delay}(ACK) &= Node_{current}(t) - Last_{ACK}(t) \end{aligned} \tag{4.1}$$

where,

$Node_{current}(t)$ is the current time at the node.

If the $Queuing_{delay}(Data)$ or $Queuing_{delay}(ACK)$ has exceeded the $Q_{thresh}(DATA)$ or $Q_{thresh}(ACK)$, the QoS scheduler raises the FLAG to 1 indicating the congestion at the AC(1) queue. When node wins the channel access, QoS scheduler checks the AC(1) queue FLAG status. If the FLAG is not raised, any AC queues which initiated the contention backoff process and whose CW counter reaches 0 first, wins the channel access, and transmits the packets present in the respective AC queue onto the channel. If the FLAG is raised for AC(1), then the QoS scheduler checks current average queue utilization $Q_{AC(i)util}$ for AC(2) and AC(3), which is computed as,

$$Q_{AC(i)util} = \frac{Q_{avgloadAC(i)}(t)}{Q_{size}} \tag{4.2}$$

where,

$Q_{avgloadAC(i)}$(t) is the average queue load at each AC(i) for $1 < i \leq 3$,

Q_{size} be the size of the AC queue in a node

The average load of the AC $Q_{avgloadAC(i)}$(t) is computed as,

$$Q_{avgloadAC(i)}(t) = (\delta) * Q_{loadAC(i)}(t) + (1 - \delta) * Q_{avgloadAC(i)}(t - 1) \tag{4.3}$$

where,

$Q_{loadAC(i)}(t)$ denotes the current queue load,

δ is a constant in the range [0, 1]

If the queue utilization factor of AC(2) or AC(3) is above the Q_{Thresh} level, then the AC which initiated the contention backoff process first will win the channel access and forwards the data from the respective queue. The Q_{Thresh} is the variable limit defined within the range of 70-75% of Q_{size}. If queue utilization of AC(2) or AC(3) is less than Q_{Thresh}, then the QoS scheduler dynamically assigns the highest priority to AC(1) by tuning the MAC layer parameters such as contention window, and TXOP to overcome congestion at AC(1).

4.3.3 Dynamic Adaptation of MAC Parameters

EDCF achieves considerable improvement over DCF in providing QoS guarantees for delay-sensitive applications by using different MAC parameter values for different ACs. The CW range and AIFS parameters are set to low values for high priority traffic in AC(3) and AC(2), and high values for low priority traffic in AC(1) and AC(0), such that the high priority traffic gets most of the bandwidth and have to wait for less time for getting a chance to transmit. As a consequence of this, low priority traffic tends to get starved in IEEE 802.11e wireless networks [112][113]. When there is an internal collision, the internal scheduler gives access to the higher priority queue.

The EDCF uses the binary exponential backoff algorithm in which for each access category, the predefined values of the initial backoff window size $CW_{min}[i]$, and the maximum backoff window size $CW_{max}[i]$ are used. Whenever there is a transmission failure, the Contention Window size for a particular AC(i) is doubled till it reaches the maximum value, CW_{max} [i], i.e.,

$$CW[i] = min\{2 * CW[i], CW_{max}[i]\} \tag{4.4}$$

On successful transmission, the CW[i] is reset to its minimum value CW_{min}[i], i.e.,

$$CW[i] = CW_{min}[i] \tag{4.5}$$

The EDCF assumes that once a successful transmission occurs the factors causing failure of transmission no longer exist. This may not be the case always and it may result in a higher rate of collisions, further degrading the performance. Also, in the current IEEE 802.11e standard, there is no way of prioritizing the packets corresponding to the same flow inside a queue.

4.3.3.1 Dynamic Adaptation of Contention Window

This section proposes a method to dynamically prioritize the lower priority TCP packets to gain more transmission opportunities. The lower priority TCP packets are prioritized according to the traffic load in the network. The main motivation behind this work is to ensure that lower priority TCP traffic is not starved.

In the proposed scheme, the incrementing of the contention window in case of a transmission failure as well as its resetting in case of collision is done in a non-uniform manner within the Contention Window range. Whenever there is a FLAG raised by a QoS scheduler for AC(1) and there occurs the event of unsuccessful transmission for TCP packets, the protocol checks current $CW[i]$ value. If it is less than twice that of its $CW_{min}[i]$, its CW is increased at a faster rate by multiplying by a factor of 1.5, as proposed in [114]. This will ensure that, AC(1) queue gets a faster access to channel without much of the packet drop. The algorithm for setting the contention window is listed in Algorithm 3.

Algorithm 3 Contention Window increasing Algorithm

```
if i ≥ 2 then
    CW[i]=min{CW[i]+1, CW_max[i]}
else
    if i = 1 then
        if CW[i] < 2*CWmin then
            CW[i]=min{CW[i]*1.5, 2*CW_min[i]}
        end if
    end if
end if
```

Whenever there is a successful transmission of TCP packets in AC(1), in the proposed protocol, $CW[i]$ value is not reset to $CW_{min}[i]$ as in case of traditional EDCA. Instead, AC(1) contention window is linearly decremented, if $CW[i]$ is less than twice that of $CW_{min}[i]$ otherwise it is decremented by a factor of 0.5, as a slow decrement scheme. This will ensure that during high TCP traffic flow, the AC(1) gets channel access in minimum time to accommodate TCP transmission faster. For video or voice traffic the Contention window is linearly decremented. The algorithm for resetting the contention window is listed in Algorithm 4.

Algorithm 4 Contention Window Reseting Algorithm

```
if i ≥ 2 then
    CW[i]=max{CW[i]-1,CW_min[i]}
else
    if i = 1 then
        if CW[i] < 2*CWmin[i] then
            CW[i]=max{CW[i]-1, CW_min[i]}
        else
            CW[i]=max{0.5*CW[i], 2*CW_min[i]}
        end if
    end if
end if
```

4.3.3.2 Dynamic Adaptation of TXOP Limit

Most of the existing work that addressed QoS issues in IEEE 802.11e, guaranteed a strict differentiation between ACs. Service differentiation is guaranteed between the ACs by using different values for the AIFS or CW parameters. Hence, the different flows belonging to the same AC have the same probability of transmission. This may be inefficient if different flows belonging to the same AC have different QoS requirements. Therefore, it is necessary to differentiate and prioritize different flows within the same AC. To address this issue, the proposed protocol controls the channel access through the use of the TXOP limit parameter.

When an AC wins the contention for the channel, it transmits as many frames as available in its queue successively provided that the duration of transmission does not exceed the specified TXOP limit [6]. This process is popularly known as TXOP burst, which eliminates the additional contention overhead and control packets transmission. The TXOP limit during which the node can send more than one data frame depends on the AC and the physical rate defined in 802.11 a/b/g. For example, for background and video priority, the TXOP limit ranges from 0.2 ms to 3 ms in an 802.11a/g network and 1.2 ms to 6 ms in an 802.11b network. The default TXOP scheme assigns a fixed TXOP limit to all the flows of the same service class. The IEEE 802.11e provides the flexibility to tune the TXOP, but it does not specify hard protocol to accomplish it. A salient feature of the TXOP operation is that, if a large TXOP is assigned and there are not enough packets to be transmitted, the TXOP period is ended immediately to avoid wasting bandwidth.

The proposed ABET-EDCA introduces an adaptive TXOP scheme. The scheme maintains the average queuing time of TCP packets buffered in the AC(1) queue. Each time when the status FLAG is raised for AC(1) queue, QoS scheduler dynamically computes the TXOP limit value at runtime based on average queuing time and the status of the AC(1) queue. Dynamic adjustment of TXOP helps to forward the stranded TCP packets which has exceeded the queuing threshold time, in bursts. This reduces the TCP traffic packet delay by reducing the number of channel accesses required to transmit long waiting TCP packets and maximizes the throughput. The extended TXOP is computed as,

$$TXOP[i] = originalTXOP[i] + supplementaryTXOP[i] \tag{4.6}$$

$$originalTXOP[i] = N_i * \left[\frac{M_i}{R_i} + 2 * SIFS + ACK\right] \tag{4.7}$$

$$N_i = \frac{P_i * SI}{M_i} \tag{4.8}$$

where,

P_i represents the mean data rate,

M_i the packet size,

R_i the channel data rate

and SI is the service interval.

$$supplementaryTXOP[i] = \frac{Queue[i] * L_i}{DR_i} \tag{4.9}$$

where,

L_i = MSDU size in bits,

DR_i = minimum physical transmission rate of flow in bits per sec,

and Queue[i] = number of stranded packets in the queue of the flow *i*.

When a transmission failure occurs during a TXOP, the station does not start a backoff procedure. Instead, it retransmits the failed frame after SIFS if there is enough time left in the TXOP to complete the transmission.

4.3.4 Network Event Detection Scheme

Several approaches have been proposed to optimize TCP performance in ad hoc networks by distinguishing between packet losses due to congestion and those due to other events. Loss of packet in ad hoc could be due to any network event like link failure, channel error, buffer overflow. The existing layer architecture does not support sharing of such information with non adjacent layer. More often packet loss is assumed due to buffer overflow, and such misinterpretation leads to initiate congestion control measures.

At the PHY layer, interference and fading may result in bit errors and packets loss. For wireless links, the bit error rate is several orders of magnitude higher than wired links. The TCP protocol was originally designed for wired networks, and its congestion avoidance mechanism does not consider link errors as a possible reason for packet errors or losses. Instead, TCP interprets packet losses caused by bit errors as congestion. This can significantly degrade the performance of TCP over wireless networks, when TCP unnecessarily invokes congestion control, causing reduction in throughput and link utilization. TCP faces challenges at all lower layers in the network stack in MANETs, especially due to the congestion control mechanism which has problems differing between congestion and other network communication events.

If packet loss is not due to congestion, TCP should not invoke the congestion control procedure, but take appropriate measures based on the type of event that caused the packet to be dropped. For instance, if a packet loss is due to lossy channels, TCP only need to retransmit the lost packets. ABET-EDCA proposes a diagnostic method to correctly detect the network event at the lower layer, and initiate an appropriate action at the upper layers.

4.3.4.1 Link Failure Detection

During the data transmission, on not receiving two consecutive ACKs, routing layer might incorrectly conclude it as a link failure event. To overcome this, the protocol runs a network event failure test to detect and differentiate a possible network events, including route disconnections, and buffer overflow, that may cause packet loss. A physical layer information is acquired to initiate corrective action according to the type of event. To detect the link failure, the ABET-EDCA implements the scheme discussed in Section 3.3.8.

4.4 Simulation Results

The performance of the proposed ABET-EDCA is evaluated using extensive simulations on the NS2.35 [96] simulator. In the simulation, the network comprised of 50 mobile nodes spread over an area of 1000 X 1000 m with an average speed of 5 m/s and pause time of 10 s is assumed. The node's mobility was modeled using the random waypoint mobility model. The network traffic comprised of both best effort TCP traffic and delay sensitive voice traffic. To ensure continuity of TCP traffic generation, an FTP file transfer was simulated. During TCP connections, the FTP clients upload a file to the FTP server. The File size was assumed to be 30 KBytes uploaded within every 20 s. VoIP flow was used to generate UDP voice traffic, which was encoded using ITU-T G.711 [115]. The average source bit rate and packet size of UDP voice traffic is assumed to be 64 kbps and 160 bytes respectively. Table shows the various parameters used at MAC and Physical layer. All of the simulations are run for 600 seconds and while making the measurements, the initial 100 seconds were ignored to allow the network to stabilize. In order to average out the randomness in the node mobility pattern, each scenario is repeated 10 times (i.e., ten unique traffic movement patterns) and mean of each performance parameter is computed.

TABLE 4.2: Scenario variations

Varying parameters	Constant parameters
TCP Connections = 3 - 15	Pause time = 10 s VoIP Connection = 1 Speed = 5 m/s
VoIP Connections= 1 - 5	Pause time = 10 s TCP Connections =10 Speed = 5 m/s
Node Speed = 2 - 10 m/s	Pause time = 2 s TCP Connections = 10 VoIP Connection = 2

4.4.1 Performance Evaluation

To reflect the advantages of ABET-EDCA, simulated results are compared with the results of IEEE 802.11e EDCA and E-TCP [108]. Similar to ABET-EDCA, E-TCP protocol is implemented to improve the TCP fairness in MANETs. AODV is used as the routing protocol at the network layer for all the three schemes. The performance is assessed by analyzing goodput, average TCP packet delay and TCP retransmissions. In the simulation, the protocols are compared and evaluated by varying three parameters as shown in Table 4.2.

4.4.1.1 Effect of Varying TCP Connections

The amount of TCP traffic generated is directly proportional to the number of TCP connections. In this section, the proposed protocol is evaluated by varying the TCP traffic connections from 3 to 15. Only one VoIP connection is established, and is randomly rotated among the nodes during the simulation period. All other simulation parameters are kept same as mentioned in the Table 4.2. The main motivation behind the simulation setup is to evaluate the TCP traffic performance in ABET-EDCA in the presence of UDP traffic.

Figure 4.2-4.4 shows plot of TCP packet delay, Goodput and TCP retransmission per second for IEEE 802.11e EDCA, E-TCP and ABET-EDCA protocols. It can be observed from Figure 4.2 that, as the TCP load increases in the presence of high priority constant VoIP traffic, the contention delay increases, which in turn result in increase in TCP packet delay. Increase in packet delay will result in more retransmissions thereby reducing goodput of the system which is as shown in Figure 4.3 and Figure 4.4. The delay curve of IEEE 802.11e EDCA shows the sharp incline. This is due to the fact that low priority TCP traffic gets fewer opportunities to access the channel in the presence of high priority VOIP traffic. This eventually increases retransmission of TCP data packets because of delayed transmission of TCP ACK packets. Hence, there is a degradation in goodput performance. The performance of E-TCP is better compared to that of IEEE 802.11e EDCA. E-TCP provides the high priority to ACK packets thereby reducing retransmission attempts. Hence, there is an improvement in delay and goodput performance. But as the traffic load increases, the network capacity is reduced and TCP data packet contention time is increased which degrades the overall performance. The proposed ABET-EDCA outperforms E-TCP.

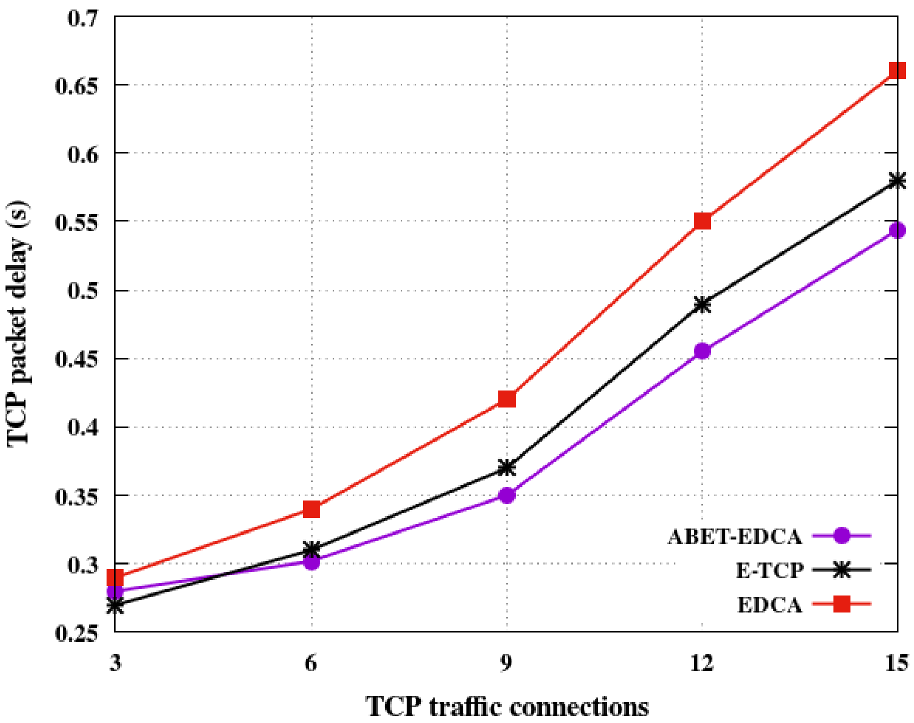


FIGURE 4.2: TCP packet delay v/s TCP connections

Because of priority queue implementation of AC(1) queue in ABET-EDCA, the ACK packets gets the higher priority compared to TCP data packets. ABET-EDCA protocol also allows tuning of CW and TXOP parameters. The dynamic adaptation of CW parameter under the heavy TCP traffic load increases the channel contention

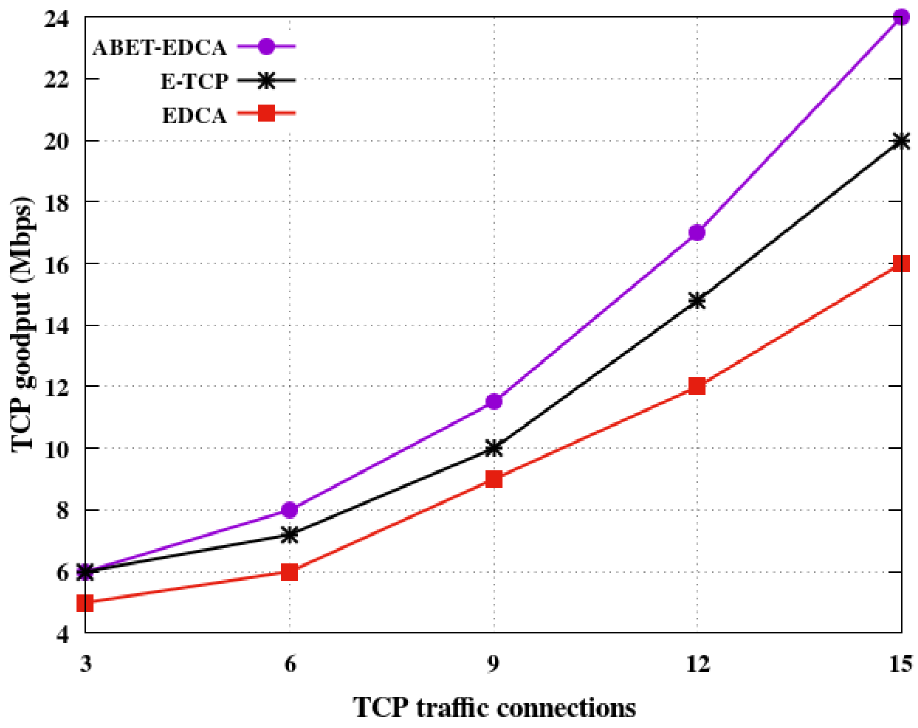


FIGURE 4.3: TCP goodput v/s TCP connections

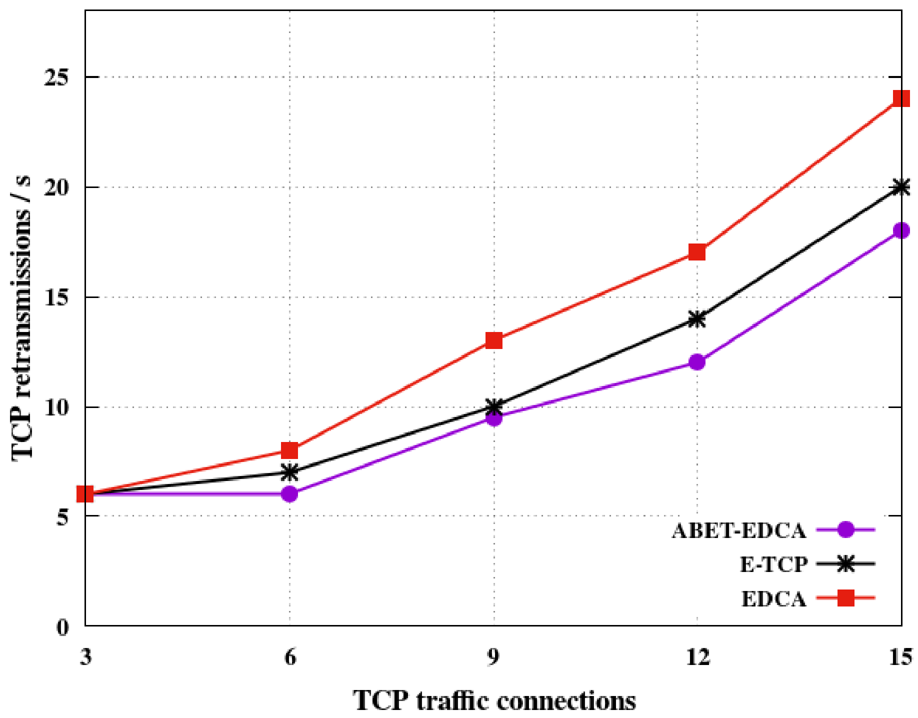


FIGURE 4.4: TCP retransmission v/s TCP connections

probability of AC(1) and the tuning of TXOP parameter helps to send permissible TCP data packets along with ACK packets in burst. Hence, the average delay of TCP packets is lesser compared to other protocols. This ensures higher number of TCP packets successfully reaches the destination and are acknowledged back on time thereby improving the TCP goodput and retransmission performance.

4.4.1.2 Effect of Varying Voice Connections

In this simulation, the voice traffic load is varied by varying the number of voice connections between 1 to 5. The TCP traffic consists of 10 FTP connections and after every 20 seconds the FTP file transfer is initiated. The main motivation behind this simulation setup is to evaluate the TCP traffic fairness in the presence of high priority UDP traffic.

Figure 4.5 - 4.7 illustrates the plot of TCP packet delay, goodput, and packet retransmissions per sec v/s voice traffic connections for IEEE 802.11e EDCA, E-TCP and ABET-EDCA protocol. As the high priority UDP traffic load increases, the low priority TCP packets find it more difficult to get the channel access. In IEEE 802.11e EDCA, the UDP packet gets the majority of channel access, making the TCP packets to starve for a longer duration. This results in an overflow of queue which leads to packet drops and increase in retransmission due to timeout as shown in Figure 4.7. This subsequently results in increased packet delay and reduced goodput. In E-TCP, although ACK packets get the highest priority, TCP data packets do not get any preference in high UDP traffic and hence the performance deteriorates.

It can be observed from the Figure 4.5 - 4.7 that although the packet delay increases with the UDP traffic load, the ABET-EDCA shows considerably better performance compared to other two protocols. This is because, whenever the AC(1) queue gets the channel access because of the adaptive contention window mechanism, the TXOP limit computed dynamically, helps to send permissible TCP data packets along with ACK packets to reach the destination and get acknowledged instantly for successfully delivered packets. Hence, ABET-EDCA shows better goodput and packet retransmission are minimal.

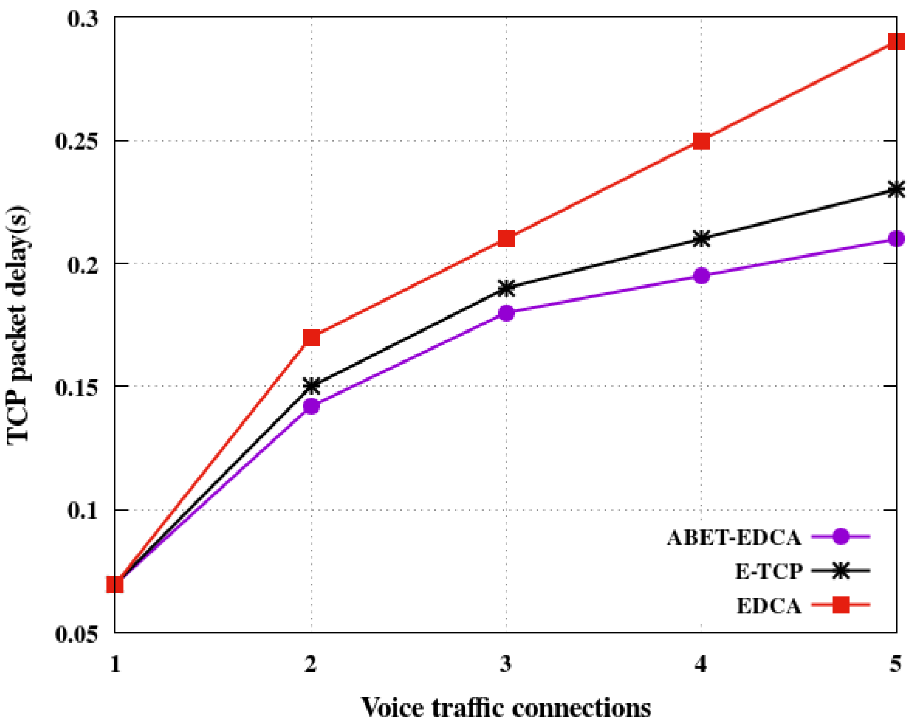


FIGURE 4.5: TCP packet delay v/s Voice connections

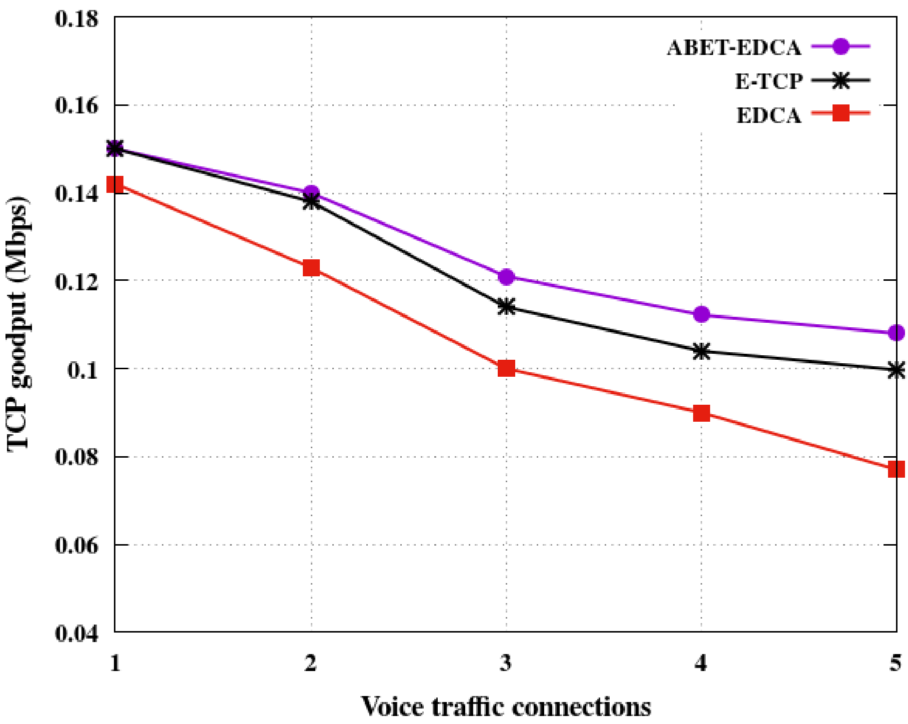


FIGURE 4.6: TCP goodput v/s Voice connections

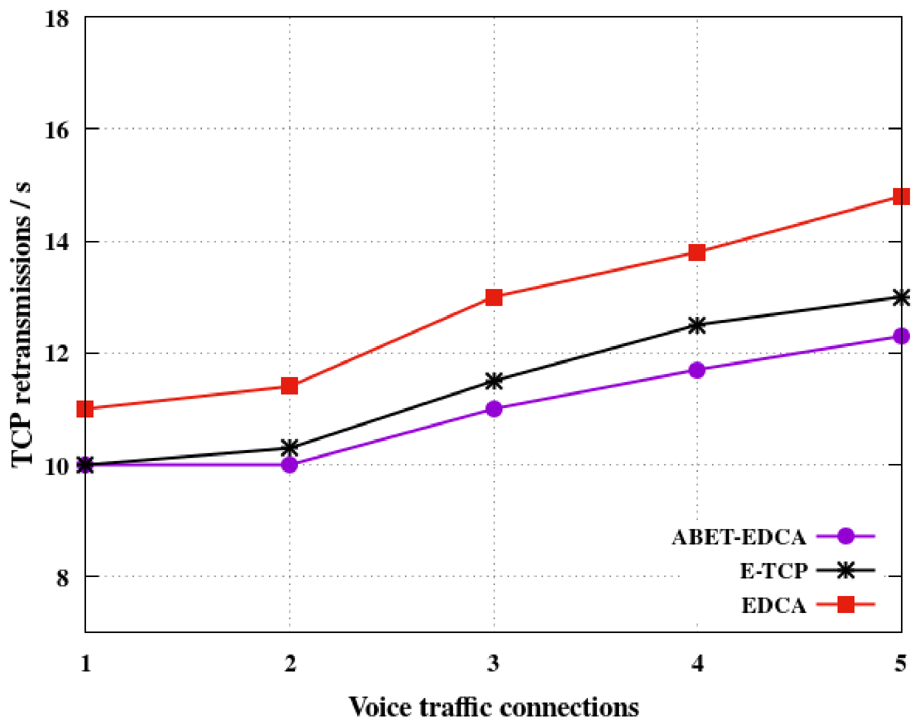


FIGURE 4.7: TCP retransmission v/s Voice connections

4.4.1.3 Effect of Varying the Node Speed

In this simulation setup, the nodes speed is varied from 2 m/s to 10 m/s to evaluate ABET-EDCA performance under mobility. The pause time of 2 s is assumed. To achieve this, 10 TCP connections are assumed, in which an FTP file is uploaded after every 20 s. Two VoIP connections are initiated for 50% of the network simulation time. All other simulation parameters are assumed same as mentioned in Table 4.2.

Figure 4.8-4.10 shows the TCP packet delay, TCP goodput and TCP packet retransmissions result analysis of the three protocols. As the node's speed increases, a possibility of link failure also increases. In IEEE 802.11e EDCA and E-TCP, performance degrades with the increase in node's speed. When ACK packets are lost or delayed due to a link failure, both the protocols assume that there is congestion and initiates packet retransmissions. This results in additional packet delay and also affects the TCP goodput. The proposed ABET-EDCA shows comparatively improved performance than other two protocols in terms of packet delay, TCP goodput and retransmissions when the node speed is above 6 m/s. This is because, ABET-EDCA

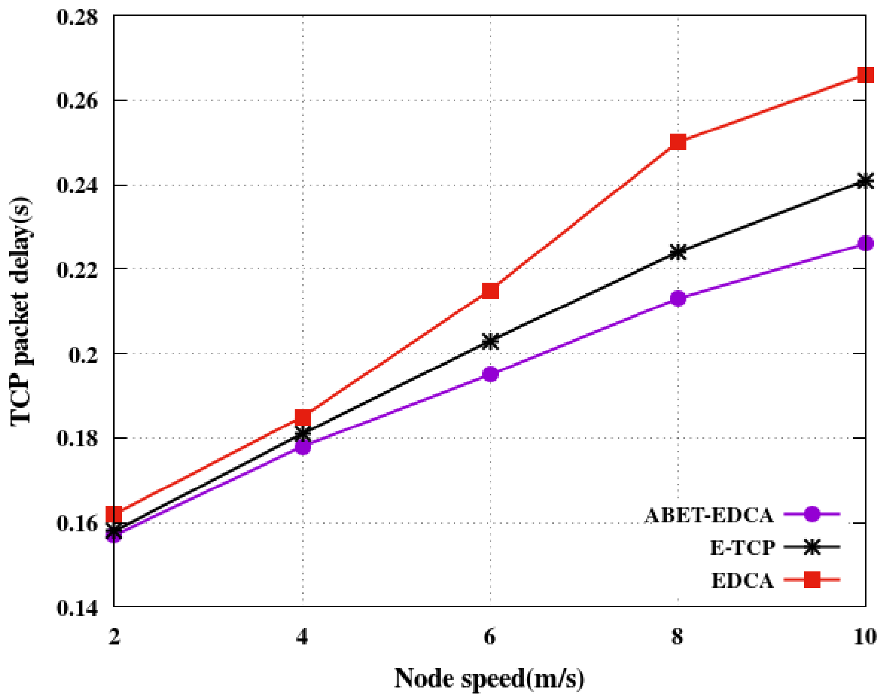


FIGURE 4.8: TCP packet delay v/s Node speed

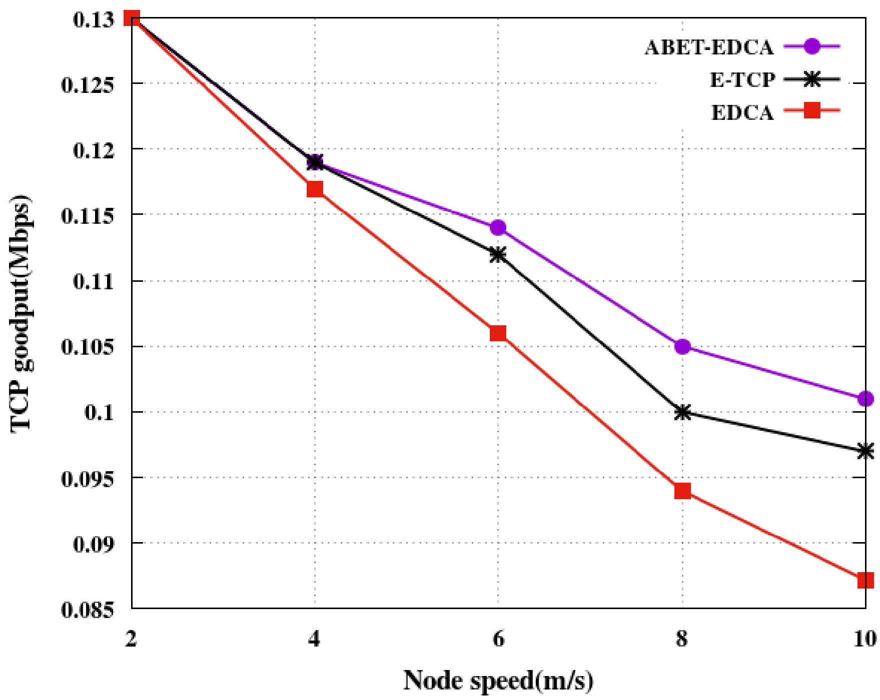


FIGURE 4.9: TCP goodput v/s Node speed

has a mechanism to detect the route failure or predict the nodes moving out of interference range. This helps in preventing the delayed ACK packets from retransmitting and thereby reducing wastage of the channel bandwidth.

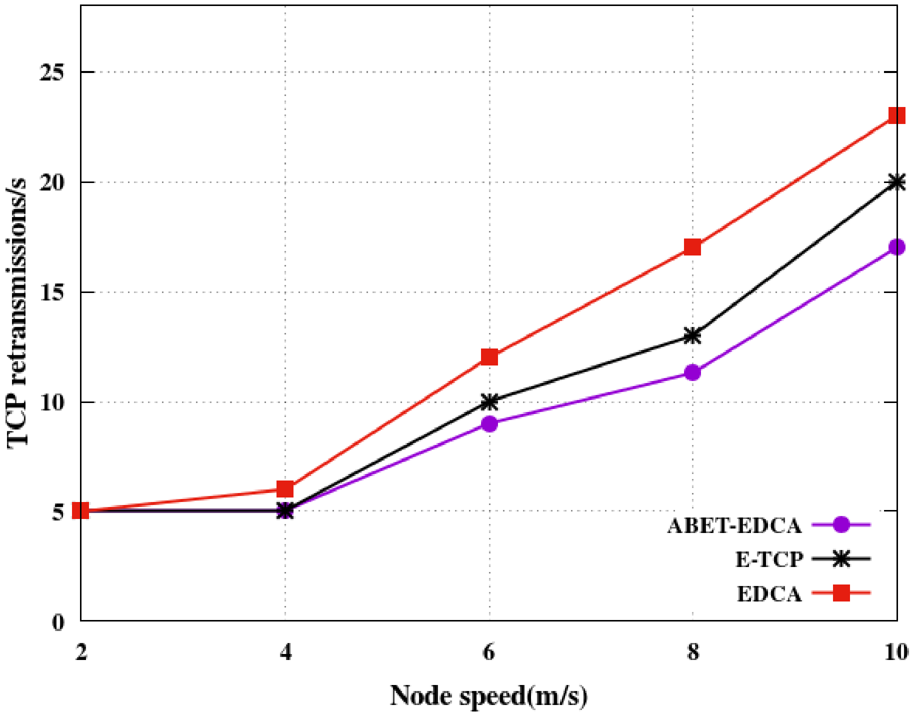


FIGURE 4.10: TCP retransmission v/s Node speed

4.5 Conclusion

In this chapter, a congestion control method for TCP over 802.11e called Adaptive Best Effort Traffic Scheduler for EDCA (ABET-EDCA) is proposed. The chapter investigates how to enhance the performance of low priority TCP-based applications in the presence of high priority UDP-based traffic. ABET-EDCA scheme prioritizes the TCP ACK packets during congestion by tuning the MAC parameters based on traffic load conditions. During the high TCP traffic flow, the protocol simultaneously implements a dynamic TXOP adaptation at the MAC level to send permissible TCP data packets along with ACK packets to reach the destination and get acknowledge instantly for successfully delivered packets. This results in minimal retransmissions improving the average packet delay. The effectiveness of the proposed algorithm has been shown by extensive simulations carried with heterogeneous traffic like FTP and VoIP traffic. The ABET-EDCA shows the improved fairness for TCP traffic compared to IEEE 802.11e and E-TCP under varying heterogeneous traffic load and node mobility. The ABET- EDCA shows better performance in terms of TCP packet delay, TCP goodput and packet retransmission.

Chapter 5

Multi Objective Cross Layer Optimization for IEEE 802.11e

5.1 Introduction

The user demand for multimedia application over the Internet has been rapidly growing in the past few years. VoIP, and especially Internet Protocol Television (IPTV) and Video on Demand (VoD) applications are gaining an ever increasing popularity. These applications are highly sensitive to delay and jitter, which affect the quality of the application output. This increasing popularity places new challenges to networks and devices that are used to connect the networks [116].

The IEEE 802.11e was initially proposed for wireless LANs with APs to provide QoS for real time traffic. Even though, the service differentiation is done at the MAC layer by IEEE 802.11e to ensure QoS, excessive traffic flow belonging to the same AC may result in performance degradation. This is due to the increase in the level of contention among the flows and may result in overloading of nodes. Service differentiation at MAC layer combined with a routing layer solution that can detect and avoid the overloaded nodes that are busy forwarding high priority packets, can minimize the packet drop and maximize the throughput.

ABET- EDCA protocol proposed in Chapter 4, prioritizes the TCP ACK packets thereby improving best effort traffic performance. Extending this idea, in this chapter, a Multi Objective Cross Layer Optimization (MOCLO) protocol is proposed to improve the performance of both real time and best effort traffic in terms of packet delay, and throughput. The MOCLO protocol proposes a modification to IEEE 802.11e MAC layer to prioritize and protect the important packets within each class of traffic flow. It uses cross layer interaction between PHY-MAC-Network layers to find multiple optimal route to support different flows in AC based on priority thereby reducing the load on particular route. The performance of the network with MOCLO protocol is thoroughly evaluated through the extensive simulations, which highlight the advantages of proposed approach.

5.2 A Literature Review

In wireless networks with access points, recent research has revealed how to achieve QoS guarantee in terms of fairness and delay [117] [118]. However, in MANETs there are more challenges due to channel interference, limited resource availability, signal fading, dynamic topological changes caused by nodes' mobility, and low data rate due to energy constraints in the communication path. In a congestion prone network, routing the delay sensitive multimedia packet through less congested route is a challenging task [84]. When such data packets get buffered in a queue beyond the threshold limit, the packet drop rate increases thereby degrading the throughput and delay performance[119]. In such situations, priority scheduling can enhance the throughput and delay performance of traffic with multimedia delay sensitive packets.

At the physical layer, issues like channel interference, multipath and channel fading makes it difficult to offer QoS [120]. At the MAC layer, shared channel access causes exposed and hidden terminal problems, which results in complications associated with bandwidth reservation and channel fairness. At the network layer, routing protocol should adapt quickly to the topological changes due to nodes' mobility. Some of the research proposals found in the literature, deals with handling QoS issues related to single layer like physical layer, MAC, or routing layer [121]. The physical layer of 802.11 supports various transmission modes based on wireless channel characteristics. The data rate adaptation is a method to change dynamically the data transmission rate and adapt to corresponding channel coding and modulation techniques. Most of the proposed rate adaptation schemes[122][19] consider only the wireless channel conditions at PHY layer. In ideal channel conditions, the network throughput is directly proportional to the data transmission rate. In MANETs, the channel is error-prone and vary over time. Hence, the predetermined high rate transmission for delay sensitive multimedia traffic may result in an increased packet drop over the channel and degrading the network throughput.

IEEE 802.11e WLAN performance is mainly dependent on EDCA parameters and have static values. To meet dynamic requirements of QoS in MANET, EDCA parameters should be made adaptive. Most of the recent work is based on changing three EDCA parameters such as TXOP, AIFS and the contention window (CW) size.

There have been efforts to improve IEEE 802.11e QoS by optimizing EDCA parameters based on different traffic types. To meet the specification of QoS, Pablo Serrano et al. [123] proposed a tuning mechanism for IEEE 802.11e EDCA which dynamically tune the parameters of EDCA for both realtime and non real time traffic to maximize the performance. Yang Xiao et al. [124] proposed global data parameter control for the IEEE 802.11e EDCA. The approach provides good differentiation among different ACs and good fairness among real-time streams within the same AC. L. Romdhani et al. [125] proposed an adaptive service differentiation scheme for QoS enhancement in IEEE 802.11 called Adaptive EDCF (A-EDCF) scheme, wherein the contention window of each traffic class is adapted according to the estimated collision rate in order to improve the goodput of the traffic under heavy loads. S. Prasetya et al. [126] proposed an adaptive contention window algorithm for QoS improvement with 802.11e EDCA. The contention window adjustment is based on the number of stations involved and collision probability in the network. Tinnirello et al. in [127], observed that the only CW differentiation is not enough and hence proposed a scheme with both minimum contention window and AIFS. The scheme proposed different AIFS for each

AC thus providing good differentiation in the access delay between ACs while CW_{min} provides good control of throughput differentiation.

There are various QoS aware routing protocols. Perkins et al. in [12] proposed a QoS-AODV routing protocol where the optimal path from source to destination is found using hop count as a basis along with two other QoS requirements, delay and bandwidth. The merit of QoS-AODV is that it excludes some nodes that do not satisfy the QoS requirements before establishing the route and hence reduces unnecessary transmission of RREQ thereby saving routing establishment process overhead. Qi Xue et al. proposed an on-demand ad hoc QoS-aware routing protocol AQOR [13] based on metrics like available bandwidth estimation and end-to-end delay measurement.

The key components of MANET design is routing, data rate selection and channel access scheduling. These components are distributed across various layers of the protocol stack. The routing is managed at the network layer, channel scheduling is handled at the MAC layer, whereas the data rate is controlled sometime at MAC layer or physical layer. In recent times it has become apparent that a traditional layered network approach of separating routing, channel scheduling, and data rate control is not efficient for MANETs. Also, the approach is inflexible and incapable of coping up with the dynamics of MANETs supporting real time multimedia applications. Despite many researchers proposing several mechanisms to improve QoS of real-time traffic, the experimental performance obtained are not optimal. This is mainly because, at MAC layer, EDCA parameters are fixed and cannot be made adaptive to the variation in traffic characteristics and load conditions. Most of the work discussed above, assumes the ideal channel condition or identical link conditions among the participating host while designing the solution, which may not be true in a realistic wireless environment. With the heterogeneous traffic flow in the network, tuning only the EDCA parameters may not provide desired differentiated QoS. Providing QoS service only at the MAC layer which is not sufficient as overall QoS achievement also depends on the cooperation of other layers in the protocol stack. For example, at the network layer, the existing on demand routing protocol for ad hoc network do not consider application level QoS requirements while making a routing decision.

There are primarily three main challenges for QoS support in MANETs:

1. Handling time-varying network conditions: IEEE 802.11e does not take into account varying network conditions like channel condition, network load, and topological changes due to nodes mobility. An increase in the number of users in the network, leads to longer channel access defer periods and higher collision probability. The Degradation channel condition can weaken the QoS differentiation mechanism of IEEE 802.11e, so that it does not work as intended.

2. Adapting to varying application profiles: The QoS requirement is application specific. Estimating QoS requirements correctly is crucial while designing and tuning the medium access mechanism. Poor estimation leads to unacceptable delays, buffer overflows and inefficient resource utilization.

3. Managing link layer resources: Since IEEE 802.11e is a MAC layer enhancement, there is a need for some kind of link layer cooperation, so that link layer resources can be optimally managed. Additionally, an admission control scheme, and scheduling algorithms should be designed. Admission control optimally allocates the resources to the traffic flow based on the service request. Scheduling algorithm handles packets at the network layer and decides which packets to forward. Both of these are crucial in providing QoS in wireless networks. IEEE 802.11e amendment does not specify any admission control process or specifies any scheduling algorithm.

To address above three challenges, a Multi Objective Cross Layer Optimization (MOCLO) between PHY-MAC-Network layers are proposed. The proposed cross layer approach helps in achieving the QoS goal for real time and best effort traffic by implementing necessary components like node disjoint multipath route, priority scheduling, rate adaptive congestion control, and distributed MAC.

The following are the key features of the MOCLO protocol:

1. Each of the AC's in 802.11e is implemented as a priority queue scheduler rather than the classical FIFO queue to prioritize the transmission of the traffic.

2. A congestion aware QoS metric based disjoint multipath routing protocol is designed, to support multiple QoS factors.

3. A differentiated service routine for different AC has been designed by implementing a mapping function between the MAC and the network layer queues based on their QoS requirement.

4. The cross layer approach is further enhanced by exploiting the multi-rate link adaptation function to select the appropriate transmission rate per frame basis based on channel state information obtained.

5.3 Multi Objective Cross Layer Architecture Design

Applications are categorized into real-time and best effort based on their sensitivity to packet delay. Even though preferential treatment to real-time data is given by EDCA at MAC layer, network performance degrades when additional real-time flows are injected into the network, resulting in the loss of delay sensitive packets. Increase in real time traffic results in high priority queue buildup and eventually leads to increased contention among the nodes during channel access. This section introduces QoS architecture which assists both the real-time and best effort communication in MANET environments. Node mobility is one of the main design complication for MANETs, which results in connection disruptions, leading to an escalated packet loss. The major cause of this issue is that 802.11e manages the problem of QoS only in the MAC layer. Hence the primary aim is to utilize the cross layer information to contemplate a multi-layered QoS metric for routing. The Figure 5.1 shows the cross layer architecture proposed in MOCLO algorithm. The congestion control component is extremely important in the QoS architecture. For computing QoS aware node disjoint multiple paths, physical layer, and MAC layer share the node mobility and the queue load information respectively with the Network layer. Distributed QoS MAC is used in the architecture design which is based on an enhanced IEEE 802.11e EDCF function. The priority queue scheduler at the MAC layer services the packets in the order of highest priority to lowest priority sequence. The MAC layer provides the congestion status of ACs to the transport layer. The transport layer, in turn, uses this congestion information to adjust data rate dynamically. The mapping function maps the high priority packets at MAC layer with the most efficient path in the network layer.

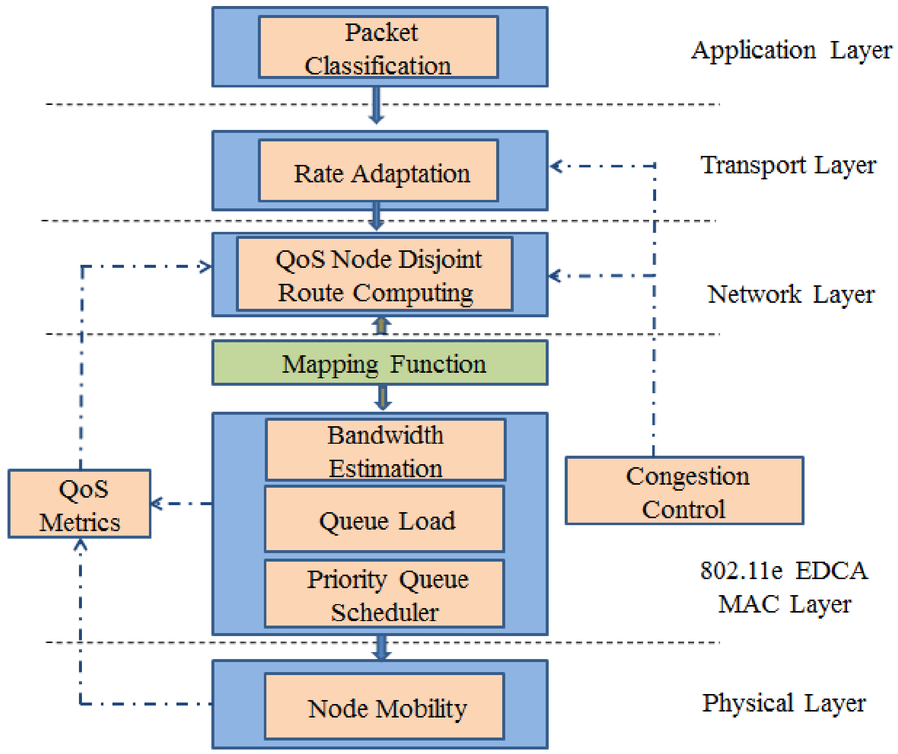


FIGURE 5.1: Proposed cross layer architecture in MOCLO

5.3.1 QoS Metrics

Following three metrics are used in the MOCLO algorithm to improve the QoS support for IEEE 802.11e.

1. Relative Mobility Factor: The mobility of the nodes is a complicated factor that significantly affects the effectiveness and performance of the routing protocol. Routing overhead can be reduced and more steady routes can be found by preventing nodes with excessive mobility to be a part of the route selection mechanism. Hence the mobility factor is used to minimize the impact of route failure due to node movement.

The relative mobility of a node can be calculated based on the received signal strength which is an indicative of the distance between transmitting and receiving node. In an ideal channel condition, Friis' free space propagation model can be used to model the mobility. In Friis free space [128] propagation model, the physical distance between any two neighboring nodes exhibit inverse-square dependence of the ratio of received and the transmitted power. But in a practical scenario, the estimation of this parameter is difficult due to the varying channel conditions. Hence, in the proposed scheme, the power ratio of two consecutive Hello packets is used for this estimation. The relative mobility $M_Y^{ref}(X)$ at node Y w.r.t. node X is computed as [129],

$$M_Y^{ref}(X) = 10\log_{10}\frac{RxPr^i_{X\rightarrow Y}}{RxPr^{i-1}_{X\rightarrow Y}} \tag{5.1}$$

where,

$RxPr^i_{X\rightarrow Y}$ is the current received signal power of packet from X to Y,
$RxPr^{i-1}_{X\rightarrow Y}$ is the previously received signal power of packet from X to Y.

A low value for $M_Y^{ref}(X)$ means that node Y has less mobility with respect to its one hop neighbor X and vice versa. If, $RxPr^i_{X\rightarrow Y} < RxPr^{i-1}_{X\rightarrow Y}$ then $M_Y^{ref} < 0$ which indicates that the two nodes are moving away with respect to each other. On the other hand, $RxPr^i_{X\rightarrow Y} > RxPr^{i-1}_{X\rightarrow Y}$, indicates that the nodes are moving closer to each other.

For a node with n neighbors, there will exist n such values for $M_Y^{ref}(X)$. Therefore, aggregate local mobility value $M_Y^{aggr}(X)$ at node Y is calculated by calculating the variance of the entire set of relative mobility samples $M_Y^{ref}(X)$ with respect to zero, where X_j is a neighbor of Y:

$$M_Y^{aggr} = var[M_Y^{ref}(X_j)]^n_{j=1} = E\left[\left(10\log_{10}\frac{RxPr^i_{X\rightarrow Y}}{RxPr^{i-1}_{X\rightarrow Y}}\right)^2\right] \tag{5.2}$$

A low value of $M_Y^{aggr}(X)$ indicates that Y is relatively less mobile with respect to its neighbors, On the contrary, a high value of $M_Y^{aggr}(X)$ indicates that Y is highly mobile with respect to its neighboring nodes. If for the node Y, relative mobility value is in the range of the $0.08 < M_Y^{aggr}(X) < 0.5$, which indicates a high mobility for a node and the proposed algorithm avoids the route with such nodes.

The relative mobility metric $M_Y^{aggr}(X)$ calculation is based on two successive packet transmissions received within the time interval t. It is possible that during the time interval t, certain nodes may move out or come into range of node Y. So to deal with this issue, nodes which do not participate in two successive transmissions to node Y are excluded from the calculation. Thus, the nodes which have been in the neighborhood of Y for time interval t qualify for the mobility metric calculation. This calculation is performed at every node, and thus, each node will maintain an aggregate relative mobility metric.

2. Queue Utilization Factor: In EDCA, when packets arrive from the application layer to the MAC layer of a node, they are placed into the respective AC's based on their priority. During high multimedia traffic flow, the network throughput decreases heavily due to several factors like starvation of packets, collision, or buffer overflow. Hence, the queue load at the node's interface is an important factor which indicates the congestion level. In the proposed algorithm, the average queue load at each AC(i), where $1 < i \leq 3$, is estimated by calculating the weighted moving average of each AC's queue load over time $\triangle$t, which gets recomputed after specific intervals.

$$Q_{avgloadAC(i)}(t) = (\alpha) * Q_{loadAC(i)}(t) + (1 - \alpha) * Q_{avgloadAC(i)}(t - 1) \tag{5.3}$$

where,

$Q_{avgloadAC(i)}$(t) is the average queue load at each AC(i),

$Q_{loadAC(i)}$(t) is the current load at AC(i),

α is a constant smoothing factor between 0 and 1.

The average queue utilization $AC(i)_{util}$ for each of the AC(i) is computed as:

$$Q_{AC(i)util} = \frac{Q_{avgloadAC(i)}(t)}{Q_{size}} \tag{5.4}$$

where, Q_{size} be the size of the AC queue in a node.

3. Bandwidth Estimation: In a delay tolerant network, to gain the high performance, the proposed protocol considers the available end to end bandwidth between neighbor nodes. In wireless ad hoc networks, bandwidth is a shared resource within the interference range of one hop nodes. Due to the random nature of the wireless channel, available bandwidth also changes randomly. Increase in traffic load over the path leads to increase in MAC layer overhead in terms of contention, which in turn results in packet drop. Hence, estimating the effective bandwidth plays a vital role in network efficiency. The protocol finds the residual bandwidth available at each node that can support a stream of traffic. While estimating the available bandwidth, the MAC overhead and data forwarding rates of neighbor nodes in the interference range are considered. The available bandwidth at node is computed as given as,

$$BW = DR_c - \left(\frac{\sum_{i=1}^{\eta} (\gamma_i + \beta_i)}{\eta} \right) \tag{5.5}$$

where,

DR_c is the data rate of the wireless channel.

γ represents the aggregate data generation rate of all the nodes which are in the interference range of the node that is computing available bandwidth which also includes the self-data traffic rate,

γ_i indicates the value of γ at the i^{th} index of averaging window,

β_i represent the MAC layer overhead at the i^{th} index of averaging window,

η represent the average window size,

The proposed MOCLO protocol uses weight based QoS metric by assigning the weights dynamically to each node and these weights are known as network utilization weight. Each node computes its network utilization weight NU^i at a regular interval of time. The metrics are normalized by their maximum values with some multiplicative factors and is given as,

$$NU^{i} = \rho_1 \left(\frac{Q_{avgloadAC(1)}(t)}{Q_{AC(1)max}} + \frac{Q_{avgloadAC(2)}(t)}{Q_{AC(2)max}} + \frac{Q_{avgloadAC(3)}(t)}{Q_{AC(2)max}} \right) + \rho_2 \left(\frac{BW}{BW_{max}} \right) + \rho_3 \left(\frac{M_Y^{ref}(X)}{M_Y^{max}(X)} \right) \tag{5.6}$$

where,

$AC(1)_{max}$, $AC(2)_{max}$, and $AC(3)_{max}$ are the maximum MAC queue size,

BW_{max} is the maximum bandwidth at the node in the network,

$M_Y^{max}(X)$ is maximum possible mobility for a node.

$\sum_{i=1}^{3} \rho_i = 1$ denote the adjustment factor that shows the importance of each QoS parameter (bandwidth, queue load, and node mobility) on the path.

Each of the QoS metric is assumed to have the same priority for the purpose of the our study. This weighting factor can be varied to alter the significance of each metric.

5.3.2 QoS Aware Disjoint Multipath Routing

The MOCLO algorithm considers a multi QoS metric for the admission control, according to the network utilization weight and the congestion factor. The algorithm adapts a source routing mechanism, where, only the destination node replies to the source in an unicast manner to pick a node disjoint route with minimum threshold route utilization value even if intermediate nodes have the route entry in the route cache. This is because, the route record in route cache of intermediate node may not satisfy the QoS requirement of the new route request. The source node then selects the route with the minimum network utilization weight value. Hence, more stable routes with less congestion are selected.

5.3.2.1 Route Discovery

The proposed MOCLO protocol uses AODV as the base routing protocol with modifications to accommodate QoS metrics during route discovery. The algorithm uses multipath approach so that in case of node mobility, the data transmission can continue through an alternate secondary path. At the time of route discovery, the source node sends the RREQ packets to the destination. The <source-address, sequence number> tuple uniquely identifies the RREQ packet so that nodes can refrain from transmitting duplicate RREQ packets. The intermediate nodes on receiving RREQ packets, cache them in a Request Received Table. To ensure cycle-free disjoint paths, each intermediate node compares newly arrived RREQ packet sequence number with Request Received Table entry and drop all the RREQ packets with the sequence number lower than that of the cached RREQ packet. The incoming RREQ packets are also dropped by a node when its mobility feature value $M_Y^{aggr}(X)$ is above 0.5 indicating high mobility, or if the queue load of any of the three mentioned ACs namely, AC(1), AC(2), AC(3) goes beyond a threshold limit Q_{thresh}.

The RREQ packet header is modified to accommodate the minimum network utilization value NU^i. The source node initially sets this value to be 0. The Figure 5.2 demonstrates the route request forwarding scheme of MOCLO protocol at the intermediate node. An intermediate node may receive more than one RREQ packets. On receiving multiple RREQ within time $RREQ_{wait}$, the intermediate node selects the RREQ packet with a minimum value of NU^i. The intermediate node, then updates the NU^i field in the RREQ packet by adding its own NU^i value and broadcast it to its neighbours.

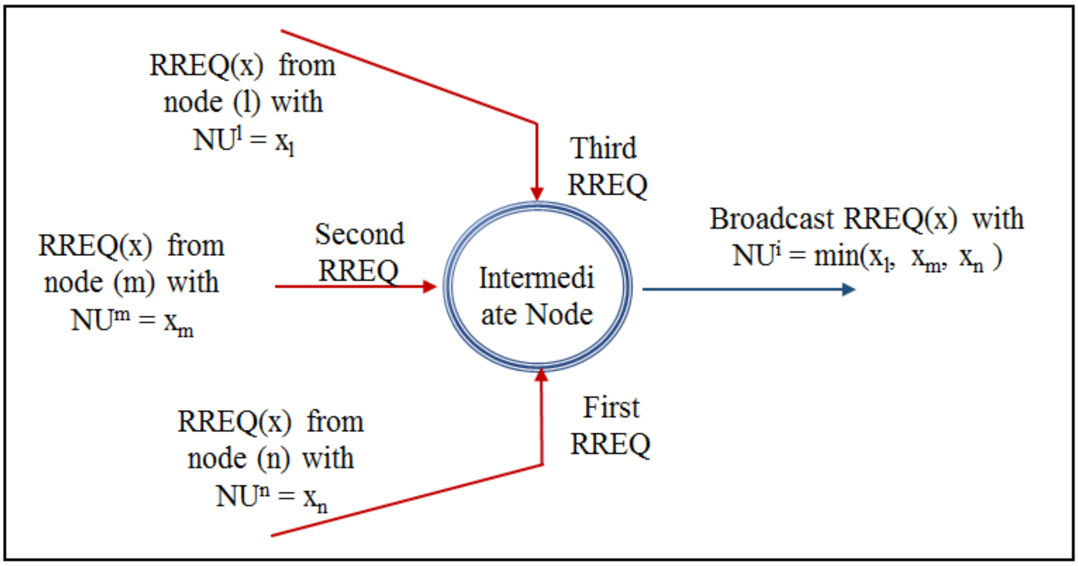


FIGURE 5.2: RREQ forwarding process at intermediate node

5.3.2.2 Route Reply

The destination node may receive more than one RREQ packet. When the destination node receives the first RREQ packet, it waits for a period $RREQ_{wait}$. After the timeout, the destination node selects three routes with least network utilization value NU^i, and sends a route reply (RREP) packet along the selected paths.

The modified RREP packet includes an additional field NU^i to ensure that the route selected for RREP packet forwarding meets the requirements of QoS criteria and is initially set to 0. The intermediate node route table is also modified to contain an additional field called "seen-flag" to store a flag value, which is initialized to 0. The intermediate node on receiving the RREP packet for the first time, makes an entry in the route table and sets the seen-flag to 1. The intermediate node, then updates the NU^i field in the RREP packet by adding its own computed NU^i value and forwards it towards the source node. In case a duplicate RREP packet for the same RREQ message is received by the intermediate node, it is dropped if seen-flag is 1. This mechanism ensures that every intermediate node participates in only one of the multiple paths found between the source and the destination. The source node on receiving the first RREP packet waits for an additional $RREP_{wait}$ time to receive more RREP packets. After $RREP_{wait}$ timeout, the source node chooses the most optimum path as the primary route, on the basis of minimum NU^i.

5.3.3 Priority Queue Scheduling

Transmitting the delay sensitive packets using highest priority will enhance the QoS of the application. The MOCLO protocol proposes the priority scheduling scheme to provide high preference to high priority packets placed in MAC queues without starving them for a longer duration. In MOCLO, the priority queue scheduling scheme is proposed for both multimedia video traffic in AC(2) and best effort traffic in AC(1).

5.3.3.1 Prioritization of Video Packets

MPEG-4 is one of the popular video compression standards and is mainly targeted for internet based multimedia video streaming used today. The encoded video stream is composed of a series of Group of Picture (GOP) with three different types of frames, I frames (Intra-coded), P frames (Predictive-coded), and B frames (Bi-directionally predictive-coded). I frames are the least compressible and don't require the information about other video frames to be decoded. P frames use the information from previous frames (I frame and P frame) to decompress and are more compressible than I frames. B frames use both previous and forward frames as reference data to get the highest amount of data compression. Figure 5.3 shows a typical GOP pattern with frames IBBPBBPBBI, where the curved arrows show the coding dependencies, and the number below the individual frame stands the importance level of that frame. Therefore, I and P frames are considered to be the most important video frames compared to B frames in a video sequence. To achieve higher performance, a scheme which supports higher priority for I and P frames is needed. In 802.11e EDCA, the

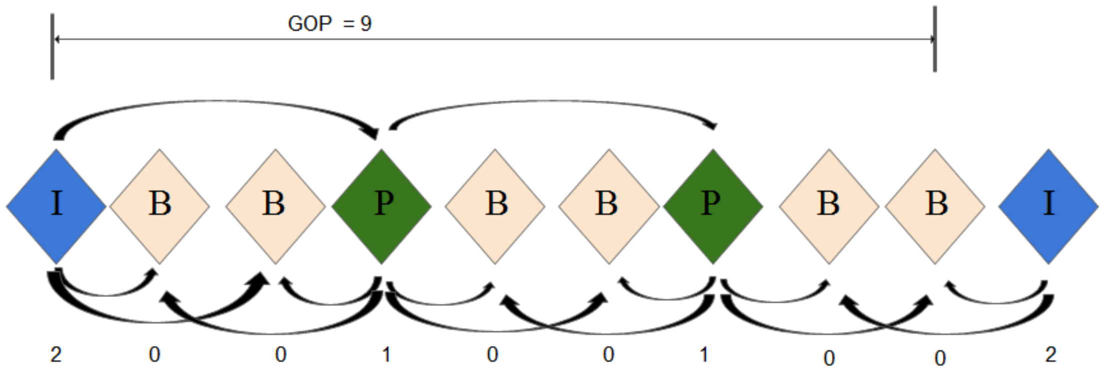


FIGURE 5.3: Group of pictures in encoded video stream

AC's are implemented using the simple queue data structure with "first-come, first serve" policy and each traffic class belongs to one of the AC queue, in which packets are ordered as they arrive. This ordering affects directly the way the packets are removed from the queue and hence not suitable in an application where priority is required. In case of multimedia video traffic, during the high traffic load, buffering the I and P frames for long duration may lead to delay in decoding the frames at the receiver node, thereby degrading the system performance. The proposed MOCLO protocol uses a priority queue rather than conventional FIFO queue to improve the system performance. The use of priority queue prevent higher priority packets from starving in queue for longer period thereby efficiently utilizing the network resources. The priority queue is implemented using heap data structure [111].

To ensure a delay-sensitive multimedia video stream communication within desired delay time limits, MOCLO proposes a solution to control the queuing delay of each frame in the queue entirely based on the priority of video frames. The application layer passes the video frames (I/P/B) information to differentiated service code point (DSCP) field of Network layer and then to the MAC layer. Based on the frame information received, MAC layer will perform the mapping to allocate video frames (I/P/B) into the intended AC queue. In the proposed scheme, the multimedia video packets arrive at AC(2). In AC(2), the I and P packets are prioritized compared to B packets, such that these packets will be transmitted on priority.

5.3.3.2 Prioritization of ACK Packet

Since most of the network traffic is still dominated by TCP traffic, it is important to address the concerns of the TCP traffic. To address the issue, the MOCLO uses a scheme proposed in Chapter 4, Section4.3.1, wherein the TCP ACKs are prioritized to improve the TCP throughput while using the IEEE 802.11e MAC. The AC(1) queue is implemented using a priority queue allocating the ACK packets with the highest priority. This ensures the TCP packets that managed to get successfully transmitted are acknowledged instantaneously. This effective mechanism helps in enhancing the network performance. The Figure 5.4 demonstrates the packet prioritization at AC(2) and AC(1).

5.3.4 QoS Mapping between the MAC and Network Layer

Currently, in most of the existing multipath routing protocols, both high priority traffic and low priority traffic is routed over a single route [92] [94] and if the route with the optimal QoS metric is selected every time, then this route would get more loaded than

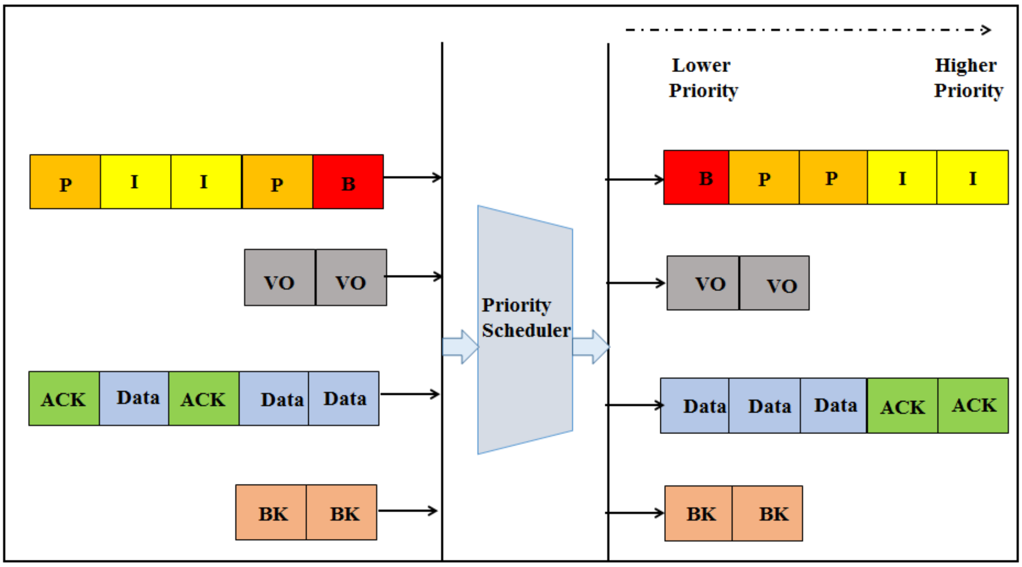


FIGURE 5.4: Priority queue scheduler for AC(2) and AC(1)

others. This may lead to congestion and results in packet loss. Also, some of the nodes along the chosen route may die due to depletion of battery. Hence, load balancing in MANET is important to minimize the probability of disconnecting or partitioning the network thereby ensuring long network lifetime.

The MOCLO proposes an effective service differentiated routine for different AC's by implementing a mapping function between the MAC and the network layer queues for each of the AC's based on their QoS requirements. To improve the performance of real time applications and maximize the channel utilization, the transmission of high priority packets must be given higher privilege. Therefore, the proposed scheme provides a different route for each class of service. There are two cases:

1. When the source and the destination connection pair is different for the real time and best effort traffic:
   For a real time traffic, the source node uses the best optimal route to send I and P-frames, and the second best route is used to send B-frames. For best effort traffic, the source uses the best optimal route to send all best effort traffic dequeuing ACK packets first.

2. When the source and the destination connection pair is same:
   The proposed scheme maps the most optimal route to send the highest priority

real time packets, while the best effort traffic packets get transmitted through the secondary route.

Figure 5.5 demonstrates the mapping mechanism proposed in MOCLO protocol.

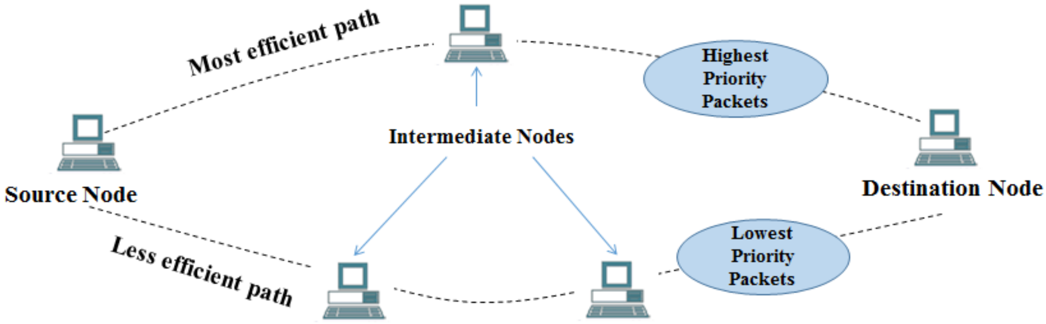


FIGURE 5.5: Multipath scheme to split the traffic

5.3.5 Data Rate Adaptation

Real time applications need service differentiation compared to other forms of traffic flow. The delivery of such delay sensitive packets outside the time boundary is considered as lost packets. Thus, it is vital to integrate the QoS with the rate adaptation technique to improve delay performance of high priority multimedia traffic over others. Most of the existing rate adaptation algorithms adopt probe-based approaches that search for the most optimal transmission rate using the hit-and-trial methods and focus only on the existing channel state information, ignoring the congestion-based factors affecting the network throughput . Such a probe-based approaches result in suboptimal performance, because of the inefficient technique to constantly tune the contention parameters due to the changing channel conditions. A network is said to be congested when the offered traffic load surpasses the available resource capacity. Such a condition leads to increased buffer space demand in intermediate nodes over the data path, which in-turn increases the packet drop ratio due to limited resource availability. This results in retransmission of packets, which further degrades the network throughput. To address these issues, we extend the cross layer approach in the proposed MOCLO by implementing an efficient link rate adaptation algorithm with a methodology based on traffic congestion to decide the most suitable transmission rate. In the proposed algorithm, the congestion status of a node is indicated by *Alertlevel*. Let $Q_{avgloadAC(i)}$ be the average queue load of the i^{th} AC queue and Q_{size} is the maximum size of the queue. In *Alertlevel*, when $Q_{avgloadAC(i)} \geq Q_{thresh}$, the load balancing procedure described in Section 3.3.8. The Q_{Thresh} is defined in the range of $75\% - 80\%$ of Q_{size}. When a source in the route has some data packet to send to some destination node, it finds a valid path in the route table. On finding the appropriate path, the node forwards the data packets to the next hop as per the rate DR_i at which last packet was routed. Nodes adapt to various data rates dynamically.

5.4 Simulation Results

The performance of MOCLO protocol is analyzed using the discrete event network simulator NS2.35 [96] which supports complete physical and enhanced MAC layer models for simulating wireless ad hoc networks over 802.11e. A network consisting of 50 mobile nodes in an area of $1000 * 1000$ m is simulated. The nodes move with an average speed varying from 5 m/s to 10 m/s with a pause time of 10 s. In the simulation, for modeling the node movement, the random waypoint model is used. The network is flooded with a mixture of best effort TCP traffic and delay-sensitive video traffic. TCP traffic is generated by simulating an FTP server client environment, where TCP connections are established between the server and the client when the FTP clients upload/download a file to/from the FTP server. A 30 KB file is assumed to be transferred every 20 seconds for an FTP pair of connections. An MPEG-4 video codec [130] with a data rate 250 KBytes is used to simulate the video traffic and is played at 30 frames per second and the connection is randomly switched among the nodes during the simulation. The IEEE 802.11e MAC layer with different queue priorities has also been implemented so as to map the packet at MAC layer to network layer queue with different priorities. The length of each AC queue is set to 50 packets. The starting time of all applications are set at a random time interval of 10 s - 40 s and continues till the end of the simulation time in order to create high traffic load conditions. The simulation is executed for 20 runs to average the value using different seed values. Table 3.1 shows the physical layer parameters set for the simulation.

5.4.1 Performance Evaluation

The performance of the proposed MOCLO protocol is compared with the system that uses QoS-AOMDV [79] routing protocol at the network layer and IEEE 802.11e EDCA at the MAC layer in MANETs. The QoS-AOMDV is a cross layer based multipath routing protocol, which focuses on energy conservation and load balance. It constructs the weighted route, based on metrics namely, residual energy and queue length of a node. In QoS-AOMDV, data transmission is split among the available paths which are discovered during route discovery phase based on the energy balance among the nodes.

In the simulation, two networks are compared and evaluated by varying following two parameters:

1. Connection time, which is expressed as a percentage of the total time for one complete video connection.
2. Traffic load by varying the number of video stream connections from 5 to 10.

5.4.1.1 Effect of Varying the Video Connection Time

The multimedia video connection tends to take longer time for data transfer even with high priority access to the resources along the route, whereas in low priority TCP connection, a file takes few seconds. Therefore, to evaluate the effect of video connection on the performance of two protocols, results are analyzed at regular percentage of connection time. As the focus is on QoS support for real time video streaming application, a video stream is added with a single pair of the connection, which lasts for 50 s. After 50 s, the connection is randomly switched between different pairs of source and destination nodes. Ten different TCP connections are injected, in which 30 KBytes file is transferred between source and destination pair within every 20 s.

Figure 5.6 shows the plot of average video packet delay v/s video traffic connection time for proposed MOCLO and QoS-AOMDV+EDCA protocols. It is observed from the figure that, as the connection time progresses, the delay curve increases for both the protocols. This is mainly because of more collisions due to increase in load which in turn increases the exchange of control information between the nodes further worsening congestion at node queues. It can be observed from the figure that, the MOCLO protocol shows better performance compared to QoS-AOMDV+EDCA protocol. This is because, MOCLO protocol controls the congestion by adjusting the data rate of the flow based on QoS metric queue load size and available bandwidth at the individual node. And also, it maps the high priority and low priority traffic with different available efficient routes thereby reducing the load on single efficient path.

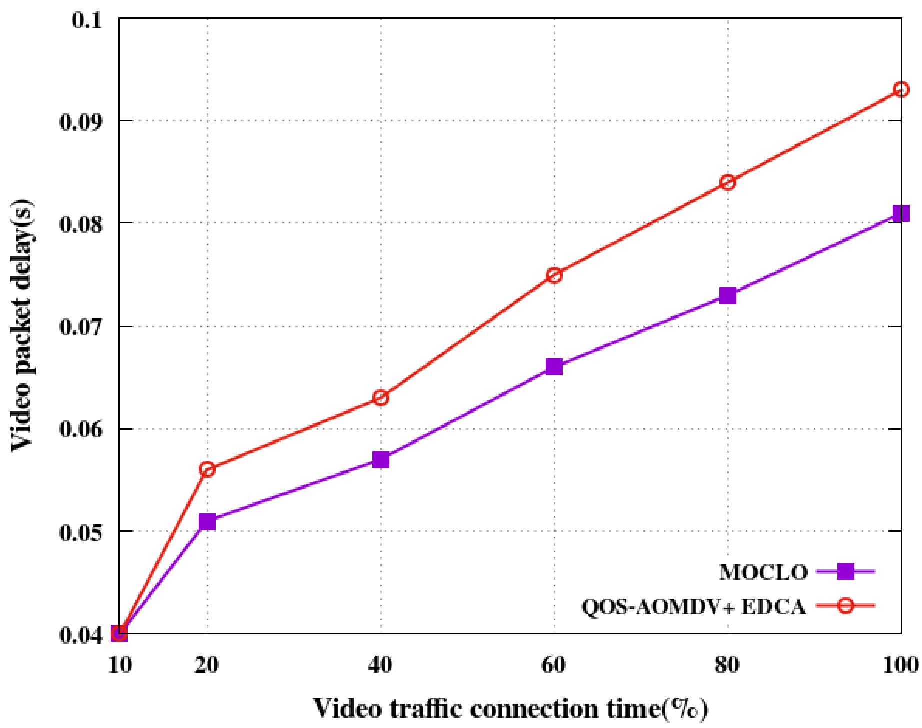


FIGURE 5.6: Video packet delay v/s Video connection time

Figure 5.7 shows the plot of average TCP packet delay versus video traffic connection time for MOCLO and QoS-AOMDV+EDCA protocol. It is observed from the figure that, when video connection is inactive, the TCP packet delay is very low. But as the video transfer connection progresses, being the highest priority traffic, gets, the higher preference to channel access and bandwidth at each node. This results in starvation for TCP packets leading to retransmissions. It is observed from the figure that MOCLO protocol shows better delay performance compared to QoS-AOMDV + EDCA. In MOCLO protocol, the traffic is split and mapped to a different available efficient path which reduces the load on single efficient path. Also, since within AC(1), TCP ACK packets are given higher priority, there is less delay in acknowledgement of data packets, which further reduces the end to end delay for data packets.

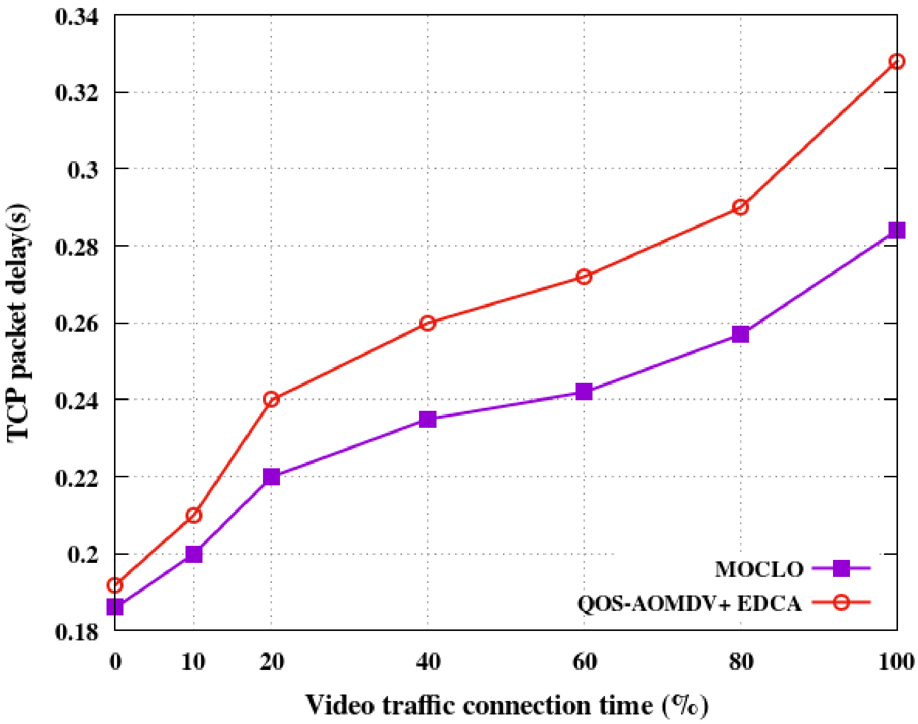


FIGURE 5.7: TCP packet delay v/s Video connection time

To evaluate QoS provisioning, throughput analysis for video packets and TCP packets is carried out and results obtained are plotted as shown in Figure 5.8 and 5.9. In general, the throughput is mainly affected by the increase in traffic load due to excessive channel contention and retransmissions. The throughput of TCP traffic degrades in the presence of high priority video traffic. In both the protocols, when the high priority video connection is completely absent, the TCP traffic shows better throughput. As the video connection progresses, frequent collisions increases and the TCP packets are deprived of channel access. This results in retransmission of packets and hence reducing the throughput of the TCP packets. The proposed MOCLO protocol achieves better performance as it provides significantly higher throughput for video and TCP traffic compared to QoS-AOMDV+EDCA protocol. Apart from having the advantage of splitting the traffic in multiple paths in the proposed protocol, a congestion notification is also provided as the feedback to the packet forwarding node. This helps node to control the traffic flow by adjusting its data rate and eventually reducing packet losses and retransmissions which is evident in Figure 5.10. While in QoS-AOMDV+EDCA, there is no mechanism to tackle the delayed ACK packet, and hence shows a poor performance.

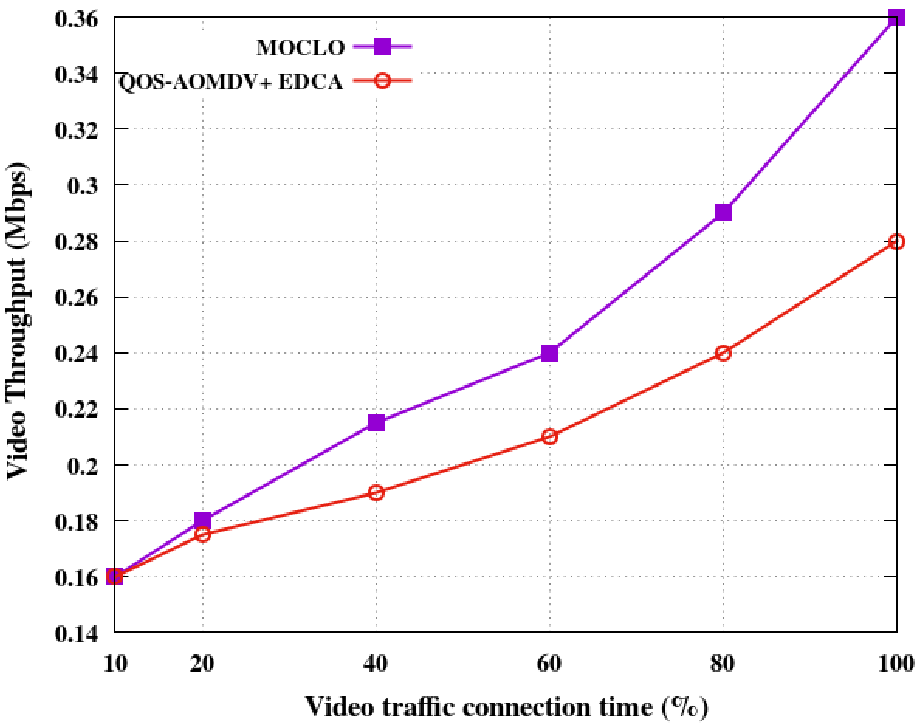


FIGURE 5.8: Video throughput v/s Video connection time

To analyze the effect of priority within an AC, the frame loss v/s video traffic connection time is plotted which is as shown in Figure 5.11. From the figure, it can be observed that, MOCLO protocol protect the important frames, namely I and P frames by sending them through the best available path using high priority AC. The AC is implemented using priority queue data structure to dequeue the packet faster based on higher priority. Till the 40% of connection time, the frame loss in MOCLO is almost equal to zero, whereas by the end of the connection time, frame loss is less than 10%. In QoS-AOMDV+EDCA, the frame loss is very high as there is no mechanism to protect the important frames.

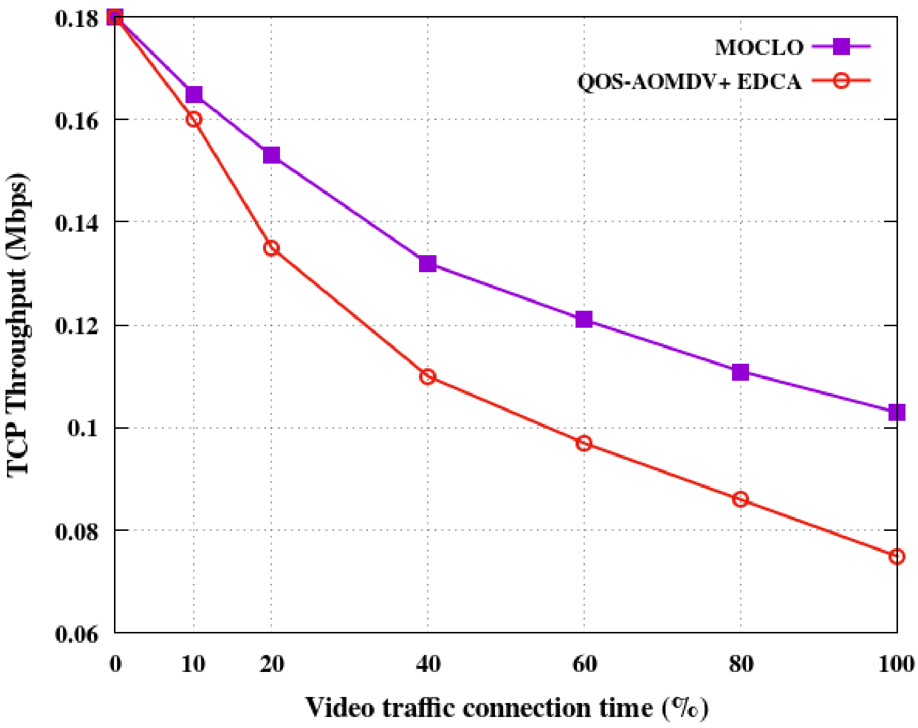


FIGURE 5.9: TCP throughput v/s Video connection time

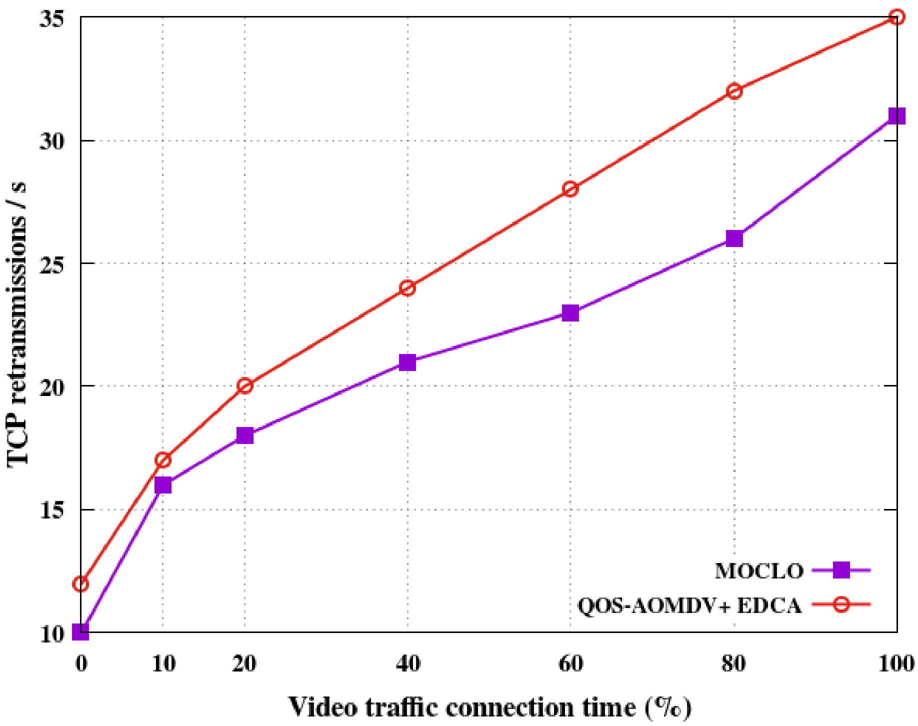


FIGURE 5.10: TCP retransmission v/s Video connection time

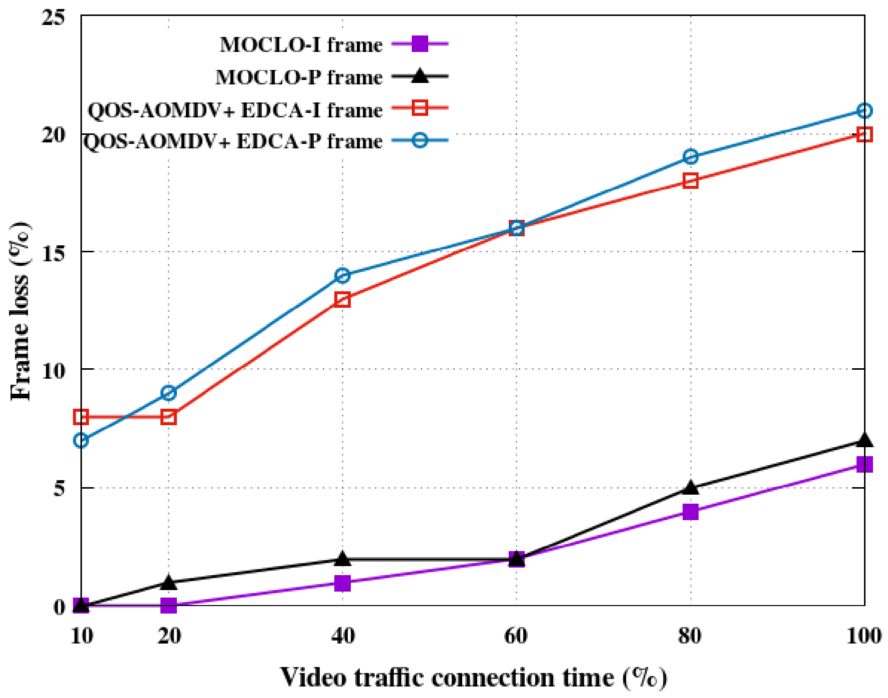


FIGURE 5.11: Frame loss percentage v/s Video connection time

5.4.1.2 Effect of Varying Video Traffic Flows

The proposed protocol is also evaluated by varying the number of concurrent video stream connections. The network traffic load is generated by injecting 5 TCP connections and video stream connections varied from 5 to 10 connections. These connections are randomly switched among the nodes during the simulation.

Figure 5.12 to Figure 5.15 illustrates delay and throughput performance of video and TCP packets in MOCLO and QoS-AOMDV+EDCA protocol under varying traffic load. It is evident from all the figures that MOCLO protocol outperforms the QoS-AOMDV+EDCA Protocol. This is mainly due to the route selection procedure adopted, priorities within the AC's and rate adaptation in the proposed protocol. Rate adaptation mechanism in MOCLO protocol allows packets to be transmitted at higher data rates whenever feasible. Figure 5.16 shows the percentage of I and P frames as a function of traffic load variation. It is observed from the figure that frame loss percentage for both protocols increases with the increase in video traffic load. The difference in loss percentage of two protocols, performance curve gets wider as the connection increases. The MOCLO outperforms the compared protocol because, it

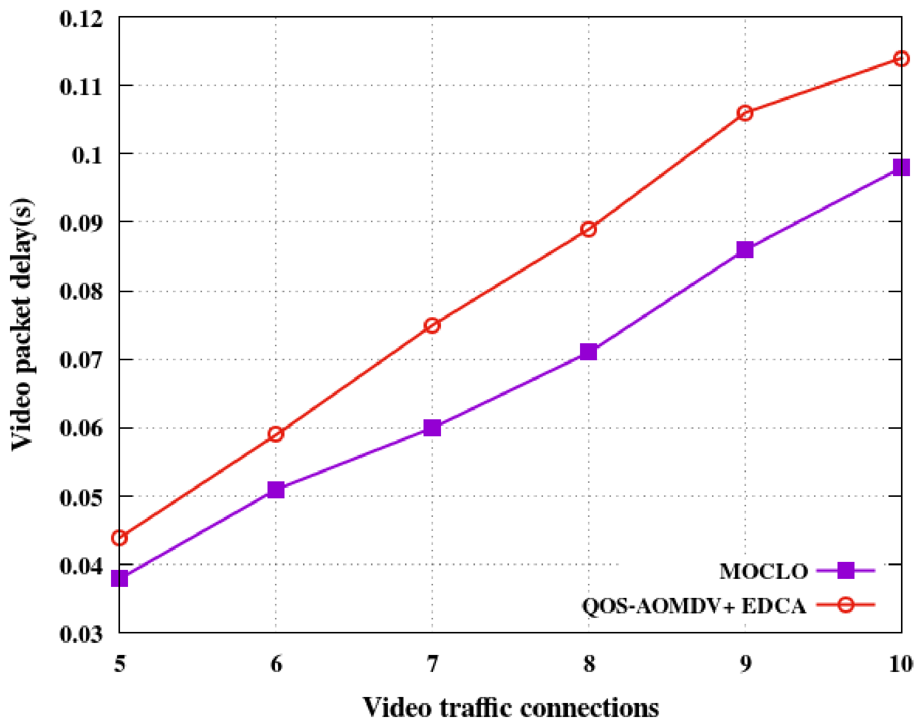


FIGURE 5.12: Video packet delay v/s Video connections

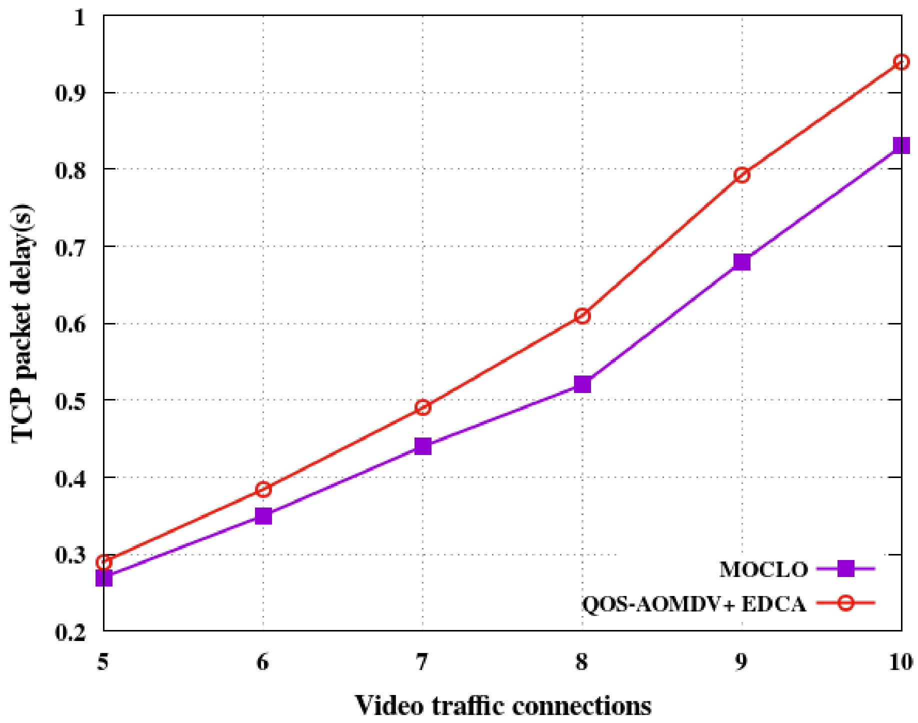


FIGURE 5.13: TCP packet delay v/s Video connections

provides high priority for I and P frames in AC and transmits on a more stable and less congested path. Thus the packet loss is minimized.

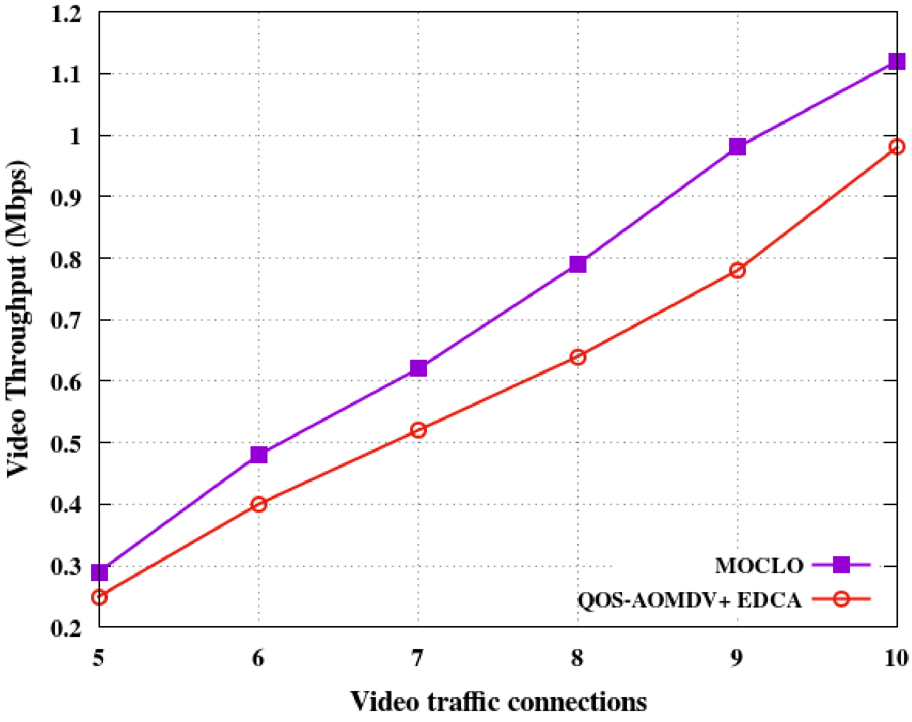


FIGURE 5.14: Video throughput v/s Video connections

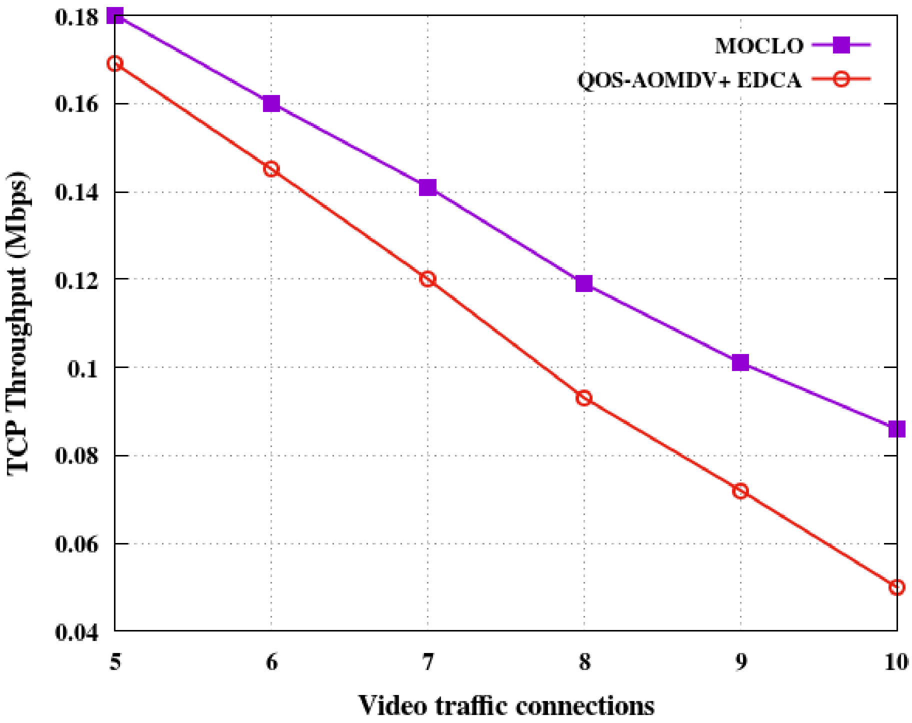


FIGURE 5.15: TCP throughput v/s Video connections

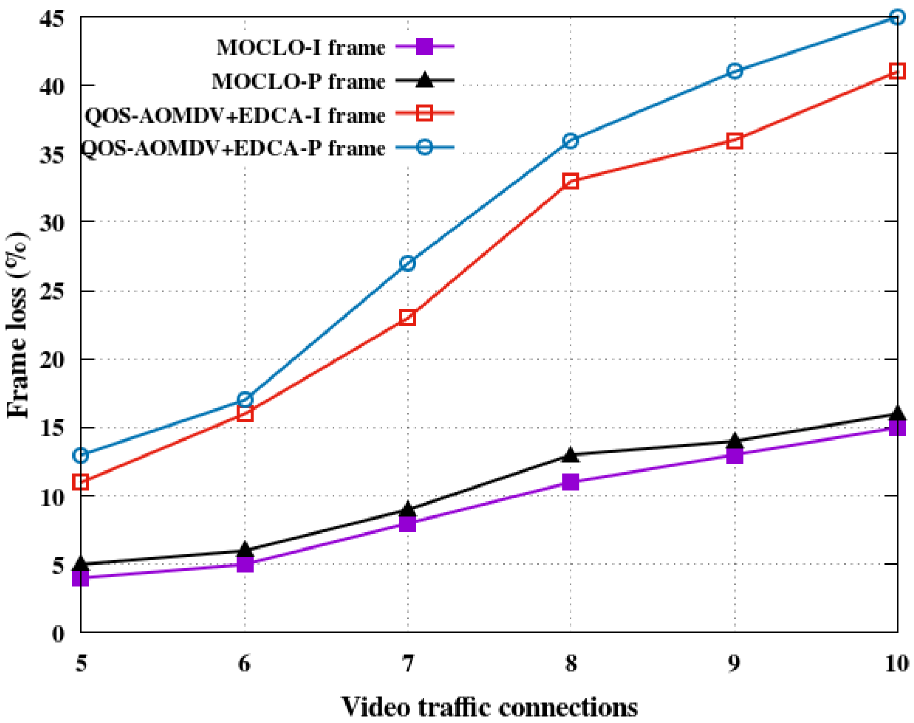


FIGURE 5.16: Frame loss percentage v/s Video connections

5.5 Conclusion

In this chapter, MOCLO protocol is presented and analyzed with reference to QoS-AOMDV + EDCA protocol. The on demand and congestion aware MOCLO protocol is based on cross-layer approach for QoS provisioning in MANETs. It implements priority scheduler for each AC to optimize the transmission of important packets in video and the best effort TCP trac. Being a congestion aware protocol, it discovers stable node disjoint paths and to reduce the load on the single path, traffic is split into multiple paths. Multirate functionality of the MOCLO protocol helps to send the packets at higher data rates based on queue load of a node. In order to evaluate the performance of the proposed protocol, simulations were conducted using network simulator NS-2.35 Through extensive simulations, it is observed that the proposed MOCLO protocol outperforms QoS-AOMDV+EDCA protocol in terms of different metrics namely packet delay, throughput and important packet loss under different conditions such as varying traffic connections and traffic connection time.

Chapter 6

Adaptive Multi QoS Cross Layer Cooperative Routing in MANETs

6.1 Introduction

The growing demand for higher data rates in wireless data transmission has enforced serious challenges on network design. In some of the applications like battlefield communications, emergency operations, search and rescue, and disaster relief operations, the desired high data rate coverage cannot be achieved through the direct transmission. These applications have to meet certain QoS levels in multiple domains, such as, end-to-end delay, packet delivery ratio, reliability, and so on. However, due to channel fading, dynamic network topology due to node mobility, and severe constraints on energy, achieving multiple QoS requirements at the same time in MANETs is a challenging task. This motivated the innovation of a new technology, known as cooperative communication [31].

Recently, cooperative communication has received considerable interest of the research fraternity in wireless networks. The basic idea of the cooperative communication is to allow the nodes in the network to cooperate with each other in data communication [31]. The cooperative communication has provided higher reliability since the probability of all the channel links to the destination to go down is very less. Cooperative communication improves the system performance, typically in terms of increased capacity, resource utilization, improved transmission reliability, throughput and spatial diversity [131][132].

In wireless systems, physical layer multirate adaptation techniques are widely supported to adapt to different channel conditions. These techniques make a compromise between the data rate and the range. The long range data communication must operate at low data rates to minimize error rate. In MANETs, when a node transmitting at higher rate moves away from the destination may result in increased error rate [133].

Usage of cooperative communication together with rate adaptation techniques can enable the nodes to adapt their data rates to match the channel conditions and node mobility. Combining both these techniques can provide a substantial throughput improvement in direct and conventional multihop network. Hence, this chapter proposes an Adaptive Multi-QoS Cross layer Cooperative Routing protocol (AMCCR) that enhances the performance while guaranteeing multiple QoS requirements for MANETs. The protocol uses cross layer communication between PHY, MAC, and network layers. The proposed protocol uses an energy aware, end-to-end routing scheme to optimize the trade-off between end-to-end delay, and energy of the system. AMCCR uses an adaptive mechanism to select the transmission mode, i.e., direct or relay mode so that a low data rate host can be assisted by high data rate relay nodes. AMCCR protocol also supports dual-hop half-duplex communication via selected relay node by coding technique. The proposed algorithm is validated by extensive simulations, and compared with CD-MAC [134] protocol, CODE protocol [135] and traditional non-cooperative 802.11 DCF.

6.2 A Literature Survey

To cope up with varying channel conditions, rate adaptation techniques are widely used to dynamically adjust the modulation technique and the data rate. IEEE 802.11b provides four data rates, 1, 2, 5.5 and 11 Mbps whereas IEEE 802.11g provides eight data rates, 6, 9, 12, 18, 24, 36, 48, and 54 Mbps using different modulation schemes.

The rate adaptation can be either sender-based or receiver-based. In sender based technique, based on the channel conditions, the sender node uses the history of previous transmissions to determine the next transmission rate. In receiver based technique, the transmission rate of the sender is decided by the receiver [122][19]. To estimate the channel condition more accurately, the receiver node measures the instantaneous channel conditions by sensing the signal strengths of the control frames during each transmission and sets the transmission data rates accordingly for sender node. However, these techniques are not very effective when the channel condition is poor, since only low data rates can be supported and the mobility aspect of MANETs makes the problem worse. Therefore, multirate capability with cooperative communication is a better solution.

The cooperative communication exploit the broadcast nature of the wireless medium by transforming single-antenna terminals into a virtual antenna array [136]. Thus, multiple signals are transmitted from source and intermediate forwarding (relays) nodes through uncorrelated channels to the destination and provide benefits of spatial diversity. Due to signal processing at each intermediate relay node, this technique achieves robustness against channel variations due to fading. Using cooperative communication, nodes in a wireless network can cooperate with each other to send data to the destination thereby achieving higher reliability and increased system throughput through resource aggregation. An example of single-relay cooperative communication network is shown in Figure 6.1. When a source node S transmits a signal to a destination node D through its direct path (S-D), other nodes, which are in the communication range of S can overhear the signal. If a node say, R_1 is in the range of S and is in cooperative mode, then it can act like a relay node and forward the signal received from S to D. Thus, D receives two signals: the original one transmitted from S through the direct path (S-D) and second one forwarded by R through the relayed path (S-R_1-D). The cooperative communication in a multi-hop network is as shown in Figure 6.2 where, I is the intermediate node between source S and destination D. The cooperation is achieved through set of relay nodes along the data transmission route. In order to facilitate cooperative communication, there are certain important issues needs to be addressed. These issues include making decision about the cooperative transmission scheme, relay node selection, resource allocation, channel state information, and the cooperative routing metrics.

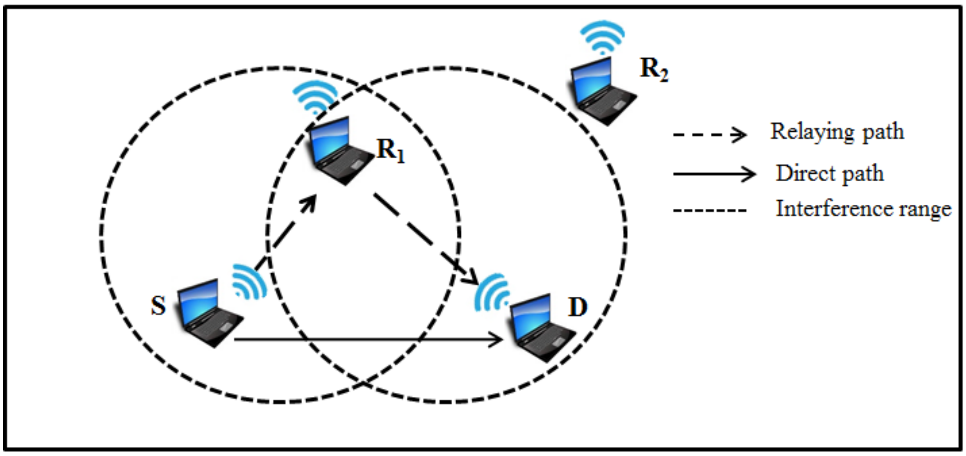


FIGURE 6.1: Single relay cooperative communication network

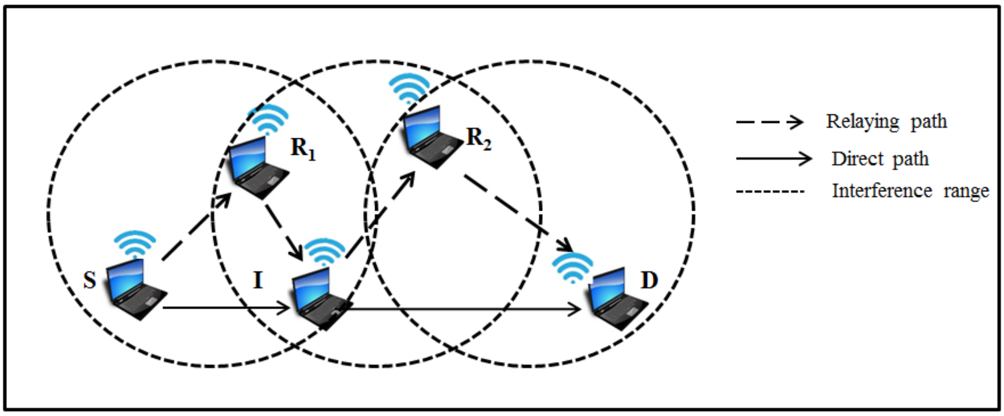


FIGURE 6.2: Multi-hop cooperative communication network

Initially, the idea of cooperative communication was mainly concentrated on the physical layer with the advances in the techniques such as modulation, coding, etc. [137]. The cooperative strategy at the physical layer include selecting cooperative relaying schemes, choosing power for signal transmission and scheme for selecting appropriate relay nodes. Store and forward(SF) [138], amplify-and-forward(AF) [139], decode-and-forward (DF) [140] are the most commonly used relaying schemes. The innovation of cooperative communication is not restricted only to the physical layer.

To subjugate the issue of node mobility, it would be ideal to expose the physical layer information for the cooperation decision at higher layers of protocol stack. In the recent time, the cooperative MAC scheme in ad hoc wireless networks has also attracted much attention [141]. These schemes use handshaking techniques to reserve the channel and avoid the collision problems. These schemes exchange the information between the physical and MAC layers for relay selection with the criteria related to rate adaptation and power control. Due to random nature of the wireless channel, MAC layer should know when to initiate the cooperative transmission. Thus, the cooperative MAC should utilize the assistance from a relay node to forward the data using better link adaptation techniques that supports higher data rates to enhance the network throughput.

Sai Shankar N et al. proposed a protocol called Cooperative communication MAC (CMAC) [32] with a minor modifications to the standard IEEE 802.11 DCF. When the source node sends the data packets and if destination node receives it with errors or fails to receive, then the source node selects a cooperative relay node to forward the lost data packet. Though CMAC provides the reliability of data transmission and throughput enhancement, it assumes that the link between any nodes is ideal and error-free.

To address the issue of low throughput caused by low-data-rate nodes, Pei Liu et al. proposed a MAC protocol for WLAN known as CoopMAC protocol [33] in which low data rate nodes are supported by high data rate nodes through two-hop transmission. Table driven based CoopMAC protocol maintains a table, called a CoopTable, which is updated based on the information gathered using the overheard transmissions. The protocols choose to send packets at a higher data rate, using relay node in a two-hop manner instead of a low data rate direct transmission. In CoopMAC, the relay node itself decides whether to cooperate or not, based on its local information maintained in the cooperative table. This protocol is not suitable for multihop ad hoc networks as it does not have any provision to deal with hidden node and exposed node problem.

Niraj Agarwal et al. proposed cooperative MAC called UtdMAC [142], in which the data packets are transmitted through the relay whenever a direct transmission fails due to fading. In UtdMAC, it is assumed that the relay is selected in prior and will be ready to transmit in a cooperative method whenever necessary. Thus this protocol does not have to deal with much of the relay selection overhead and management. The protocol designed is suitable for fixed networks.

Sangman Moh et al. proposed Cooperative Diversity MAC (CD-MAC) [134] protocol, which is based on the IEEE 802.11 DCF. In CD-MAC, the transmission of multiple copies of a data stream is distributed among the cooperating nodes, which act as a virtual antenna array. The cooperating nodes encode the data by using orthogonal codes and simultaneously transmit it to the destination. The packet scheduling technique of CD-MAC is similar to that of ARQ scheme, where cooperation is triggered when a direct transmission of a control packet fails. A source node sends an RTS packet to the destination node. If the destination replies with a CTS before the timeout period expires, then CD-MAC does not initiate cooperation. Otherwise, it activates the cooperation in the next phase. The source node intimates the need for cooperation to relay nodes through repeated Cooperative RTS packet called C-RTS. The destination node replies to a C-RTS, with C-CTS simultaneously to relay and source nodes. During cooperation using distributed space-time coding technique, source node sends a packet to the relay node in the first phase, During second phase, both the source and relay nodes simultaneously transmit the packet to the destination node. In CD-MAC, each node maintains an estimate of neighbor nodes link quality through periodic broadcast of Hello packets. However, due to repeated transmission of control packets, the nodes suffer from energy depletion and increase in end-to-end latency. Also, the relay node unavailability is not addressed properly.

Recently, there has been increased interest in protocols for MANETs to exploit the significant interactions between various layers of the protocol stack for the performance enhancements [143]. The state information available at the PHY and MAC layers can be utilized at higher layers, in particular, the routing layer to realize a fully cooperative network [144]. However, the problem of combining routing with cooperative diversity, has received very little attention.

The cooperative routing takes the advantage of the broadcast nature of wireless communication to transmit the message in each hop through relay nodes as well. Therefore, cooperative routing mechanism allows multiple nodes along a path to coordinate together to transmit a message to the next hop as long as the combined signal at the receiver node satisfies the given Signal-to-Noise Ratio (SNR) threshold value. The receiver node at each hop selects the best of received signals (direct or relayed), or aggregates them to produce a stronger signal.

H. Xu, et al. proposed Cooperative Multipath Routing protocol (CMPR) [145] which constructs an energy-efficient node-disjoint cooperative multi-path routing while satisfying the bandwidth constraint on each path. This algorithm consists of two steps: multi-path route construction and cooperative relay node assignment. The first step includes calculating a cost function of each link, based on the routing objective under the direct (or non-cooperative) transmission path. In the second step, k minimum cost node disjoint paths are constructed from the source to the destination node. The CMPR algorithm utilizes a method based on the dynamic programming for relay node assignment on each path. To select the best relay node of a link, each node in the network will be assigned a weight, which represents the amount of performance achievement if that particular node acts as a cooperative relay node for that specific link in the constructed path.

A. M. Akhtar et al. proposed Power Aware Cooperative Routing (PACR) protocol [146] which is a distributed cooperative routing that takes the active radio electronic power into account while constructing the route that consumes minimum-power to transmit a message from a source to its destination node. PACR is mainly proposed for cooperative routing in cellular wireless mesh networks. In this algorithm, a source node, while selecting its neighboring nodes, uses their location information to calculate an optimal transmission distance for cooperative communication. PACR assumes all the eligible relay nodes are at equal distance from destination node. Hence, equal power is allocated to each node in a cooperative group, therefore, it is not a optimal cooperative routing algorithm.

S. Chen et al. proposed Energy-Balanced Cooperative Routing (EBCR) protocol [147] that performs cooperative communication for each hop along the path to maximize the network lifetime. Initially, the route is found by employing non-cooperative routing strategy, such as energy-efficient ad hoc routing protocol or shortest path based routing algorithms. In order to apply cooperative communication at each link, a set of relay for each transmitter node along the path is selected. A potential relay set is defined as any neighbors of the transmitter node that are closer to the transmitter than the receiver node. Given a potential relay set with size k, the remaining lifetime of the transmitter node and its relay set is calculated. Then, all of these k + 1 nodes are sorted in a descending order by their remaining lifetime. Finally, the algorithm greedily picks those nodes with a larger remaining lifetime and uses them as the relay nodes until the cumulative signal strength at the receiver node is greater than or equal to the detection threshold, i.e., until the received signal strength meets the detection requirement. In this algorithm a central controller is required to collect the potential relay set and pick the best relay node to prolong the network lifetime.

In the cooperative routing schemes discussed above, the primary objective is focused on energy efficiency [148], bandwidth efficiency [149], enhanced throughput [150], outage probability [151], etc. Most of the cooperative routing algorithms designed are application specific as they satisfy one or two QoS parameters. Hence, it is difficult to make comprehensive cooperative routing protocol satisfying needs of all applications. Further, significant effort has not been done in QoS provisioning for cooperative routing in wireless networks, especially in the context of achieving multiobjective QoS services. The existing cooperation techniques mentioned above, do not consider cross layer coordination between physical layer, MAC and network layer.

Hence, this chapter proposes an AMCCR protocol which is an energy efficient distributed cross layer cooperative routing protocol for MANETs. The protocol uses potential cooperation gain to find the optimal route that satisfies multi-QoS cooperative metrics. The cross layer coordination between the MAC and network layer is used to select an optimal next hop, while the cross layer coordination between the MAC and physical layer is used to select the best relay.

The main features of the proposed protocol are as follows:

1. An end-to-end delay efficient, energy-aware routing protocol is proposed in order to increase the network lifetime.

2. The proposed protocol exploit the cross layer approach at the PHY, MAC and routing layer for rate adaptation, relay selection, route discovery and resource allocations.

3. Using the channel state information at the physical layer, the proposed algorithm determines, whether cooperation on the link is necessary or not. In case the cooperation on the link is necessary, multi QoS metric is used to determine the potential relay nodes for a cooperative transmission over each link. The cooperative mode activation is implemented at the MAC layer.

4. The proposed protocol exploit the cross layer approach at the routing layer for selecting the next hop node, based on the relay selected at the MAC layer that reflects the potential cooperation gain to find the optimal paths. The best relay node selection strategy is executed in distributed manner.

5. The proposed algorithm is analyzed with single relay participation for cooperative scheme, and further extended for supporting network coding scheme.

6. The best relay selection is enabled with collision avoidance mechanism.

6.3 AMCCR Cross Layer Architecture

The AMCCR cross layer scheme uses an energy efficient QoS routing and an adaptive cooperative MAC to meet the requirements of the application. During the routing phase, multiple paths are discovered between source and destination nodes. The adaptive cooperative MAC is used to identify the candidate relay node satisfying certain QoS metrics, to cooperate the data transmission between two intermediate hops in the paths discovered during the routing phase.

6.3.1 System Model

Consider a mobile wireless ad hoc network comprising of nodes V_i, V_{i+1},..., V_n. These nodes are randomly distributed and transmit signal to their neighbor node either directly, or using a single relay node. The system design does not consider more than one relay node, just to keep the complexity low. IEEE 802.11g is considered at the PHY layer, which uses different modulation techniques to support multiple data rates of 6, 9, 12, 18, 24, 36, 48, and 54 Mbps. The PHY and MAC layer headers as well as control packets are transmitted at a fixed rate of 1 Mbps.

The slow Rayleigh fading channel model [152] is considered, as the analysis focuses on the case of slow fading to capture scenarios in which delay constraints are on the order of the channel coherence time. A half duplex constraint is imposed on all the relay nodes, i.e., relay nodes cannot transmit and listen simultaneously. The packet transmissions are multiplexed in time. The channel is assumed to be symmetric, and therefore, the communication in either direction experiences same channel fading. Let R_i represents the relay node in the i^{th} hop, thus the set of n relays is denoted by R_i, R_{i+1}, ...,R_n. Let $h_{V_i,V_{i+1}}$, h_{V_i,R_i}, $h_{R_i,V_{i+1}}$ represents the channel gain from the node V_i to the next hop V_{i+1}, node V_i to relay R_i and relay R_i to node V_{i+1} respectively. The noise $\eta_{V_i,V_{i+1}}$, η_{V_i,R_i}, and $\eta_{R_i,V_{i+1}}$ are modeled as zero mean complex Gaussian random variables with variance N_0.

We model the cooperation with signal combining approach in two phases. In Phase 1, the source node V_i broadcasts its information to the next hop V_{i+1}, which is overheard by neighbor R_i nodes. The signal received at the next hop and the neighbor relay node are represented as $y_{V_i,V_{i+1}}$ and y_{V_i,R_i} respectively and is given by [131],

$$y_{V_i,V_{i+1}} = \sqrt{P1}h_{V_i,V_{i+1}}x + \eta_{V_i,V_{i+1}} \tag{6.1}$$

$$y_{V_i,R_i} = \sqrt{P1}h_{V_i,R_i}x + \eta_{V_i,R_i} \tag{6.2}$$

where,

$P1$ is the power transmitted at the node V_i,

x represents the transmitted information symbol.

In phase 2, the neighbor node R_i, processes the received signal either by Amplify-and-Forward (AF) or Decode-and-Forward (DF) techniques which is described below.

6.3.1.1 Adaptive Cooperation with Amplify-and-Forward (AF)

In phase 2, under the AF technique implementation, the selected relay terminal simply amplifies the received signal and forwards it to the next hop with transmission power $P2$. As stated in [131], under AF, the achievable rate C_{AF} between V_i and V_{i+1} with R_i as relay node is given as,

$$C_{AF} = WI_{AF}(V_i, V_{i+1}, R_i,) \tag{6.3}$$

where, W represents the bandwidth of channels at node V_i and relay R_i, and the average mutual information $I_{AF}(V_i, V_{i+1}, R_i)$ between the input and the outputs achieved by independent and identically distributed (i.i.d) complex Gaussian inputs and is given as [153],

$$I_{AF}(V_i, V_{i+1}, R_i) = log_2\left(1 + SNR_{V_i,V_{i+1}} + \frac{SNR_{V_i,R_i} * SNR_{R_i,V_{i+1}}}{SNR_{V_i,R_i} + SNR_{R_i,V_{i+1}} + 1}\right) \tag{6.4}$$

where,

$$SNR_{V_i,V_{i+1}} = \frac{P1}{\eta_{V_i,V_{i+1}}} \left|h_{V_i,V_{i+1}}\right|^2 \tag{6.5}$$

$$SNR_{V_i,R_i} = \frac{P1}{\eta_{V_i,R_i}} \left|h_{V_i,R_i}\right|^2 \tag{6.6}$$

$$SNR_{R_i,V_{i+1}} = \frac{P2}{\eta_{R_i,V_{i+1}}} \left|h_{R_i,V_{i+1}}\right|^2 \tag{6.7}$$

The signal received at the next hop V_{i+1} sent by R_i is represented as,

$$y_{R_i,V_{i+1}} = \frac{\sqrt{P1P2}}{\sqrt{P1\left|h_{V_i,R_i}\right|^2 + N_0}} h_{R_i,V_{i+1}} h_{V_i,R_i} x + \eta'_{R_i,V_{i+1}} \tag{6.8}$$

where,

$$\eta'_{R_i,V_{i+1}} = \frac{\sqrt{P2}}{\sqrt{P1\left|h_{V_i,R_i}\right|^2 + N_0}} h_{R_i,V_{i+1}} \eta_{V_i,R_i} + \eta_{R_i,V_{i+1}} \tag{6.9}$$

The η_{V_i,R_i} and $\eta_{R_i,V_{i+1}}$ are assumed to be independent. Thus, $\eta'_{R_i,V_{i+1}}$ is modeled as a zero-mean complex Gaussian random variable with a variance

$$\left(\frac{P2\left|h_{R_i,V_{i+1}}\right|^2}{P1\left|h_{V_i,R_i}\right|^2 + N_0} + 1\right)N_0$$

The next hop node, on receiving the signal from the relay and previous hop node, detects the symbols transmitted with the knowledge of the channel gains $h_{V_i,V_{i+1}}$ and $h_{R_i,V_{i+1}}$ respectively. The next hop node implements a Maximum Ratio Combiner (MRC) technique [154] to decode the signals received from the previous hop node and the relay node. The MRC output at the next hop is given as,

$$y = a_1 y_{V_i,V_{i+1}} + a_2 y_{R_i,V_{i+1}} \tag{6.10}$$

where, a_1 and a_2 is given as,

$$a_1 = \frac{\sqrt{P1}h^*_{V_i,V_{i+1}}}{N_0}$$

and

$$a_2 = \frac{\sqrt{\frac{P1P2}{P1\left|h_{V_i,R_i}\right|^2+N_0}}h^*_{V_i,R_i}h^*_{R_i,V_{i+1}}}{\left(\frac{P2\left|h_{R_i,V_{i+1}}\right|^2}{P1\left|h_{V_i,R_i}\right|^2+N_0} + 1\right)N_0} \tag{6.11}$$

where,

h^* is the conjugated channel gain corresponding to the received symbol.

Assuming the average energy of transmitted symbol x in Equation 6.1 and Equation 6.16 is 1, the SNR of MRC output is represented as,

$$\gamma = \gamma_1 + \gamma_2 \tag{6.12}$$

where γ_1 and γ_2 are given as,

$$\gamma_1 = \frac{P1\left|h_{V_i,V_{i+1}}\right|^2}{N_0}$$
$$\text{and} \tag{6.13}$$
$$\gamma_2 = \frac{1}{N_0}\frac{P1P2\left|h_{V_i,R_i}\right|^2\left|h_{R_i,V_{i+1}}\right|^2}{P1\left|h_{V_i,R_i}\right|^2 + P2\left|h_{R_i,V_{i+1}}\right|^2 + N_0}$$

6.3.1.2 Adaptive Cooperation with Decode-and-Forward (DF)

In phase 2, for DF technique implementation, if the selected relay terminal is able to decode the symbols of the received information from the forwarding node V_i correctly, then it re-transmits the information with power P2 to the next hop. We assume, if the SNR received at the relay is greater than the threshold, then the symbol will be correctly decoded. As stated in [131], under DF, the achievable rate C_{DF} between V_i, V_{i+1} via R_i is given as,

$$C_{DF} = WI_{DF}(V_i, V_{i+1}, R_i) \tag{6.14}$$

where, the average mutual information $I_{AF}(V_i, V_{i+1}, R_i)$ between the input and the outputs achieved by i.i.d. complex Gaussian inputs, is given by,

$$I_{DF}(V_i, V_{i+1}, R_i) = min\left\{log_2(1 + SNR_{V_i,R_i}), log_2(1 + SNR_{V_i,V_{i+1}} + SNR_{R_i,V_{i+1}})\right\} \tag{6.15}$$

The signal received at the next hop V_{i+1} node sent by R_i is given as,

$$y_{R_i,V_{i+1}} = \sqrt{P2}h_{R_i,V_{i+1}}x + \eta_{R_i,V_{i+1}} \tag{6.16}$$

The channel gains $h_{V_i,V_{i+1}}$, $h_{R_i,V_{i+1}}$, h_{V_i,R_i} are assumed to be known at the receiver, but not at the transmitter and are assumed to be independent of each other. The next hop node V_{i+1}, on receiving the signal from the relay and previous forwarding node, detects the symbols transmitted with the knowledge of the channel gains $h_{V_i,V_{i+1}}$ and $h_{R_i,V_{i+1}}$ respectively. The next hop implements a MRC. The MRC maximizes the SNR at the receiver so that the bit error rate is minimized. The combined signal at the next hop is represented as,

$$y = a_1 y_{V_i,V_{i+1}} + a_2 y_{R_i,V_{i+1}} \tag{6.17}$$

where a_1 and a_2 are computed to maximize the SNR of MRC output at the next hop node, and are represented as,

$$a_1 = \frac{\sqrt{P1} h_{V_i,V_{i+1}}}{N_0}, \text{ and } a_2 = \frac{\sqrt{P2} h_{R_i,V_{i+1}}}{N_0} \tag{6.18}$$

Assuming the average energy of transmitted symbol x in Equation 6.1 and Equation 6.16 is 1, the SNR of MRC output is represented as,

$$\gamma = \frac{P1 \left|h_{V_i,V_{i+1}}\right|^2 + P2 \left|h_{R_i,V_{i+1}}\right|^2}{N_0} \tag{6.19}$$

6.3.2 Energy Aware QoS Routing

The network lifetime is directly related to energy efficiency. The energy efficiency can be measured by the duration of the time over which the network can maintain a certain performance level. Therefore, the energy of the node is a crucial QoS factor while designing the routing algorithm.

AMCCR proposes an energy efficient routing scheme in order to maximize the network lifetime while forwarding the packets through energy constrained nodes. Routing to maximize the network lifetime is more advantageous compared to minimum energy routing. Minimum energy routing protocols, tends to forward most of the data packets through the energy efficient path discovered during the route discovery phase. Hence, nodes in the energy efficient path exhaust very soon. Hence, it is vital to route the packets in a network by balancing the lifetime of all the nodes, so that desired network performance can be achieved for a longer time. Therefore, the energy efficiency is not only measured by the power consumption, but in a more general sense it can be measured by the duration of time over which the network can maintain a certain performance level.

Similar to AODV routing protocol, when a source node has data to send to the destination node, it broadcasts a route request packet, RREQ to its neighbors. The <source-address, broadcast-id> pair is used to identify the RREQ uniquely to control the overhead created by flooding and reduce the repeated transmission of packets. To support energy efficiency, and not to add extra overhead, the proposed scheme modifies the RREQ format of AODV, where the hop count and reserved field in RREQ frame is replaced with E_{thresh}, and $E_{residual}$ respectively. E_{thresh} is a minimum threshold value of energy that the node can have and $E_{residual}$ is the residual energy of the node. The $E_{residual}$ field is initialized to a maximum value by the source.

6.3.3 Adaptive Cooperative MAC

To optimize the network performance, AMCCR exploits the MAC cooperation, while forwarding the packets from source to destination. Once the route from the source to the destination is discovered, two adjacent nodes on the route can select a relay node for cooperation to satisfy QoS requirements. For example, consider a sample network as shown in Figure 6.3, where the path is discovered between source V_i to the destination V_j. Two adjacent nodes in the chosen path, say V_i and V_{i+1}, can select and utilize the relay node R_1 or R_2, which are in the interference range of both these nodes for adaptive cooperative transmission. For each hop in the routing path, determining the need for

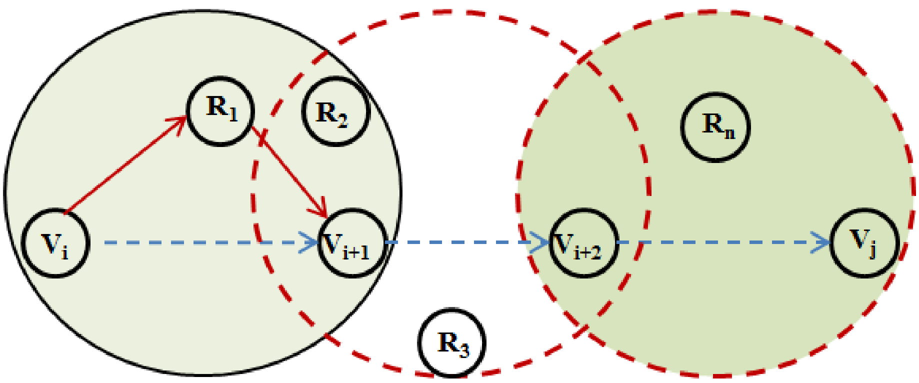


FIGURE 6.3: Multi-hop cooperative routing path

relay node between two adjacent nodes and selecting the optimal one from a number of available relay nodes is challenging. To address this problem, AMCCR proposes an adaptive cooperative MAC. The primary focus of adaptive cooperative set up in the MAC layer is to meticulously allocate the resources and to engage cooperative relay nodes in setting up the cooperative environment. The nodes willing to communicate over the 802.11 network use four way handshake procedure to eliminate the hidden terminal issue.

The proposed AMCCR protocol uses receiver based rate adaptation technique RBAR [19], where the receiver dictates the data rate for the sender. When the node V_i transmits the RTS with the payload of length L to V_{i+1}, due to the broadcast nature of the wireless channel, neighboring nodes R_i and R_{i+1} overhear the RTS. To qualify as a relay node, R_i checks channel allocation for any ongoing communication in the interference region through network allocation vectors (NAV) to alleviate hidden terminal problems. If NAV is not set, it decodes the RTS and determine the appropriate data rate $DR_{V_i - Ri}$ between the V_i and R_i based on the estimated signal noise ratio (SNR) and bit error rate (BER). Similarly, when V_{i+1} receives RTS, it determines data rate $DR_{V_i - Vi+1}$ between the V_i and V_{i+1} and transmits CTS incorporating the $DR_{V_i - Vi+1}$ to node V_i. If NAV of neighboring node R_i is not set, then it decodes the CTS transmitted from the node V_{i+1} and selects the appropriate data rate $DR_{R_i - Vi+1}$.

6.3.3.1 Best Relay Selection Scheme

Several neighboring nodes may qualify to participate in the adaptive cooperation transmission. The next challenge is to select the efficient neighbor node as a relay node, among the qualified neighboring nodes, for cooperation with minimal overhead. In the existing schemes discussed in literature, the relay selection mechanism mostly incurs extra overhead due to complicated interactions among the neighboring nodes. When every neighbor node starts transmitting the request to the source to become a relay node, it will lead to wastage of bandwidth and energy of the nodes, which may further incur additional delay. Thus, it is essential to analyze the impact of complicated interactions among the neighboring nodes on cooperative diversity performance and minimize it to the lowest possible level. In order to achieve multi QoS objectives such as energy efficiency, delay and high throughput, the AMCCR protocol needs to make decision on the following:

1. Centralized relay selection versus Distributed relay selection: In centralized relay selection technique [135] [33], the relay selection is done in a passive listening mode with centralized control. In this technique, all the neighbor relay's channel state information (CSI) is accumulated and compared, which increases complexity and delay. The scheme proposed in [135] requires periodic broadcast of readiness to participate message by each neighbor node to its one hop neighbors, irrespective of whether cooperative mode is needed or not. Whereas, in case of distributed relay selection technique, nodes in the network dynamically select among the qualified relay nodes based on the local network information available.

2. Fixed relaying scheme versus Adaptive relaying scheme: In the fixed relaying scheme, data transmission always happens via relay node in the cooperative manner, even when the destination node can directly receive and decode the data packets transmitted from the source correctly. Hence, time slot used by the relay to forward the data packets is a waste of resources and takes double the time to transmit packets than the direct transmission. To counter these problems, adaptive relay selection would be more recommendable, wherein packets will be transmitted cooperatively only if the network demands else packets will be transmitted non- cooperatively.

3. Single versus Multiple relays: To improve the network performance, few authors in [155][156] proposed to use multiple relay nodes with the intent of increasing the diversity gain. However, with the multiple relay nodes participating in cooperative communication creates a larger interference area and causes additional coordination overhead. Thus, affecting the overall throughput. The authors in [157] proved that single relay node achieves the same diversity gain as that of multiple relay nodes.

In AMCCR protocol, a distributed, adaptive, single relay selection method is employed. In the proposed protocol, selection of best relay node is carried out when the direct transmission from the transmitting node V_i to receiving node V_{i+1} in a multi hop network fails due to fading, or the relay path transmission time is better than the direct path. In such a scenario, the neighboring nodes with a potential to be a relay node, participates in the relay node selection process by using the local information collected by it. The neighbor nodes will have to satisfy certain QoS metric test to qualify as the best relay node among the other competing neighboring nodes. If the direct transmission path between transmitting node and receiving node satisfies the QoS requirements, the relay nodes will not participate in the communication and the protocol will be reduced to simple DCF.

6.3.3.2 QoS Metrics for Relay Selection

The various QoS metric tests a neighbor node has to pass to qualify as a relay node to support cooperative MAC mechanism are as follows:

1. Transmission time: The first QoS metric is the transmission time. The neighbor nodes, on hearing the RTS and CTS between transmitting node V_i and receiving node V_{i+1}, estimates the data rate from the SNR of receiving signals. Then, it estimates the cooperative transmission time that would incur between the node V_i to node V_{i+1} , if it participates as a relay node. The cooperative transmission time T_{coop} and the direct transmission time T_{direct} are computed as,

$$T_{coop} = \frac{L}{DR_{V_i - R_i}} + \frac{L}{DR_{R_i - V_{i+1}}} + T_{RE} + 2 * T_{SIFS} \tag{6.20}$$

$$T_{direct} = \frac{L}{DR_{V_i - V_{i+1}}} \tag{6.21}$$

where,

L is the packet length,

T_{RE} is the transmission time for RE frame which is sent by the candidate relay node to V_i and V_{i+1} to notify its willingness to participate in the cooperative communication,

T_{SIFS} is the SIFS interval.

If the neighbor node prefers to be a relay node, the total transmission time via relay node i.e., T_{coop} should be less than the direct transmission time T_{direct}.

2. Channel contention metric: The second QoS metric focuses on the channel contention. In MANETs, due to constant topological change, nodes may cluster at certain area and there could be high inflow and outflow of data within that region leading to high interference. Therefore, the average channel contention time of the node may increase, thereby degrading the throughput of the network. In the proposed protocol, a node having a packet to forward, will run a contention counter $T_{cc}(t)$ from the start of channel contention till it wins the channel access. The average contention time for a node $T_{cc-avg}(t)$ is computed using exponential weighted moving average over time Δt and is given by,

$$T_{cc-avg}(t) = \alpha T_{cc}(t) + (1-\alpha) T_{cc-avg}(t-1) \tag{6.22}$$

where,

α is a constant smoothing factor between 0 and 1.

If the neighbor node prefers to be a relay node, the $T_{cc-avg}(t)$ should be lesser than the $T_{cc-thresh}$, specified acceptable time duration.

3. Energy utilization factor: The third metric, energy efficiency is undoubtedly one of the apt metrics for quality evaluation. The network lifetime is defined as the time from the deployment of the nodes to the instant the first node dies. So to maximize the network lifetime, data has to be routed such that energy expenditure is fairly among the nodes in proportion to their energy reserved. The energy levels of all the nodes in the network have to be balanced and the node's death due to frequent communication should be minimized to extend the network lifetime. In the proposed algorithm, the energy required by the neighbor node during cooperative communication is computed as,

$$E_{cooperative} = \frac{P_{receiver}}{DR_{V_i-R_i}} + \frac{P_{transmitter}}{DR_{R_i-V_{i+1}}} \tag{6.23}$$

where,

$P_{receiver}$ is the power required to receive the data packet from the source

$P_{transmitter}$ is the power required to transmit the data packet to the destination.

Each neighbor node also stores its residual energy, $E_{residual}$. The proposed AMCCR protocol uses a novel way of interpreting energy using a metric called the energy utilization factor to select a relay node which is not only energy efficient but also assists in improving the network lifetime. The energy utilization metric ε value is given as,

$$\varepsilon = \frac{E_{cooperative}}{E_{residual}} \tag{6.24}$$

Energy utilization metric ε is a better parameter compared to $E_{cooperative}$ or $E_{residual}$ metric. For example, consider the two neighbor nodes N_1 and N_2 satisfying the above two QoS metric tests. Let us assume that N_1 and N_2 has the residual energy of 30 J and 50 J respectively. Many of the routing protocols select the next hop based on residual energy or the minimum power needed to transmit. Let us assume, if N_1 or N_2 is selected as the relay node for cooperation, the energy required for transmission be 10 J and 25 J respectively. So, according to maximum residual energy selection criteria, N_2 is selected and after cooperative transmission the residual energy of N_1 and N_2 would be 30 J and 25 J respectively.

According to the proposed energy utilization selection metric, the value of ε for N_1 and N_2 will be $\varepsilon(N_1)$=10/30 and $\varepsilon(N_2)$= 25/50 respectively. The node with minimum ε will be selected, i.e., node N_1. Thus, after cooperative transmission the residual energy of N_1 and N_2 would be 20 and 50 respectively. Even though the ε factor does not guarantee total energy consumption minimization, it maximizes the minimum value of $E_{residual}$ and maintains energy levels of the nodes in the network in a balanced state. Therefore, this factor extends the node survival time and improves the network lifetime.

In the network, there may be several nodes which have qualified all the three QoS metric tests and are candidate nodes for best relay node selection. So, to optimally select the best relay node, each node computes the network utilization weight based on the individual QoS metric weight. Each node computes its weight W_{relay} at regular interval of time and is given by,

$$W_{relay} = \delta_1 * \frac{T_{coop}}{T_{coop-max}} + \delta_2 * \frac{T_{cc-avg}}{T_{cc-max}} + \delta_3 * \varepsilon \tag{6.25}$$

where,

$\sum_{i=1}^{3} \delta_i = 1$,

$T_{coop-max}$ is the maximum acceptable cooperative transmission time,

T_{cc-max} is the maximum acceptable channel contention time.

It can be seen from the above equation that the metrics are normalized by their maximum values with some multiplicative factor δ.

6.3.3.3 Relay Selection Overhead

Due to the mobility of nodes, multiple relay nodes may be available satisfying the QoS metric test and they communicate their willingness to source node via a special Relay Eligible (RE) packet. If all the candidate relay nodes starts contending for the channel to send RE packet, then it may lead to collisions. Such frequent collisions may reduce the cooperation opportunity and degrade system performance. Thus, it is a challenging task to select the best relay efficiently with minimum collision rate. So, it is required to make a trade-off between relay selection period and collision probability. In order to avoid the collision from multiple candidate relay nodes, each of these nodes waits for an additional time λ after SIFS time. The λ time slot is divided into multiple τ slot of length ω. The τ and ω are computed as,

$$\omega = 2 * P_{delay} + T_{switch} \tag{6.26}$$

$$\tau = \left\lfloor \frac{\lambda}{\omega} \right\rfloor \tag{6.27}$$

where,

P_{delay} is a channel propagation delay,

T_{switch} is transceiver switching time.

We further divide the time slots τ into ϕ to map W_{relay} value, which is given as,

$$\phi = (W_{relay_{max}} - W_{relay_{min}})/\tau \tag{6.28}$$

where,

$W_{relay_{max}}$ and $W_{relay_{min}}$ is the maximum and minimum possible value of W_{relay}.

Each source node starts its counter from zero after SIFS interval, and when it reaches to its W_{relay} value, wins the channel access and sends RE control frame in that time slot to node V_i and V_{i+1}.

The format of RE frame is as shown in Figure 6.4 . The priority field in RE frame is set to a value that matches with the corresponding data rate pair as per Table 6.1. If RE frame is received within the timeout, node V_i switches from direct transmission to cooperative transmission and selects the appropriate MAC scheme. No retransmission attempt is allowed if RE frames collide or the relay backoff counter reduced to zero. Table 6.2 shows the comparison of some of the existing schemes and proposed scheme

Frame control	Duration	Source Address (V_i)	Relay Address (R_i)	Next hop Address (V_{i+1})	Sequence control	Data rate (DR_{Vi-Ri})	Data rate ($DR_{Ri-Vi+1}$)	$L_{vi-vi+1}$	$L_{vi+1-vi}$	Priority	FCS
2 bytes	2 bytes	6 bytes	6 bytes	6 bytes	2 bytes	1 byte	1 byte	2 bytes	2 bytes	1 byte	2 bytes

FIGURE 6.4: Relay Eligible (RE) frame format

for relay selection. From the table, we can infer the novelty of relay selection in our proposed AMCCR protocol which employs distributed relay selection based on multi QoS metrics, initiated by the relay based on local information.

6.3.3.4 Adaptive Cooperative MAC Schemes in AMCCR

In AMCCR protocol, the data transmission can be in one of the four categories:

1. Direct transmission

TABLE 6.1: Priority for 802.11g data rate pair(V_i-R_i, R_i-V_{i+1})

Priority	**Data Rate Pair(V_i-R_i, R_i-V_{i+1})**
1	54-54, 54-48, 54-36, 54-24, 54-18
2	48-48, 48-36, 48-24, 48-54, 48-18, 36-36, 36-24, 36-54, 36-48, 36-18
3	24-24, 24-54, 24-48, 24-36, 24-18, 18-54, 18-48, 18-36, 18-24,18-18
4	54-12, 54-9, 48-12, 48-9,
5	36-12, 36-9, 24-12
6	24-9, 18-12, 18-9, 12-9
7	12-54, 12-48, 12-36, 12-24, 12-18, 12-12
8	9-54, 9-48, 9-36, 9-24, 9-18, 9-12, 9-9

TABLE 6.2: Comparison of various relay selection schemes

Protocol	Cooperation decision	Centralization	Selection overhead	Network Coding functionality	Number of QoS metric
CD-MAC	Trasmitter	Centralized	Preselect, Historical information	No	Single
rDCF	Receiver	Distributed	Preselect, Historical information	No	Single
CODE	Transmitter	Centralized	Periodic broadcast	Partial	Single
CoopMAC	Transmitter	Centralized	Passive monitoring	No	Single
utd MAC	relay	Centralized	Preselect	No	Single
AMCCR	relay	Distributed	RTS-CTS contention	Yes	Multiple

2. Sender-relay-receiver

3. Cooperative transmission scheme using DF technique

4. Cooperative transmission scheme using AF technique

These MAC schemes are adaptively selected based on priority field of RE frame. The cross layer adaptive data transmission is briefly explained in Algorithm 5. To facilitate the proper selection of transmission scheme, CTS packet is modified to accommodate a flag known as FLAG_P.

1. Direct Transmission: Based on the received signal quality transmitted from node V_i, the receiving node V_{i+1} sets its FLAG_P field while replying CTS. If the direct transmission is sufficient, i.e. SNR of received signal greater than SNR_{thresh}, then the FLAG_P is set to 0. The neighbor node while decoding the CTS, will notice the FLAG_P field and refrain from interfering with the ongoing communication if it is set to 0. Figure 6.5 demonstrates the scheme.

Algorithm 5 Adaptive Cooperative MAC data transmission

```
if RE frame not received within timeout then
    Transmission mode = direct transmission
end if
if RE frame received then
    Check the Priority field in RE frame
    if 1 ≤ Priority ≤ 3 then
        Transmission mode= sender-relay-receiver transmission
        if 4 ≤ Priority ≤ 6 then
            Transmission mode= DF cooperative transmission
        else if Priority > 6 then
            Transmission mode= AF cooperative transmission
        end if
    end if
end if
```

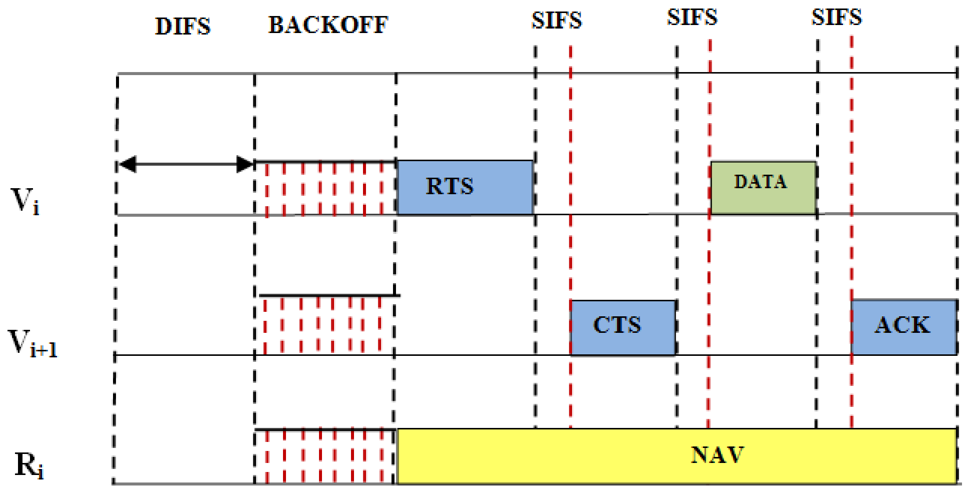


FIGURE 6.5: Direct transmission

2. Sender-Relay-Receiver Transmission: If FLAG_P is set to 1 and the priority field of RE frame contains the value between 1 to 3, then AMCCR protocol prefers simple sender-relay-receiver transmission scheme. Priority 1-3 indicates a high data rate estimated between the link $V_i - R_i$ and $R_i - V_{i+1}$ and hence, the probability of data received at relay node or the receiver node with error is very minimal. Since the transmission rate from sender node to relay node is too high, receiver node does not overhear it due to reduced range and hence minimizes the interference. Figure 6.6 demonstrates the scheme.

3. DF Cooperative transmission : When priority field of RE frame contains the value between 4 to 6, the data rate estimate from source to relay node is higher

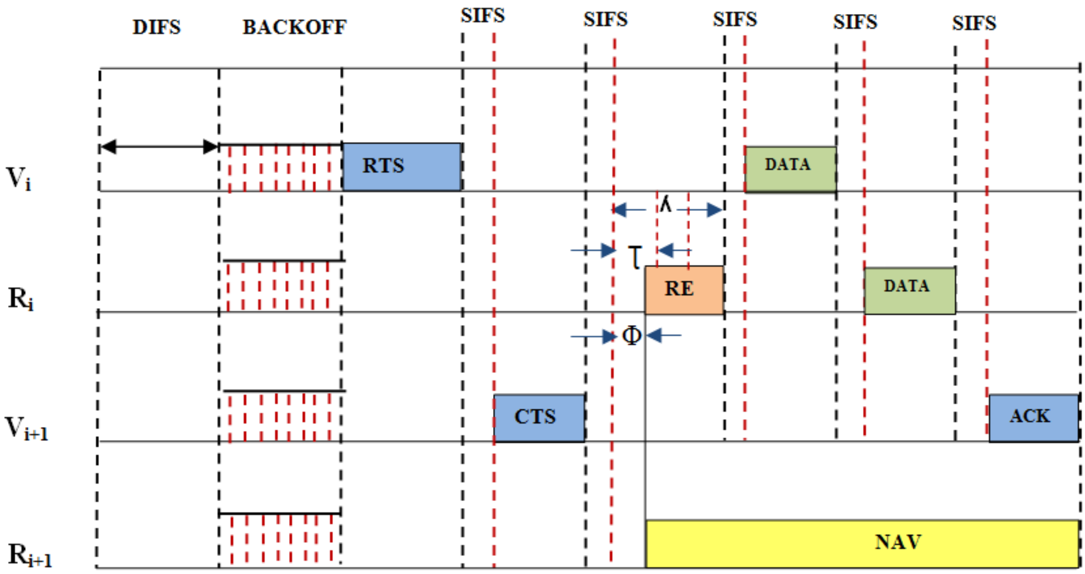


FIGURE 6.6: Cooperative transmission

compared to relay node to destination node. When the relay is located closer to the source terminal, channel quality between node V_i and relay R_i is better than that between relay R_i and node V_{i+1}. When relay R_i moves away from the source, the BER increases and the relay node may send erroneous bits to the destination. Under this situation, DF technique is always better than that of the AF technique and guarantee a performance diversity of order two [30]. Therefore, when FLAG_P is set to 1 and the priority field of RE frame contains the value between 4 to 6, AMCCR protocol prefers DF Cooperative transmission. DF technique is always better than that of the AF technique and guarantee a performance diversity of order two [30].

4. AF Cooperative transmission: If FLAG_P is set to 1 and the priority field of RE frame contains the value greater than 6, the AMCCR protocol prefers the AF Cooperative transmission. Priority greater than 6 indicates low data rate estimate between the V_i - R_i and high data rate estimate between R_i - V_{i+1}. Performance in terms of BER is good when the relay node is closer to the destination node. As the relay node moves away from the destination, signal strength drops due to high data rate. Hence, cooperative scheme with AF is preferred to guarantee a performance diversity of order two as compared to DF scheme.

6.3.3.5 Relay Reassignment

Every relay node maintains a table containing the information about the ongoing neighbor node transmissions. This restrains the neighbor nodes from unnecessarily participating in ongoing cooperative transmission. During cooperative transmission, relay node may fail to satisfying the relay qualifying criteria due to nodes mobility or channel conditions.The existing cooperative protocols mentioned in the literature does not address the issue in case a selected relay goes offline or is unavailable. In such cases, AMCCR proposes a mechanism to automatically re-elect the new relay node. The neighbor nodes on not hearing the RE frame from the current relay node, will spontaneously initiate the transmission of RE frame in the next RE time slot. So, whichever neighbor node is successful in forwarding the RE frame without collision, replaces the previous relay node as the current best relay node. The efficient relay selection scheme proposed in AMCCR has the following characteristics:

- Relay selection is time efficient.
- The Collision probability of relay selection is minimized.
- Relay selection dynamically adapts to time varying channel condition and node mobility.
- Relay selection is done in a distributed manner.
- No hidden node problem.

6.3.4 Cooperative Routing

In a mobile ad hoc network, cooperation at the MAC layer can be exploited at the network layer to enhance the system performance. Unlike the traditional routing protocol, we exploit the MAC cooperation, while selecting the route from source to destination. Thus, cooperative routing is activated whenever the opportunity of cooperation gain exists. Consider a simple network topology given in Figure 6.3. Let us assume that the proposed routing scheme discovers the route from node V_i to node

V_j, via nodes V_{i+1} and V_{i+2} during initial route discovery phase. During the control packet exchange between V_i and V_{i+1} , if R_i qualifies to be the best relay node to support cooperative transmission, then node V_i and V_{i+1} will update their route table with the additional entry for the relay node. Thus, the route layer will have access to cooperative link metrics of relay between two adjacent nodes. The route layer fully exploits this information while choosing the best next hop from the route table. The possible route from node V_i to V_j can be $V_i - R_i - V_{i+1} - V_{i+2} - V_j$. Hence, the cross layer mechanism utilizes cooperative diversity that helps in improving the network performance.

6.3.5 Network Coding Through Cooperative Communication

In a dual-hop half-duplex relay network, the throughput performance is reduced to half as it cannot transmit and receive simultaneously. To counter this fallback, the proposed AMCCR implements the network coding scheme in cooperative mode to support efficient communication. The network coding [158] is the technique that forwards the data received from multiple nodes by combining them into one or more output data packets. This results in a reduced number of packet transmissions, reducing the delay and enhancing the network performance. This technique is widely being accepted and can be exploited in the cooperative communication [159]. Whenever the destination node also has data to communicate back to the source, the destination node sends this information to the source via a duration field in the CTS packet by mentioning the length of packet to be communicated to the source. Consider an simple network coding example in half-duplex channels with a time-division duplex (TDD) as shown in Figure 6.7, where node V_i and V_{i+1} wants to exchange the data packets for communication through the selected relay node R_i. Network coding technique works in half duplex mode, which is divided into three phases, where the channel is time multiplexed. In phase 1, node V_i forwards the data packet $DATA_{V_i - V_{i+1}}$ of length $L_{V_i - V_{i+1}}$ bytes to relay R_i. In phase 2, node V_{i+1} forwards the data packet $DATA_{V_{i+1} - V_i}$ of length $L_{V_{i+1} - V_i}$ bytes to relay R_i. Finally, in phase 3, relay node R_i performs exclusive OR operation on $DATA_{V_i - V_{i+1}}$ and $DATA_{V_{i+1} - V_i}$ and broadcasts the result to nodes V_i

and V_{i+1}. The nodes V_i and V_{i+1} recover the received data packets destined to itself. At the end of phase 3, ACK packets are exchanged between node V_i and V_{i+1}, and transmission completes. In case an data packet gets lost, the retransmission of packet is initiated. Therefore, by implementing the network coding scheme, only single channel contention is needed. The length of packets from source to destination $DATA_{V_i-V_{i+1}}$ and destination to the source $DATA_{V_{i+1}-V_i}$ is computed through the duration values in RTS and CTS respectively. Figure 6.8 demonstrates the exchange of control packets in

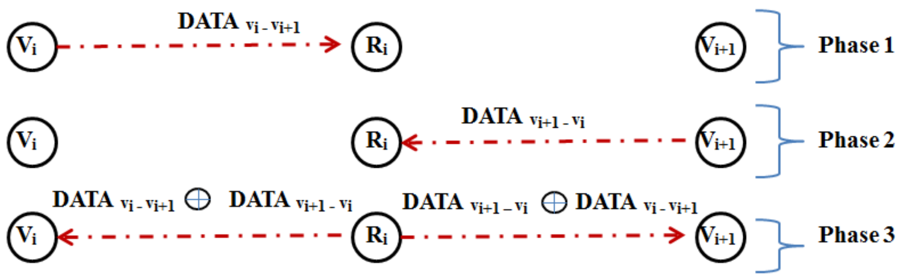


FIGURE 6.7: Example of network coding in half-duplex channel

the proposed cooperative coding scheme. The condition for a relay node to participate in cooperative network coding technique is revised and is given by,

$$\begin{aligned}\frac{L_{V_i-V_{i+1}}}{DR_{V_i-R_i}} + \frac{L_{V_{i+1}-V_i}}{DR_{R_i-V_{i+1}}} + \frac{max(L_{V_i-V_{i+1}}, L_{V_{i+1}-V_i})}{min(DR_{V_i-R_i}, DR_{R_i-V_{i+1}})} \\ +T_{RE} < \frac{L_{V_i} - V_{i+1}}{DR_{V_i-V_{i+1}}} + \frac{L_{V_{i+1}-V_i}}{DR_{V_i-V_{i+1}}} + T_{RTS} \\ +T_{CTS} + T_{SIFS} + T_{DIFS} + T_{backoff}\end{aligned} \tag{6.29}$$

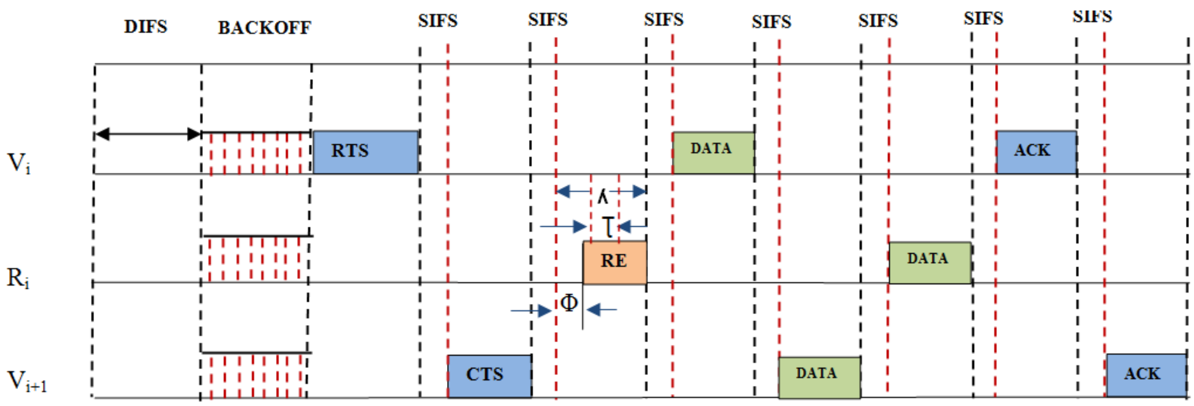


FIGURE 6.8: Cooperative transmission using Network Coding scheme

6.3.6 NAV Adaptation

The IEEE 802.11 wireless network uses a virtual carrier sensing mechanism which uses network allocation vector (NAV). NAV helps to limit the nodes from accessing wireless medium to avoid multiple interferences, and conserve energy. The transmitting node specifies the transmission time required for the frames like RTS, CTS, and DATA, during which the channel will be busy. Every node listening to channel on the network sets their NAV for which they have to defer their transmission. During direct communication, the NAV is calculated based on the transmission data rate between source and destination. But during cooperative communication, setting NAV for RTS and CTS accurately is not possible until the relay node information is available to the source and the destination nodes. Thus, in the proposed AMCCR protocol, the neighboring node updates their NAV after every frame transmission. Updating the NAV of control packets in the proposed AMCCR protocol is described below:

1. Whenever a node V_i has data to forward, it senses the channel to check if it is idle for a DCF interframe space (DIFS) time. Once it completes the required backoff time, sends RTS and reserves the channel for RTS_{NAV} time. At this instance, since V_i has no idea about the cooperation or network coding, it uses basic data rate to compute channel reservation based on direct communication.

$$RTS_{NAV} = 4 * T_{SIFS} + T_{CTS} + \frac{L_{V_i - V_{i+1}}}{DR_{V_i - V_{i+1}}} + T_{ACK} \tag{6.30}$$

where, $L_{V_i - V_{i+1}}$ is the length of the data packet, and $DR_{V_i - V_{i+1}}$ is the data rate from node V_i to V_{i+1}.

2. The node V_{i+1} on receiving RTS replies with CTS after a SIFS time and reserves the channel for NAV duration CTS_{NAV}.

$$CTS_{NAV} = 3 * T_{SIFS} + \frac{L_{V_i - V_{i+1}}}{DR_{V_i - V_{i+1}}} + T_{ACK} \tag{6.31}$$

3. When the neighboring node receives the RTS and CTS frame, checks if it is eligible to be a relay node as per best relay selection scheme. If it qualifies to be a relay node, it sends a RE frame to node V_i and V_{i+1} after a SIFS time and reserves the channel for the NAV duration of RE_{NAV}. If node V_{i+1} also has data to communicate with V_i, then V_{i+1} includes the information about dual half duplex communication in CTS packet. The relay node on finding dual half duplex communication in the CTS, prepares to forward the data using network coding scheme otherwise by a cooperative scheme.

   For cooperative scheme,

$$RE_{NAV} = 2 * T_{SIFS} + \frac{L_{V_i - V_{i+1}}}{DR_{V_i - Ri}} + \frac{L_{V_i - V_{i+1}}}{DR_{Ri - V_{i+1}}} \tag{6.32}$$

   For network coding scheme,

$$RE_{NAV} = 3 * T_{SIFS} + \frac{L_{V_i - V_{i+1}}}{DR_{V_i - Ri}} + \frac{L_{V_{i+1} - V_i}}{DR_{Ri - V_{i+1}}} + \frac{max\,(L_{V_i - V_{i+1}}, L_{V_{i+1} - V_i})}{min\,(DR_{V_i - Ri}, DR_{Ri - V_{i+1}})} \tag{6.33}$$

4. If node V_i and V_{i+1} receive RE frame within timeout, it sends the data either through cooperative technique or network coding. Otherwise, node V_i sends the data packet through a direct transmission to node V_{i+1}. $RE_{timeout}$ is given by,

$$RE_{timeout} = T_{SIFS} + T_{RE} \tag{6.34}$$

   In cooperative scheme, the NAV for data packet communication between $V_i - R_i$ and $R_i - V_{i+1}$ is given by,

$$DATA_{V_i - Ri} = 2 * T_{SIFS} + \frac{L_{V_i - V_{i+1}}}{DR_{Ri - V_{i+1}}} + T_{ACK} \tag{6.35}$$

$$DATA_{R_i - V_{i+1}} = T_{SIFS} + T_{ACK} \tag{6.36}$$

   For network coding scheme,

$$DATA_{V_i - Ri} = 4 * T_{SIFS} + \frac{L_{V_{i+1} - V_i}}{DR_{Ri - V_{i+1}}} + 2 * T_{ACK} + \frac{max\,(L_{V_i - V_{i+1}}, L_{V_{i+1} - V_i})}{min\,(DR_{V_i - Ri}, DR_{Ri - V_{i+1}})} \tag{6.37}$$

$$DATA_{V_{i+1}-Ri} = 3 * T_{SIFS} + 2 * T_{ACK} + \frac{max\,(L_{V_i-V_{i+1}}, L_{V_{i+1}-V_i})}{min\,(DR_{V_i-Ri}, DR_{Ri-V_{i+1}})} \tag{6.38}$$

$$DATA_{NC} = 2 * T_{SIFS} + 2 * T_{ACK} \tag{6.39}$$

5. On successful reception of data, receiver node V_{i+1} sends an ACK packet back to transmitter node V_i. If V_i receives ACK, within $ACK_{timeout}$, the data transmission is successful. Otherwise, the transmitter node V_i resumes the backoff procedure to contend for the channel.

   For cooperative scheme,

$$ACK_{timeout} = 2 * T_{SIFS} + \frac{L_{V_i-V_{i+1}}}{DR_{V_i-Ri}} + \frac{L_{V_i-V_{i+1}}}{DR_{Ri-V_{i+1}}} + T_{ACK} \tag{6.40}$$

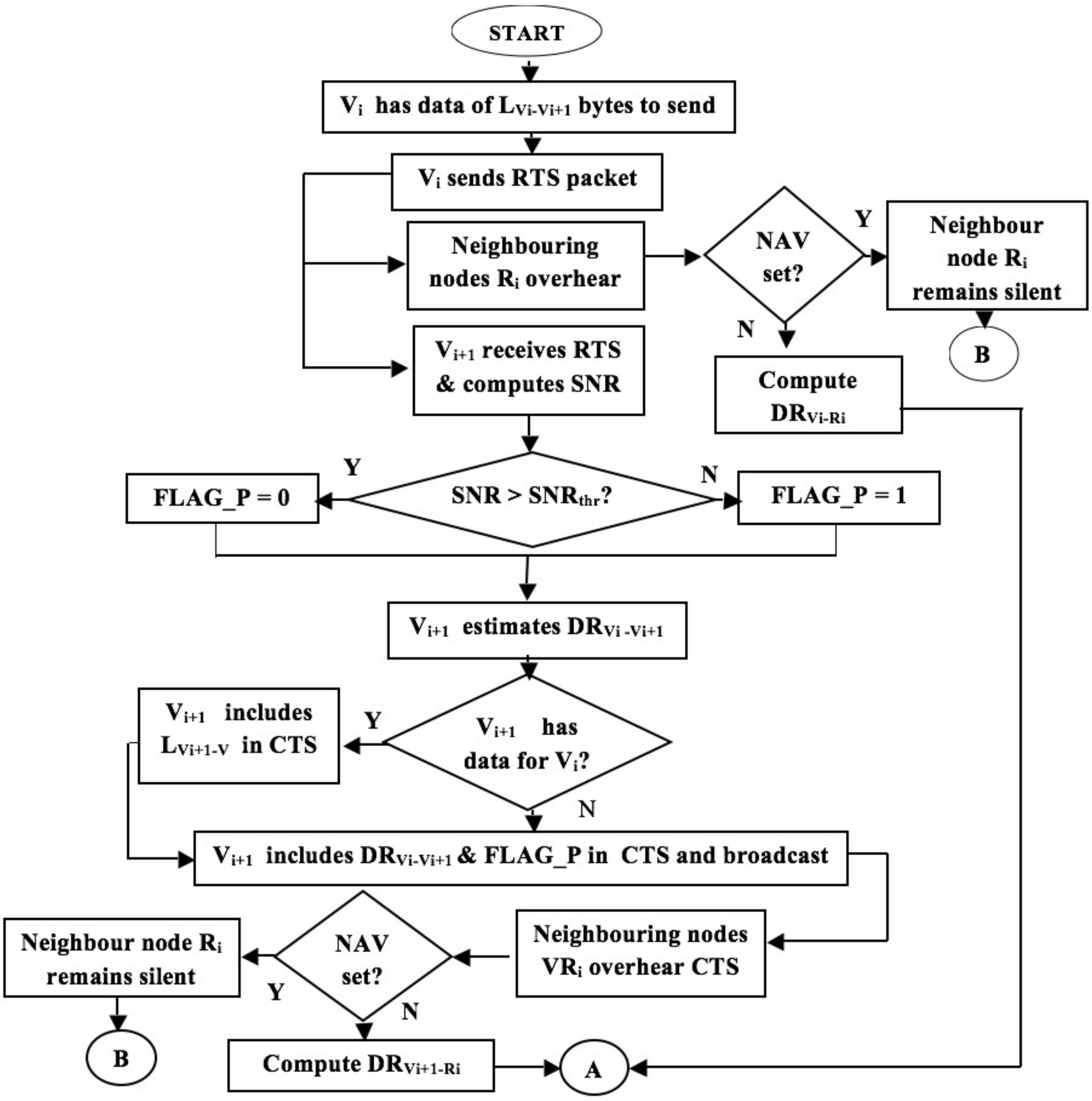
   For network coding scheme,

$$ACK_{timeout} = 2 * T_{SIFS} + \frac{L_{V_i-R_i}}{DR_{V_i-Ri}} + \frac{L_{R_i-V_{i+1}}}{DR_{Ri-V_{i+1}}} + 2 * T_{ACK} \tag{6.41}$$

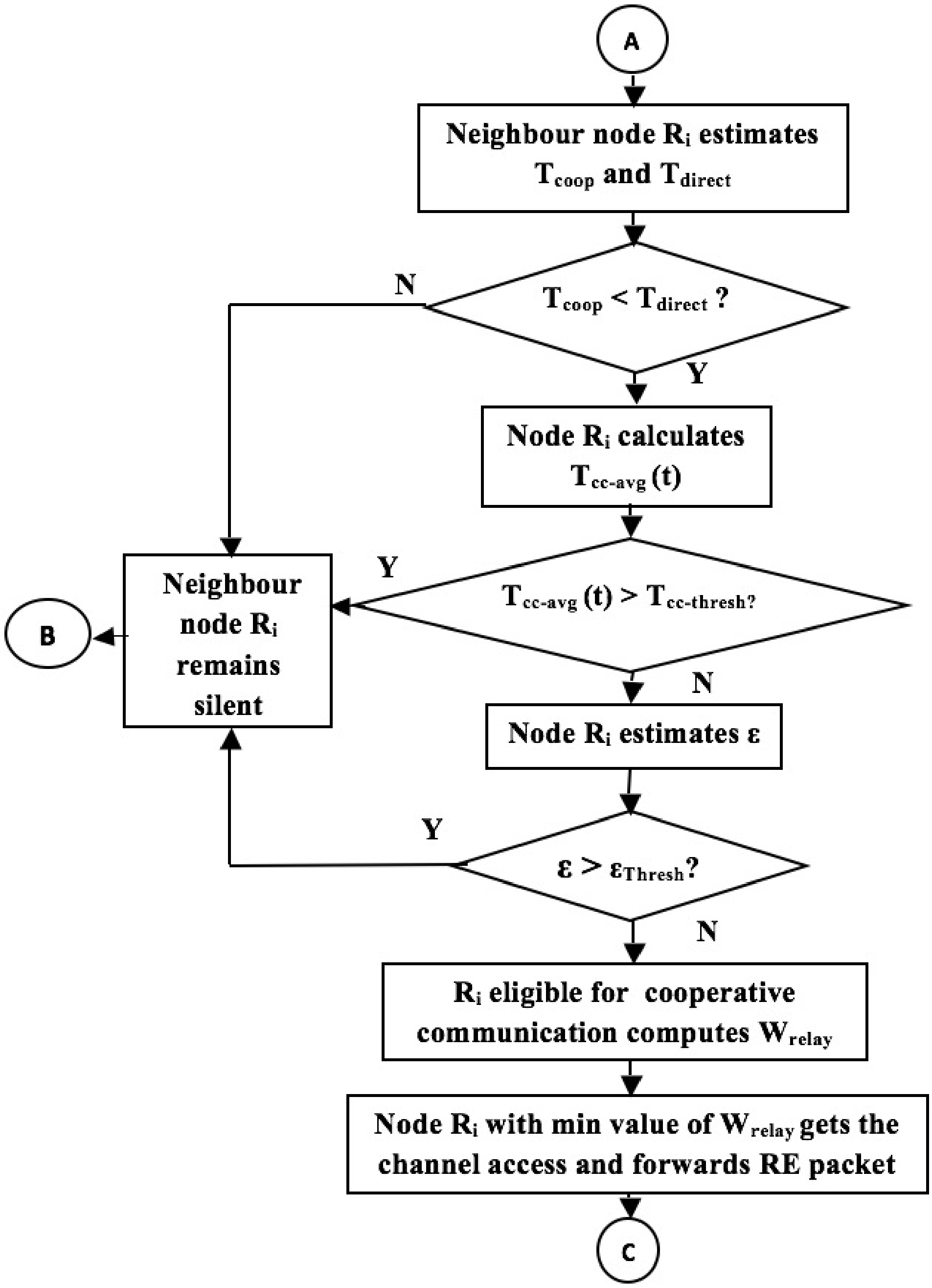
6. The nodes that receive only RTS from the source, but not the CTS, sets their NAV until the end of ACK.

6.3.7 Flowchart of proposed AMCCR protocol

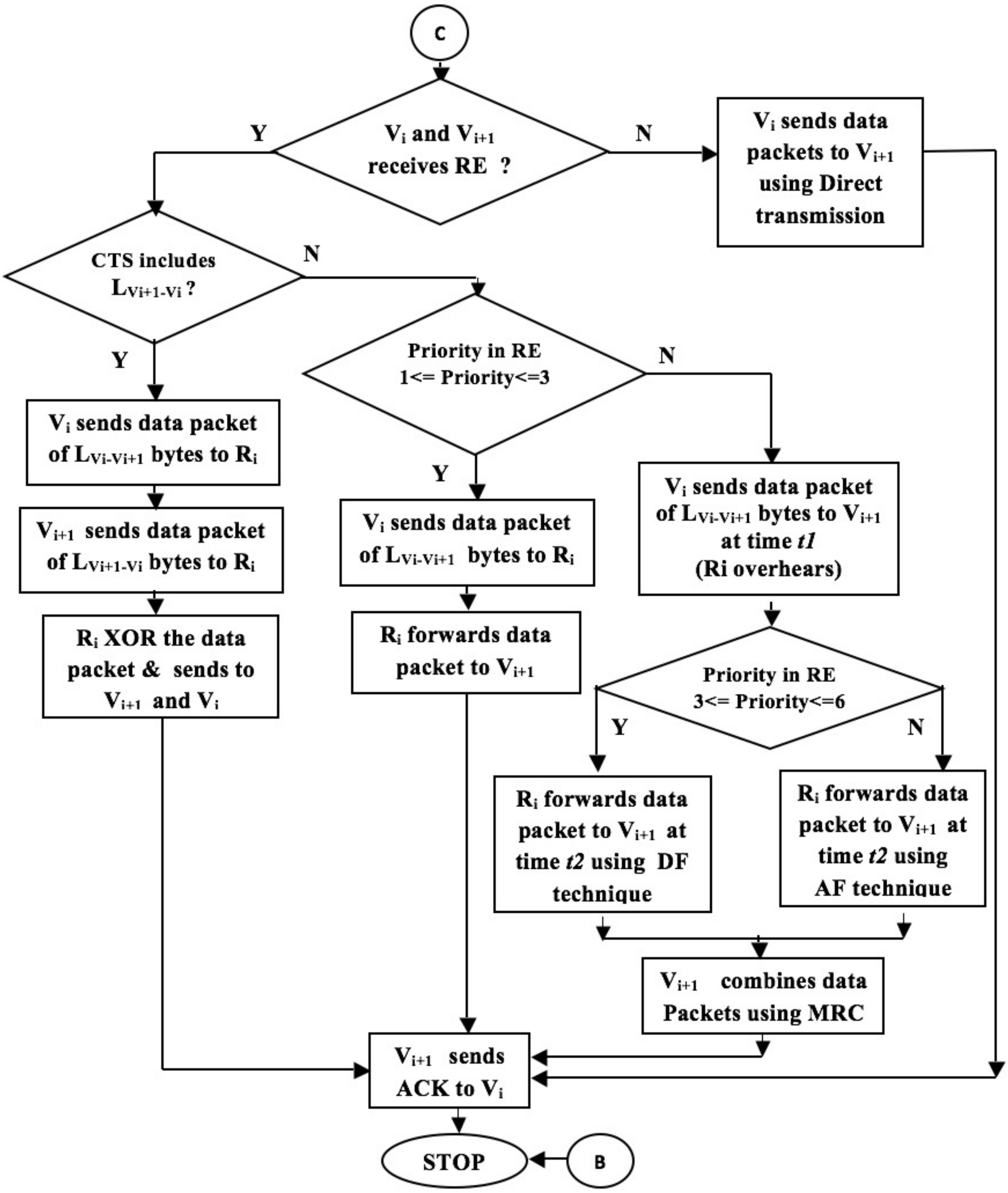
Figure 6.9a - 6.9c shows the flowchart of detailed functionality of proposed AMCCR protocol.



(A)


A
Neighbour node R_i estimates T_{coop} and T_{direct}
$T_{coop} < T_{direct}$?
N
Y
Node R_i calculates $T_{cc\text{-}avg}$ (t)
$T_{cc\text{-}avg}$ (t) > $T_{cc\text{-}thresh}$?
Y
N
Neighbour node R_i remains silent
B
Node R_i estimates ε
ε > $ε_{Thresh}$?
Y
N
R_i eligible for cooperative communication computes W_{relay}
Node R_i with min value of W_{relay} gets the channel access and forwards RE packet
C

(B)



(c)

FIGURE 6.9: (A)-(C) Detailed functionality of proposed AMCCR protocol

6.4 Simulation Setup

The AMCCR protocol is evaluated for MANET based on IEEE 802.11g. The simulation using the NS2.35 simulator [96] assumes a network consisting of mobile nodes deployed over an area of 1000 x 1000 meters. The random movement of mobile nodes is modeled using the random waypoint modeling with a speed set to 5 m/s. The environment noise level is modeled as Gaussian random variables with a noise level ranging from -83 dBm to indicate harsh communication environment to -92 dBm to indicate a good communication environment for 802.11g with a standard deviation of 1 dBm. The source generates Constant Bit Rate (CBR) traffic packets at a rate of 5 packets/s with size of 512 bytes. The source destination pairs are randomly selected. The initial energy of all nodes is set to 60 Joules and E_{thresh} is assumed to be 5 Joules. The channel between each pair of nodes is set as independent Rayleigh fading channel. The transmission rate of the data packet is computed based on the average SNR value of the received signal at the receiver. The channel in the physical layer utilizes the multi-rate transmission defined in IEEE 802.11g. The experiment is simulated for 900 s time duration, and the results are averaged after 20 executions. Each execution uses a seed value ranging from 1 to 9 for randomness in node placement and steady results. The performance of the proposed AMCCR protocol is evaluated in two modes:

1. Cooperative routing mode
2. Cooperative coding mode

6.4.1 Performance Evaluation of Cooperative Routing Scheme

AMCCR under cooperative routing mode is compared with recently developed cooperative diversity protocol CD-MAC and conventional non cooperative IEEE 802.11 DCF. The protocols CD-MAC and 802.11 DCF assumes the classical AODV routing protocol for route computation. In the simulation, protocols are compared and evaluated under two different scenarios i.e, by varying traffic flow and varying network density.

6.4.1.1 Effect of Varying Traffic Flow

Under this scenario, the concurrent TCP traffic flow connection is varied from 4 to 20 at the rate of 4 packets/s and with two different noise levels i.e., -92 dBm and -83 dBm. The number of nodes is assumed to be 50, and are randomly distributed. The source and destination nodes are randomly selected for a TCP connection pair.

Figure 6.10 shows the throughput performance v/s TCP traffic flow of proposed AMCCR, CD-MAC and 802.11 DCF protocols. From the figure, it can be observed that, in low-noise and the low traffic environment, all the three protocols perform similarly. However, CD-MAC and non- cooperative 802.11 DCF protocols, performance degrades as the traffic load increases and the environment becomes noisy (i.e., at -83dBm). Increase in load leads to increase in channel contention, which further results in increase in packet collisions. Increase in noise level makes links more unreliable. A significant amount of packets are dropped due to collision and packet error, affecting the overall throughput. In CD-MAC, when the channel quality is harsh, the probability for a relay node to decode the packet and retransmit successfully is very low. Because of the adaptive cooperative MAC scheme in AMCCR, the packets are correctly decoded and transmitted with appropriate data transmission scheme. Interestingly, the AMCCR protocol offer better performance due to optimal relay node selection from the lesser congested area with minimal overhead and providing the highest data rate for packet transmission. Hence, the throughput enhancement in AMCCR is noticeable even in harsh channel conditions.

The end to end delay performance of the AMCCR in comparison with CD-MAC and non-cooperative 802.11 DCF protocols are as shown in Figure 6.11. It can be observed from the figure that the average end to end delay in all the protocols increases with increase in traffic flow and it becomes worse in harsh environmental conditions. The performance of 802.11 DCF is very poor because it uses the fixed transmission rate. Due to cooperative diversity, CD-MAC performs better compared to 802.11 DCF. But AMCCR significantly performs better. This is mainly because of the adaptive cooperative approach in AMCCR, that uses dynamic data rate adaptation technique. This allows more packets to be transmitted faster via the relay node, reducing the

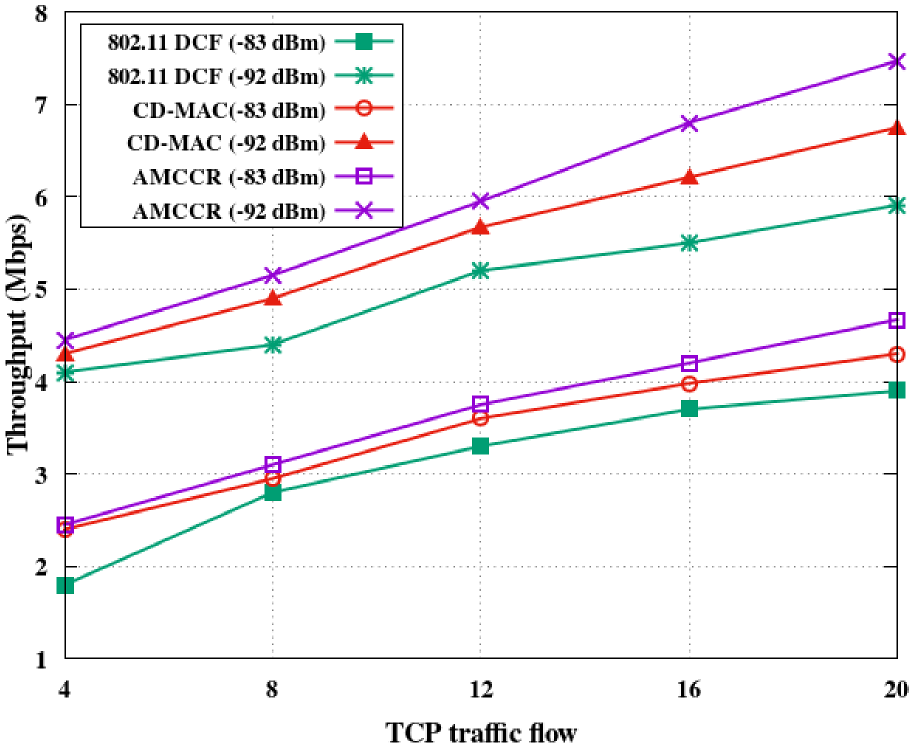


FIGURE 6.10: Throughput v/s TCP traffic flow

overall transmission time duration. In AMCCR, the network environment is accurately captured, which helps in selecting the optimal transmission mode, relay node selection at each hop and appropriate data rate for transmission. This contributes in achieving a faster end to end packet delivery.

Figure 6.12 shows the network lifetime performance as a function of increasing traffic load. It is an important performance metric for resource constrained MANETs. The nodes in the network participate in data transmission of neighboring nodes, other than their own transmissions. Thus, the network lifetime depends on the energy consumption during both phases. In AMCCR, the network lifetime is computed with respect to energy consumed per unit time. Thus the lifetime of a any node V_i is computed as $L_{V_i} = \frac{E_{initial}}{E_{consumed}}$, where $E_{consumed}$ is energy consumed per unit time. The total network lifetime is computed as $NL = min_{V_i \epsilon N}(L_{V_i})$, where N is total nodes in the network. It can be observed from the figure that the performance of AMCCR protocol is superior compared to other two protocols. Due to lack of cooperative communication in 802.11 DCF lead to poor network lifetime performance as most of the data traffic confined to a particular route leading to faster energy depletion of

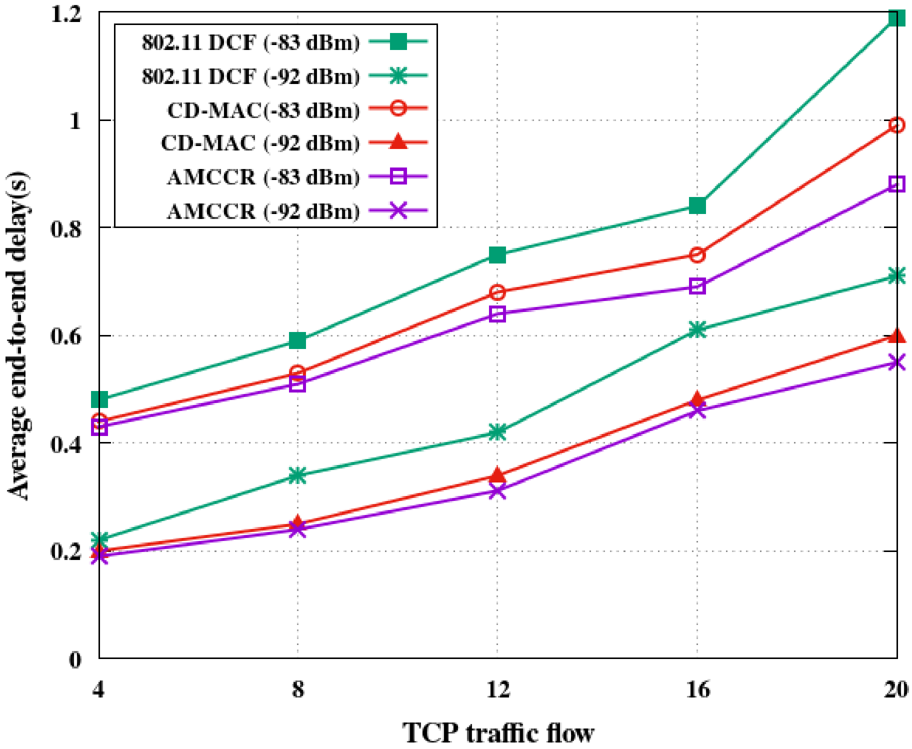


FIGURE 6.11: Average end-to-end delay v/s TCP traffic flow

nodes. Performance of CD-MAC is poor compared to AMCCR as it does not consider residual energy factors while choosing the relay node. The main reason for best performance of AMCCR protocol is that it not only selects energy efficient relay node, but also uses the relay node with appropriate data rate by instantaneously accessing channel conditions. The AMCCR also ensures uniform energy load distribution over the nodes in the network.

6.4.1.2 Effect of Varying Network Density

The proposed protocol is also evaluated by varying the network density from 20 to 100 nodes at two different noise level -92 dBm and -83 dBm. It is assumed that the number of TCP traffic flows between the randomly selected connection pair is 10.

Figure 6.13 shows the network throughput of three protocols versus node density under two noise levels. The throughput performance of AMCCR is better compared to other protocols because the protocol dynamically adapts to different channel conditions. Also, when the node density is higher, the probability that the relay nodes located in the

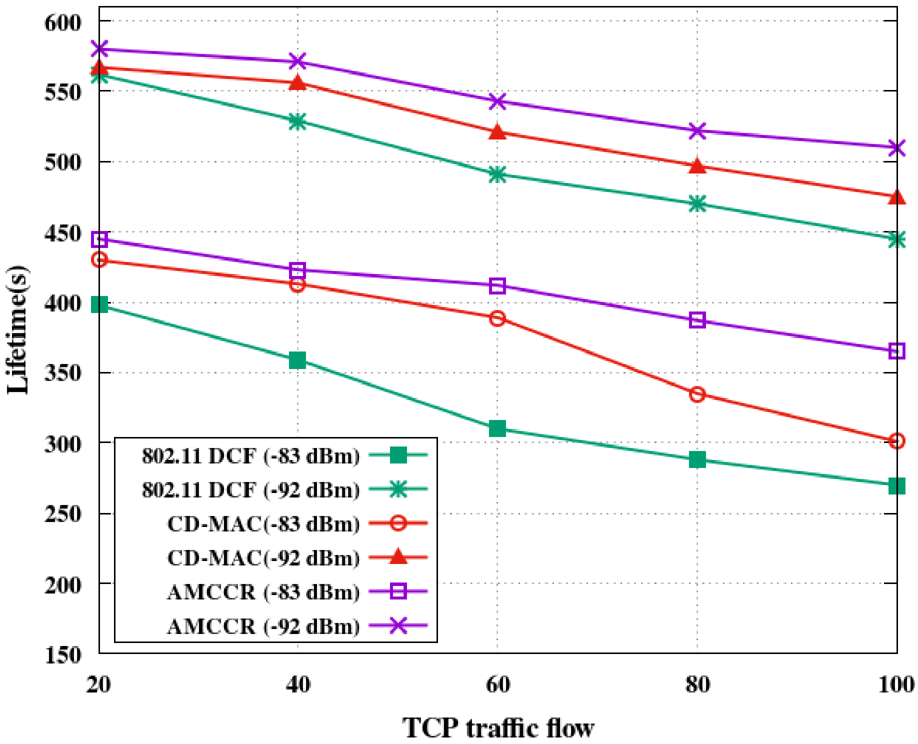


FIGURE 6.12: Network lifetime v/s TCP traffic flow

ideal positions is higher. This results in better selection of relay node for cooperative communication with higher data rates.

Figure 6.14 shows the graph of average end-to-end delay of three protocols versus node density under two noise levels. It is clear from the figure that as node density increases, end to end delay surges. This is mainly due to increased channel interference. AMCCR performs better compared to its counterparts because of its adaptive MAC scheme for selection of transmission mode, availability of multiple relay nodes at ideal position and judicious selection of one of them, and using higher data rate to transmit maximum data.

Figure 6.15 shows the network lifetime performance. From the figure it can be observed that, as the node density increases the network lifetime curve in the graph increases. The network lifetime of the proposed AMCCR performs comparatively better than the other protocols, especially in the harsh environment. As the node density increases, more the relay nodes are available with a higher energy level. The AMCCR has the longest lifetime as it optimally selects the nodes with sufficient energy level. Thus, it can be

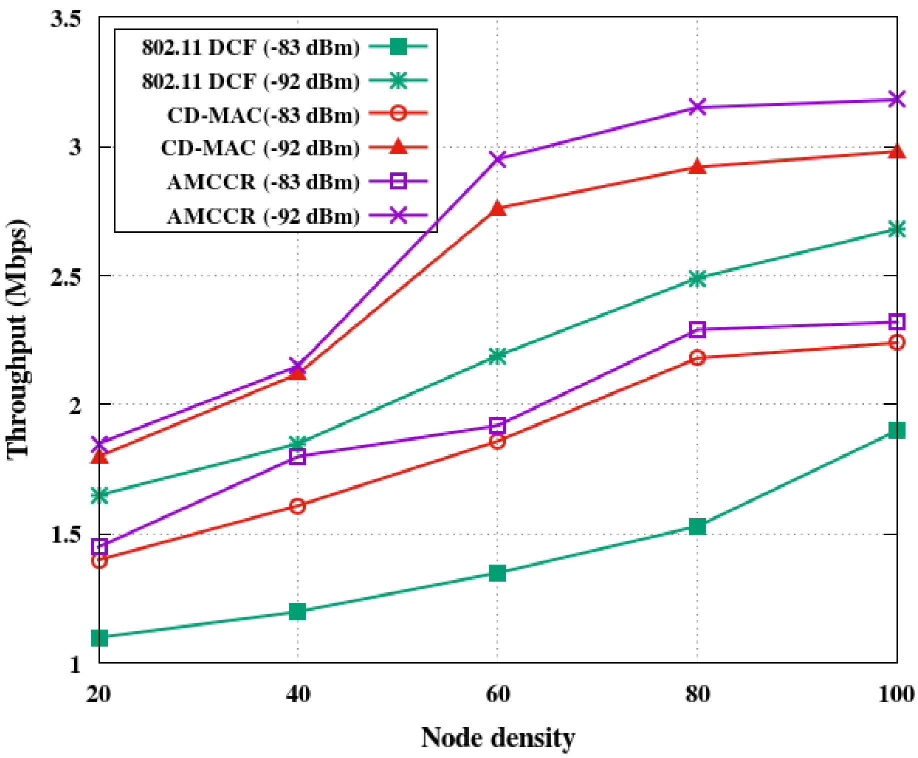


FIGURE 6.13: Throughput v/s Node density

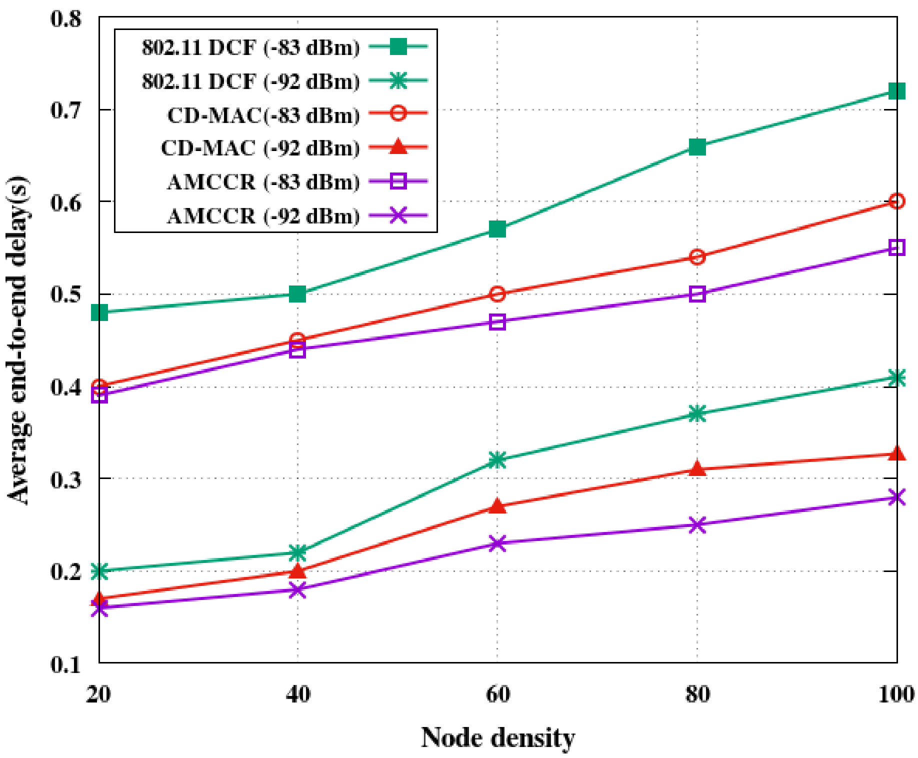


FIGURE 6.14: Average end-to-end delay v/s Node density

concluded that AMCCR is more suitable for the networks with fairly high node density and harsh environmental conditions.

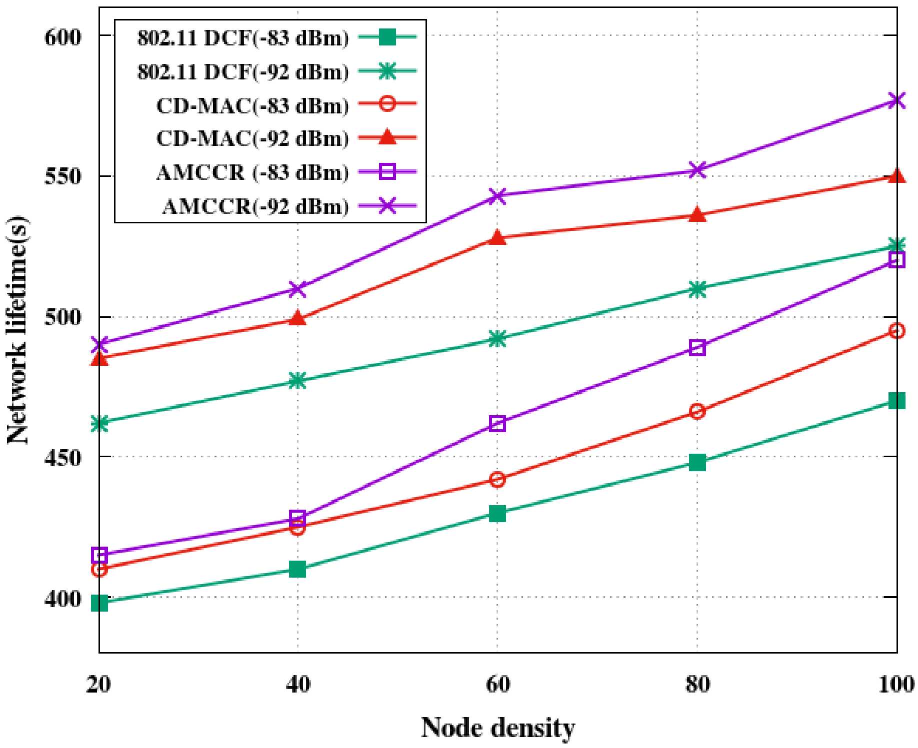


FIGURE 6.15: Network lifetime v/s Node density

6.4.2 Performance Evaluation of Cooperative Coding Scheme

To evaluate cooperative coding schemes, the AMCCR protocol is compared with CODE protocol and non-cooperative scheme 802.11 DCF. The performance of AMCCR protocol is analyzed by varying the number of UDP based concurrent VOIP traffic at two different noise levels -92dBm and -83dBm. The voice traffic is encoded into a VoIP flow using ITU-T G.711 [115], with average source bit rate of 64 kbps and the packet size of 160 bytes. The number of nodes is assumed to be 50. We assume the number of flows varies from 4 to 20, with the source and destination nodes randomly selected for a UDP connection pair.

Figure 6.16 shows the analyses of the throughput performance of AMCCR in comparison with other two protocols with varying UDP traffic. From the figure it can be observed that, when the number of VoIP calls are less, there is marginal difference in the performance of all three compared protocols. But as the number of calls increases, the proposed AMCCR shows significant improved throughput compared to CODE. When the size of the packet is small, the CODE protocol ignores the cooperation, as overhead is more compare to amount of data that is to be transferred. Whereas, AMCCR considers the benefits that can be provided by using both cooperation and network coding at the same time, while deciding whether to use cooperation or not. Being a non cooperative method, the performance of 802.11 DCF is the poor as expected.

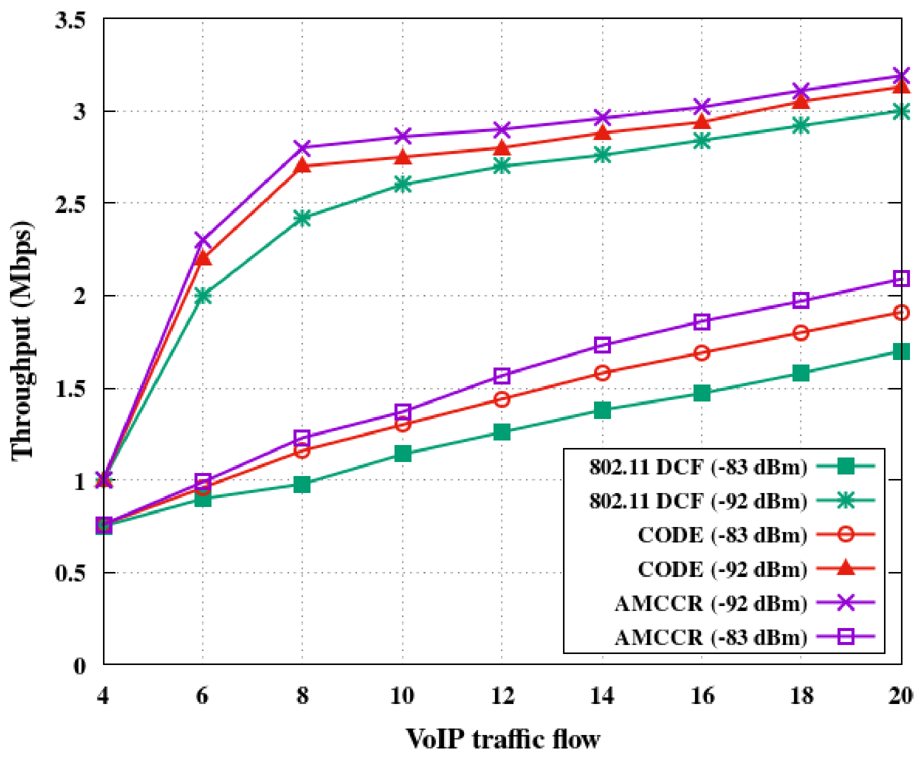


FIGURE 6.16: Throughput v/s VoIP traffic flow

Figure 6.17 shows the delay performance curve for all three protocols against the UDP VOIP traffic. The delay increases with the increase in traffic and a deteriorating environment. The AMCCR outperforms CODE with 5-10% reduction in delay. The adaptive decisions on transmission mode and best relay selections at each hop in AMCCR provides the reliability and makes it more robust to link failures, thus enhancing the delay performance.

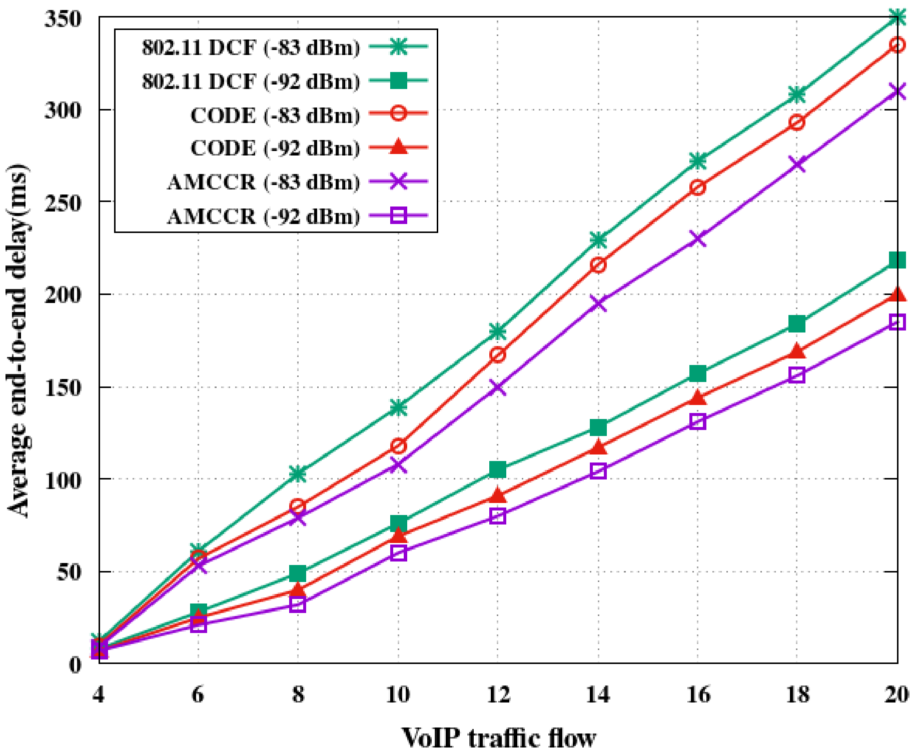


FIGURE 6.17: Average end-to-end delay v/s VoIP traffic flow

6.5 Conclusion

This chapter proposes Multi-QoS Cross layer Cooperative Routing protocol, AMCCR for multi-hop ad hoc networks. Energy aware routing protocol, AMCCR exploits cross layer communication between PHY, MAC, and network layers to achieve cooperative diversity. AMCCR uses distributed, best relay selection scheme based on multi QoS metric test. AMCCR uses an adaptive MAC cooperative mechanism to select the transmission mode, i.e., direct or relay mode so that a low data rate host can be assisted by high data rate relay node. AMCCR protocol also supports dual-hop half-duplex communication via selected relay node by coding technique. The extensive simulation conducted shows that in cooperative mode, AMCCR shows significantly better performance in terms of throughput, delay and network lifetime compared to CD-MAC and 802.11 DCF even under the harsh environment. The proposed AMCCR also outperform the CODE protocol as it exploits both cooperation and coding together to enhance system performance.

Bibliography

[1] B. P. Crow, I. Widjaja, J. G. Kim, and P. T. Sakai, "IEEE 802.11 wireless local area networks," *IEEE Communications Magazine*, vol. 35, no. 9, pp. 116–126, 1997.

[2] C.-K. Toh, *Ad hoc mobile wireless networks*. Prentice Hall PTR, 2002.

[3] I. Chlamtac, M. Conti, and J. J.-N. Liu, "Mobile ad hoc networking: imperatives and challenges." *Ad Hoc Networks*, vol. 1, no. 1, pp. 13–64, 2003.

[4] C. S. R. Murthy and B. Manoj, *Ad Hoc Wireless Networks: Architectures and Protocols*. Prentice Hall PTR, 2004.

[5] S. Chakrabarti and A. Mishra, "The handbook of ad hoc wireless networks," M. Ilyas and R. C. Dorf, Eds., 2003, ch. Quality of Service in Mobile Ad Hoc Networks, pp. 41–69.

[6] S. Misra, I. Woungang, and S. C. Misra, *Guide to Wireless Ad Hoc Networks*, 1st ed. Springer Publishing Company, Incorporated, 2009.

[7] N. Abramson, "THE ALOHA SYSTEM: Another alternative for computer communications," in *Proceedings of the November 17-19, Fall Joint Computer Conference*, 1970, pp. 281–285.

[8] E. Gregori, G. Anastasi, and S. Basagni, Eds., *Advanced Lectures on Networking, NETWORKING 2002*, ser. Lecture Notes in Computer Science, vol. 2497, 2002.

[9] A. Boukerche, B. Turgut, N. Aydin, M. Z. Ahmad, L. Bölöni, and D. Turgut, "Routing protocols in ad hoc networks: A survey." *Computer Networks*, vol. 55, no. 13, pp. 3032–3080, 2011.

[10] K. Chen and S. H. Shah, "Cross-layer design for data accessibility in mobile ad hoc networks," *Wireless Personal Communications*, vol. 21, no. 8, pp. 49–76, 2002.

[11] P. Sinha, R. Sivakumar, and V. Bharghavan, "CEDAR: a core-extraction distributed ad hoc routing algorithm," in *Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. INFOCOM '99.*, vol. 1, 1999, pp. 202–209.

[12] C. E. Perkins and E. M. Belding-Royer, "Quality of Service for Ad hoc On-Demand Distance Vector Routing," Internet Engineering Task Force, Internet-Draft draft-perkins-manet-aodvqos-01, 2001, work in Progress.

[13] Q. Xue and A. Ganz, "Ad hoc qos on-demand routing (AQOR) in mobile ad hoc networks," *J. Parallel Distrib. Comput.*, vol. 63, no. 2, pp. 154–165, 2003.

[14] W. E. Castellanos, J. C. Guerri, and P. Arce, "A qos-aware routing protocol with adaptive feedback scheme for video streaming for mobile networks," *Comput. Commun.*, vol. 77, no. C, pp. 10–25, 2016.

[15] "IEEE draft standard for information technology-telecommunications and information exchange between systems-local and metropolitan networks-specific requirements-part ii: Wireless lan medium access control (MAC) and physical layer (PHY) specifications-amendment 8: IEEE 802.11 wireless network management," *IEEE P802.11v/D15.0, September 2010*, pp. 1–428, 2010.

[16] N. S. Vo, T. Q. Duong, H. J. Zepernick, and M. Fiedler, "A cross-layer optimized scheme and its application in mobile multimedia networks with qos provision," *IEEE Systems Journal*, vol. 10, no. 2, pp. 817–830, 2016.

[17] A. Kamerman and L. Monteban, "WaveLANII: A high-performance wireless LAN for the unlicensed band," *Bell Labs Technical Journal*, vol. 2, no. 3, pp. 118–133, 1997.

[18] S. Kim, L. Verma, S. Choi, and D. Qiao, “Collision-aware rate adaptation in multi-rate WLANs: Design and implementation,” *Computer Networks*, vol. 54, no. 17, pp. 3011–3030, 2010.

[19] G. Holland, N. Vaidya, and P. Bahl, “A rate-adaptive mac protocol for multi-hop wireless networks,” in *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom ’01, 2001, pp. 236–251.

[20] X. Chen, H. M. Jones, and A. D. S. Jayalath, “Congestion-aware routing protocol for mobile ad hoc networks: CARM,” in *IEEE 66th Vehicular Technology Conference*, 2007, pp. 21–25.

[21] B. Tripathi, P. Agrawal, V. Khare, and G. R. Murthy, “Scalability study of the CADV (congestion aware distance vector) routing protocol in mobile ad-hoc networks,” in *International Conference on Signal Processing and Communications*, 2004, pp. 496–500.

[22] J.-H. Song, V. Wong, and V. C. M. Leung, “Load-aware on-demand routing (LAOR) protocol for mobile ad hoc networks,” in *The 57th IEEE Semiannual Vehicular Technology Conference*, vol. 3, 2003, pp. 1753–1757.

[23] S. Mangold, S. Choi, G. R. Hiertz, O. Klein, and B. Walke, “Analysis of IEEE 802.11e for QoS support in wireless LANs,” *IEEE Wireless Communications*, vol. 10, no. 6, pp. 40–50, 2003.

[24] T. Nilsson and J. Farooq, “A novel MAC scheme for solving the qos parameter adjustment problem in IEEE 802.11e EDCA,” in *International Symposium on a World of Wireless, Mobile and Multimedia Networks*, 2008, pp. 1–9.

[25] F. Peng and V. C. M. Leung, “Enhancing fairness and throughput of TCP in heterogeneous wireless access networks,” *IJWIN*, vol. 14, no. 3, pp. 225–236, 2007.

[26] D. J. Leith, P. Clifford, D. Malone, and A. Ng, “TCP fairness in. 802.11e WLANs,” *IEEE Communications Letters*, vol. 9, no. 11, pp. 964–966, 2005.

[27] E. Altman and T. Jiménez, "Novel delayed ACK techniques for improving TCP performance in multihop wireless networks," in *8th International Conference on Personal Wireless Communications*, 2003, pp. 237–250.

[28] C. Schurgers and M. B. Srivastava, "Energy efficient routing in wireless sensor networks," in *2001 MILCOM Proceedings Communications for Network-Centric Operations: Creating the Information Force*, vol. 1, 2001, pp. 357–361.

[29] K. J. R. Liu, A. K. Sadek, W. Su, and A. Kwasinski, *Cooperative Communications and Networking*. New York, NY, USA: Cambridge University Press, 2009.

[30] J. Boyer, D. D. Falconer, and H. Yanikomeroglu, "Multihop diversity in wireless relaying channels," *IEEE Transactions on Communications*, vol. 52, no. 10, pp. 1820–1830, 2004.

[31] A. Scaglione, D. L. Goeckel, and J. N. Laneman, "Cooperative communications in mobile ad hoc networks," *IEEE Signal Processing Magazine*, vol. 23, no. 5, pp. 18–29, 2006.

[32] S. S. N, C.-T. Chou, and M. Ghosh, "Cooperative communication mac (CMAC) - a new mac protocol for next generation wireless lans," in *2005 International Conference on Wireless Networks, Communications and Mobile Computing*, vol. 1, 2005, pp. 1–6.

[33] P. Liu, Z. Tao, S. Narayanan, T. Korakis, and S. S. Panwar, "CoopMAC: A cooperative MAC for wireless LANs," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 340–354, 2007.

[34] J. Loo, J. L. Mauri, and J. H. Ortiz, *Mobile Ad Hoc Networks: Current Status and Future Trends*, 1st ed. Boca Raton, FL, USA: CRC Press, Inc., 2011.

[35] R. Braden, D. Clark, and S. Shenker, "Integrated Services in the Internet Architecture: an Overview," RFC 1633 (Informational), Internet Engineering Task Force, 1994.

[36] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An architecture for differentiated service," United States, 1998.

[37] H. Xiao, W. K. G. Seah, A. Lo, and K. C. Chua, "A flexible quality of service model for mobile ad-hoc networks," in *IEEE 51st Vehicular Technology Conference Proceedings. VTC2000.*, vol. 1, 2000, pp. 445–449 vol.1.

[38] C.-H. Yeh, H. T. Mouftah, and H. Hassanein, "Signaling and qos guarantees in mobile ad hoc networks," in *IEEE International Conference on Communications. Conference Proceedings*, vol. 5, 2002, pp. 3284–3290.

[39] A. Pantaleo, M. Tornatore, A. Pattavina, C. Raffaelli, and F. Callegati, "Considerations on in-band and out-of-band signalling constraints in OBS networks," in *2007 9th International Conference on Transparent Optical Networks*, vol. 4, 2007, pp. 113–116.

[40] Y. Xiao and H. Li, "Evaluation of distributed admission control for the IEEE 802.11e EDCA," *IEEE Communications Magazine*, vol. 42, no. 9, pp. S20–S24, 2004.

[41] L. Xiao, J. Wang, and M. Nahrstedt, "The enhanced ticket-based routing algorithm," in *2002 IEEE International Conference on Communications. Conference Proceedings*, vol. 4, 2002, pp. 2222–2226.

[42] S. H. Shah and K. Nahrstedt, "Predictive location-based qos routing in mobile ad hoc networks," in *2002 IEEE International Conference on Communications. Conference Proceedings*, vol. 2, 2002, pp. 1022–1027.

[43] S.-B. Lee, G.-S. Ahn, X. Zhang, and A. T. Campbell, "INSIGNIA," *J. Parallel Distrib. Comput.*, vol. 60, no. 4, pp. 374–406, 2000.

[44] G.-S. Ahn, A. T. Campbell, A. Veres, and L.-H. Sun, "SWAN: service differentiation in stateless wireless ad hoc networks," in *Proceedings.Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 2, 2002, pp. 457–466.

[45] C. E. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers," in *Proceedings of the Conference on Communications Architectures, Protocols and Applications*, ser. SIGCOMM '94, 1994, pp. 234–244.

[46] S. Murthy and J. J. Garcia-Luna-Aceves, "An efficient routing protocol for wireless networks," *Mobile Networks and Applications*, vol. 1, no. 2, pp. 183–197, 1996.

[47] T.-W. Chen and M. Gerla, "Global state routing: a new routing scheme for ad-hoc wireless networks," in *IEEE International Conference on Communications. ICC 98.*, vol. 1, 1998, pp. 171–175.

[48] "Optimized link state routing protocol (OLSR)," 2003. [Online]. Available: http://portal.acm.org/citation.cfm?id=RFC3626

[49] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing*. Kluwer Academic Publishers, 1996, pp. 153–181.

[50] C. E. Perkins and E. M. Royer, "Ad-hoc on-demand distance vector routing," in *Wireless Mobile Computing Systems and Applications, 1999. Proceedings. Second IEEE Workshop on*, 1999, pp. 90–100.

[51] V. D. Park and M. S. Corson, "A highly adaptive distributed routing algorithm for mobile wireless networks," in *Proceedings of the INFOCOM '97. Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Driving the Information Revolution*, 1997, pp. 1405–1410.

[52] H. Maleki, M. Kargahi, and S. Jabbehdari, "RTLB-DSR: A load-balancing dsr based qos routing protocol in manets," in *2014 4th International Conference on Computer and Knowledge Engineering (ICCKE)*, 2014, pp. 728–735.

[53] V. C. Frias, G. D. Delgado, A. Z. Ayala, and M. A. Igartua, "MM-DSR: Multipath qos routing for multiple multimedia sources over ad hoc mobile networks," *IEEE Latin America Transactions*, vol. 5, no. 6, pp. 448–456, 2007.

[54] A. R. Hilal, A. E. Nahas, and A. Bashandy, "RT-TORA: A TORA modification for real-time interactive applications," in *2008 Canadian Conference on Electrical and Computer Engineering*, 2008, pp. 1403–1406.

[55] Z. J. Haas, M. R. Pearlman, and P. Samar, "The Zone Routing Protocol (ZRP) for Ad Hoc Networks," Published Online, IETF MANET Working Group, INTERNET-DRAFT, 2002.

[56] R. Geng, Z. Li, and L. Song, "AQMP: An adaptive QoS MAC protocol based on ieee802.11 in ad hoc networks," in *5th International Conference on Wireless Communications, Networking and Mobile Computing*, 2009, pp. 1–4.

[57] K. S. Shin, S. Y. Yun, and D. H. Cho, "Multi-channel mac protocol for qos support in ad-hoc network," in *2011 IEEE Consumer Communications and Networking Conference (CCNC)*, 2011, pp. 975–976.

[58] F. D. Rango, F. Guerriero, and P. Fazio, "Link-stability and energy aware routing protocol in distributed wireless networks (laer," *IEEE Transactions on Parallel and Distributed Systems*.

[59] C. R. Lin, "On-demand qos routing in multihop mobile networks," in *Twentieth Annual Joint Conference of the IEEE Computer and Communications Society. IEEE INFOCOM 2001.*, vol. 3, 2001, pp. 1735–1744.

[60] E. S. Crawley, R. Nair, D. B. Rajagopalan, and H. J. Sandick, "A Framework for QoS-based Routing in the Internet," RFC 2386, 1998.

[61] B. Zhang and H. T. Mouftah, "QoS routing for wireless ad hoc networks: problems, algorithms, and protocols," *IEEE Communications Magazine*, vol. 43, no. 10, pp. 110–117, 2005.

[62] E. Meyer, "ARPA Network protocol notes," Internet Engineering Task Force, RFC 0046, 1970.

[63] L. J. Miller, "The ISO reference model of open systems interconnection: A first tutorial," in *Proceedings of the ACM '81 Conference*, 1981, pp. 283–288.

[64] H. Zimmermann, "OSI reference model - the iso model of architecture for open systems interconnection," *IEEE Transactions on Communications*, vol. 28, no. 4, pp. 425–432, 1980.

[65] V. T. Raisinghani and S. Iyer, "Cross-layer design optimizations in wireless protocol stacks," *Comput. Commun.*, vol. 27, no. 8, pp. 720–724, 2004.

[66] Q. Wang and M. A. Abu-Rgheff, "Cross-layer signalling for next-generation wireless systems," in *2003 IEEE Wireless Communications and Networking*, vol. 2, 2003, pp. 1084–1089.

[67] S. Faruque, *Automatic Repeat Request (ARQ)*. Cham: Springer International Publishing, 2016, pp. 17–35.

[68] S. Jiang, D. He, and J. Rao, "A prediction-based link availability estimation for routing metrics in manets," *IEEE/ACM Transactions on Networking*, vol. 13, no. 6, pp. 1302–1312, 2005.

[69] C. E. Perkins and E. M. Royer, "Ad-hoc on-demand distance vector routing," in *Proceedings of the Second IEEE Workshop on Mobile Computer Systems and Applications*. IEEE Computer Society, 1999, pp. 90–.

[70] Y. Ping and W. Ying, "A revised aodv protocol with qos for mobile ad hoc network," in *2009 2nd IEEE International Conference on Computer Science and Information Technology*, 2009, pp. 241–244.

[71] J. N. Boshoff and A. S. J. Helberg, "Improving qos for real-time multimedia traffic in ad-hoc networks with delay aware multi-path routing," in *2008 Wireless Telecommunications Symposium*, 2008, pp. 1–8.

[72] Z. Ye, S. V. Krishnamurthy, and S. K. Tripathi, "A routing framework for providing robustness to node failures in mobile ad hoc networks," *Ad Hoc Networks*, vol. 2, no. 1, pp. 87 – 107, 2004.

[73] R. Asokan and A. M. Natarajan, "Performance evaluation of energy and delay aware quality of service (QoS) routing protocols in mobile adhoc networks," *IJBDCN*, vol. 4, no. 1, pp. 52–63, 2008.

[74] S. R. Medidi and K. H. Vik, "Quality of service-aware source-initiated ad-hoc routing," in *2004 First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks*, 2004, pp. 108–117.

[75] L. Doss, Robin Pan, "A case for cross layer design: the impact of physical layer properties on routing protocol performance in MANETs," in *Proceedings. 2006 Australian Telecommunication Networks and Applications Conference*, 2006, pp. 409–413.

[76] R. Biradar, S. Manvi, and M. Reddy, "Link stability based multicast routing scheme in MANET," *Comput. Netw.*, vol. 54, no. 7, pp. 1183–1196, 2010.

[77] T. Girici and A. Ephremides, "Joint routing and scheduling metrics for ad hoc wireless networks," in *Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers, 2002.*, vol. 2, 2002, pp. 1155–1159.

[78] Z. El-Bazzal, K. El-Ahmadieh, Z. Merhi, M. Nahas, and A. Haj-Ali, "A cross layered routing protocol for ad hoc networks," in *2012 International Conference on Information Technology and e-Services*, 2012, pp. 1–6.

[79] J. Chen, Z. Li, J. Liu, and Y. Kuo, "Qos multipath routing protocol based on cross layer design for ad hoc networks," in *2011 International Conference on Internet Computing and Information Services*, 2011, pp. 261–264.

[80] S. M. Adam and R. Hassan, "Delay aware reactive routing protocols for qos in manets: a review," *Journal of Applied Research and Technology*, vol. 11, no. 6, pp. 844 – 850, 2013.

[81] T. T. Son, H. L. Minh, G. Sexton, N. Aslam, and R. Boubezari, "A new mobility, energy and congestion aware routing scheme for MANETs," in *2014 9th International Symposium on Communication Systems, Networks Digital Sign*, 2014, pp. 771–775.

[82] M. K. Gulati and K. Kumar, "Stable energy efficient qos based congestion and delay aware routing (SEQCDR) protocol for MANETs," in *2015 International*

Conference on Communications and Signal Processing (ICCSP), 2015, pp. 505–511.

[83] Y. Lu, W. Wang, Y. Zhong, and B. Bhargava, "Study of distance vector routing protocols for mobile ad hoc networks," in *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, 2003. (PerCom 2003).*, 2003, pp. 187–194.

[84] D. A. Tran and H. Raghavendra, "Congestion adaptive routing in mobile ad hoc networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 17, no. 11, pp. 1294–1305, 2006.

[85] S. Kang, C. Yu, C. R. Das, and G. Cao, "Path-centric on-demand rate adaptation for mobile ad hoc networks," in *Proceedings of 18th International Conference on Computer Communications and Networks*, 2009, pp. 1–6.

[86] P. Chevillat, J. Jelitto, A. N. Barreto, and H. L. Truong, "A dynamic link adaptation algorithm for IEEE 802.11 a wireless lans," in *IEEE International Conference on Communications, 2003*, vol. 2, 2003, pp. 1141–1145.

[87] Y. Seok, J. Park, and Y. Choi, "Multi-rate aware routing protocol for mobile ad hoc networks," in *The 57th IEEE Semiannual Vehicular Technology Conference, 2003*, vol. 3, 2003, pp. 1749–1752 vol.3.

[88] M. Lacage, M. H. Manshaei, and T. Turletti, "IEEE 802.11 rate adaptation: A practical approach," in *Proceedings of the 7th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, 2004, pp. 126–134.

[89] Y. Xi, B. s. Kim, J. b. Wei, and Q. y. Huang, "Adaptive multirate auto rate fallback protocol for IEEE 802.11 wlans," in *2006 IEEE Military Communications conference*, 2006, pp. 1–7.

[90] M. K. Marina and S. R. Das, "Ad hoc on-demand multipath distance vector routing," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 6, no. 3, pp. 92–93, 2002.

[91] Z. Ye, S. V. Krishnamurthy, and S. K. Tripathi, "A framework for reliable routing in mobile ad hoc networks," in *Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies, IEEE INFOCOM 2003.*, vol. 1, 2003, pp. 270–280.

[92] L. Reddeppa Reddy and S. V. Raghavan, "SMORT: Scalable multipath on-demand routing for mobile ad hoc networks," *Ad Hoc Netw.*, vol. 5, no. 2, pp. 162–188, 2007.

[93] S. J. Lee and M. Gerla, "Split multipath routing with maximally disjoint paths in ad hoc networks," in *IEEE International Conference on Communications. Conference Record. ICC 2001.*, vol. 10, 2001, pp. 3201–3205.

[94] L. Wang, Y. Shu, M. Dong, L. Zhang, and O. W. W. Yang, "Adaptive multipath source routing in ad hoc networks," in *IEEE International Conference on Communications. ICC 2001.*, vol. 3, 2001, pp. 867–871.

[95] S.-T. Sheu and J. Chen, "A novel delay-oriented shortest path routing protocol for mobile ad hoc networks," in . *IEEE International Conference on Communications. ICC 2001.*, vol. 6, 2001, pp. 1930–1934.

[96] "The Network Simulator NS-2," http://www.isi.edu/nsnam/ns/.

[97] D.-J. Deng and H.-C. Yen, "Quality-of-service provisioning system for multimedia transmission in IEEE 802.11 wireless lans," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 6, pp. 1240–1252, 2005.

[98] A. Iera, A. Molinaro, G. Ruggeri, and D. Tripodi, "Improving QoS and throughput in single- and multihop WLANs through dynamic traffic prioritization," *IEEE Network*, vol. 19, no. 4, pp. 35–44, 2005.

[99] I. Aad and C. Castelluccia, "Differentiation mechanisms for IEEE 802.11," in *Proceedings IEEE INFOCOM 2001. Conference on Computer Communications. Twentieth Annual Joint Conference of the IEEE Computer and Communications Society*, vol. 1, 2001, pp. 209–218.

[100] Y. Gao, X. Sun, and L. Dai, "Achieving optimum network throughput and service differentiation for IEEE 802.11e EDCA networks," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2013, pp. 362–367.

[101] "Ieee standard for information technology–local and metropolitan area networks–specific requirements–part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications - amendment 8: Medium access control (mac) quality of service enhancements," *IEEE Std 802.11e-2005 (Amendment to IEEE Std 802.11, 1999 Edition (Reaff 2003)*, pp. 1–212, 2005.

[102] D. C. Dobhal and S. C. Dimri, "Performance evaluation of proposed-TCP in mobile ad hoc networks (MANETs)," in *International Conference on Inventive Computation Technologies (ICICT)*, vol. 2, 2016, pp. 1–6.

[103] N. Atri and R. Goyal, "A comparative analysis of MANET routing protocols over TCP," in *International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*, 2015, pp. 712–717.

[104] U. O. Ibom, "TCP performance over MANET," in *2008 International Conference on Information Networking*, 2008, pp. 1–5.

[105] S. Lohier, Y. Ghamri-Doudane, and G. Pujolle, "Cross-layer Loss Differentiation Algorithm to Improve TCP Performances in WLANs," in *11th IFIP International Conference on Personal Wireless Communications (PWC'06)*, ser. Lecture Notes in Computer Science, vol. 4217, 2006, pp. 297–309.

[106] E. C. Park, D. Y. Kim, H. Kim, and C. H. Choi, "A cross-layer approach for per-station fairness in TCP over WLANs," *IEEE Transactions on Mobile Computing*, vol. 7, no. 7, pp. 898–911, 2008.

[107] W. Liu and Y. Fang, "Courtesy piggybacking: supporting differentiated services in multihop mobile ad hoc networks," in *IEEE INFOCOM 2004*, vol. 2, 2004, pp. 1273–1283.

[108] C. Mbarushimana and A. Shahrabi, "E-TCP: Enhanced TCP for IEEE 802.11e mobile ad hoc networks," in *2009 15th International Conference on Parallel and Distributed Systems*, 2009, pp. 632–639.

[109] I. Koukoutsidis and V. A. Siris, "802.11e EDCA protocol parameterization: A modeling and optimization study," in *2007 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*, 2007, pp. 1–9.

[110] Y. H. Kim and J. B. Suk, "Performance improvement for best-effort traffic of IEEE 802.11e qos mac in mobile ad hoc networks," in *International Conference on Wireless Communications, Networking and Mobile Computing*, 2006, pp. 1–4.

[111] J. Vuillemin, "A data structure for manipulating priority queues," *Commun. ACM*, vol. 21, no. 4, pp. 309–315, 1978.

[112] Y. Xiao, "Performance analysis of priority schemes for IEEE 802.11 and IEEE 802.11e wireless LANs," *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1506–1515, 2005.

[113] R. Lenagala and Q. A. Zeng, "Starvation prevention scheme for the IEEE 802.11e EDCF using dynamic MAC layer parameters," in *2006 Wireless Telecommunications Symposium*, 2006, pp. 1–7.

[114] M. Mishra and A. Sahoo, "An 802.11 based MAC protocol for providing qos to real time applications," in *Information Technology, (ICIT 2007). 10th International Conference on*, 2007, pp. 104–109.

[115] Y. Hiwasaki and H. Ohmuro, "ITU-T G.711.1: extending G.711 to higher-quality wideband speech," *IEEE Communications Magazine*, vol. 47, no. 10, pp. 110–116, 2009.

[116] S. G. Pease, I. W. Phillips, and L. Guan, "Adaptive intelligent middleware architecture for mobile real-time communications," *IEEE Transactions on Mobile Computing*, vol. 15, no. 3, pp. 572–585, 2016.

[117] H. Zen, D. Habibi, J. Wyatt, and I. Ahmad, "Converging voice, video and data in wlan with qos support," in *5th IFIP International Conference on Wireless and Optical Communications Networks (WOCN '08)*, 2008, pp. 1–5.

[118] E. c. Park, D. y. Kim, C. h. Choi, and J. So, "Improving quality of service and assuring fairness in wlan access networks," *IEEE Transactions on Mobile Computing*, vol. 6, no. 4, pp. 337–350, 2007.

[119] J. Shijie, X. Changqiao, G. M. Muntean, G. Jianfeng, and Z. Hongke, "Cross-layer and one-hop neighbour-assisted video sharing solution in mobile ad hoc networks," *China Communications*, vol. 10, no. 6, pp. 111–126, 2013.

[120] L. H. II. and R. Tafazolli, "Qos-aware routing and admission control in shadow-fading environments for multirate manets," *IEEE Transactions on Mobile Computing*, vol. 10, no. 5, pp. 622–637, 2011.

[121] H. B. Pathan, S. V. P, and S. R. K, "Qos performance of IEEE 802.11 in mac and phy layer using enhanced oar algorithm," *Indian Journal of Science and Technology*, vol. 10, no. 9, 2017.

[122] "IEEE draft standard for information technology telecommunications and information exchange between systems local and metropolitan area networks specific requirements part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," *IEEE Std C95.2-1999*.

[123] P. Serrano, A. Banchs, P. Patras, and A. Azcorra, "Optimal configuration of 802.11e EDCA for real-time and data traffic," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 5, pp. 2511–2528, 2010.

[124] Y. Xiao and H. Li, "Voice and video transmissions with global data parameter control for the IEEE 802.11e enhance distributed channel access," *IEEE Transactions on Parallel and Distributed Systems*, vol. 15, no. 11, pp. 1041–1053, 2004.

[125] L. Romdhani, Q. Ni, and T. Turletti, "Adaptive EDCF: enhanced service differentiation for ieee 802.11 wireless ad-hoc networks," in *IEEE Wireless Communications and Networking. WCNC 2003.*, vol. 2, 2003, pp. 1373–1378.

[126] S. Prasetya, B. Rahmat, and E. Susanto, "Quality of service improvement with 802.11e edca scheme using enhanced adaptive contention window algorithm," in *2015 IEEE International Conference on Communication, Networks and Satellite (COMNESTAT)*, 2015, pp. 80–85.

[127] I. Tinnirello, G. Bianchi, and L. Scalia, "Performance evaluation of differentiated access mechanisms effectiveness in 802.11 networks," in *Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE*, vol. 5, 2004, pp. 3007–3011.

[128] D. C. Hogg, "Fun with the friis free-space transmission formula," *IEEE Antennas and Propagation Magazine*, vol. 35, no. 4, pp. 33–35, 1993.

[129] P. Basu, N. Khan, and T. D. C. Little, "A mobility based metric for clustering in mobile ad hoc networks," in *Proceedings 21st International Conference on Distributed Computing Systems Workshops*, 2001, pp. 413–418.

[130] J.-G. Kim and C. C. J. Kuo, "MPEG-4 video codec ip design with a configurable embedded processor," in *Proceedings of the 2003 International Symposium on Circuits and Systems*, vol. 2, 2003, pp. II–776–II–779.

[131] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, 2004.

[132] J. Bai, Y. Sun, and C. Phillips, "CRRP: A cooperative relay routing protocol for iot networks," in *IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2016, pp. 1–6.

[133] L. Li, D. Kaleshi, and Z. Fan, "CORELA: A cooperative relaying enhanced link adaptation algorithm for IEEE 802.11 wlans," in *8th International Symposium on Wireless Communication Systems*, 2011, pp. 382–386.

[134] S. Moh, C. Yu, S. M. Park, H. N. Kim, and J. Park, "CD-MAC: Cooperative diversity mac for robust communication in wireless ad hoc networks," in *IEEE International Conference on Communications. ICC 2007.*, 2007, pp. 3636–3641.

[135] K. Tan, Z. Wan, H. Zhu, and J. Andrian, "CODE: Cooperative medium access for multirate wireless ad hoc network," in *4th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, 2007, pp. 1–10.

[136] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. part i. system description," *IEEE Transactions on Communications*, vol. 51, no. 11, pp. 1927–1938, 2003.

[137] A. M. Akhtar, M. R. Nakhai, and A. H. Aghvami, "On the use of cooperative physical layer network coding for energy efficient routing," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1498–1509, 2013.

[138] T. Tat and A. Girinsky, "Evolution of store-and-forward techniques," *IEEE Transactions on Communications*, vol. 22, no. 9, pp. 1297–1301, 1974.

[139] D. F. Tseng and Y. Y. Lin, "Amplify-and-forward cooperative relaying for single-carrier block transmission," in *IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, 2009, pp. 3124–3128.

[140] K. Lee, "Novel decoding algorithms for decode-and-forward cooperative beamforming systems," *IET Communications*, vol. 8, no. 17, pp. 3012–3018, 2014.

[141] A. Azgin, Y. Altunbasak, and G. AlRegib, "Cooperative mac and routing protocols for wireless ad hoc networks," in *IEEE Global Telecommunications Conference. GLOBECOM '05.*, vol. 5, 2005, pp. 6 pp.–2859.

[142] N. Agarwal, D. ChanneGowda, L. N. Kannan, M. Tacca, and A. Fumagalli, *IEEE 802.11b Cooperative Protocols: A Performance Study*, 2007, pp. 415–426.

[143] Q. Zhang and Y. Q. Zhang, "Cross-layer design for qos support in multihop wireless networks," *Proceedings of the IEEE*, vol. 96, no. 1, pp. 64–76, 2008.

[144] B. Gui, L. Dai, and L. J. Cimini, "Routing strategies in multihop cooperative networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 2, pp. 843–855, 2009.

[145] H. Xu, L. Huang, C. Qiao, Y. Zhang, and Q. Sun, "Bandwidth-power aware cooperative multipath routing for wireless multimedia sensor networks," *IEEE Transactions on Wireless Communications*, vol. 11, no. 4, pp. 1532–1543, 2012.

[146] A. M. Akhtar, M. R. Nakhai, and A. H. Aghvami, "Power aware cooperative routing in wireless mesh networks," *IEEE Communications Letters*, vol. 16, no. 5, pp. 670–673, 2012.

[147] S. Chen, M. Huang, Y. Li, Y. Zhu, and Y. Wang, "Energy-balanced cooperative routing in multihop wireless ad hoc networks," in *IEEE International Conference on Communications. ICC 2012.*, 2012, pp. 307–311.

[148] A. S. Ibrahim, Z. Han, and K. J. R. Liu, "Distributed energy-efficient cooperative routing in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 10, pp. 3930–3941, 2008.

[149] R. Lu, X. Lin, H. Zhu, X. Liang, and X. Shen, "Becan: A bandwidth-efficient cooperative authentication scheme for filtering injected false data in wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 1, pp. 32–43, 2012.

[150] W. Li and L. Kai, "An throughput-optimized cooperative routing protocol in ad hoc network," in *3rd IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications*, 2009, pp. 1255–1258.

[151] M. Mohammadi, H. A. Suraweera, and X. Zhou, "Outage probability of wireless ad hoc networks with cooperative relaying," *CoRR*, vol. abs/1301.5687, 2013.

[152] *Channel models for digital transmission*. Boston, MA: Springer US, 2005, pp. 19–35.

[153] W. Su, A. K. Sadek, and K. J. Ray Liu, "Cooperative communication protocols in wireless networks: Performance analysis and optimum power allocation," *Wireless Personal Communications*, vol. 44, no. 2, pp. 181–217, 2008.

[154] D. G. Brennan, "Linear diversity combining techniques," *Proceedings of the IEEE*, vol. 91, no. 2, pp. 331–356, 2003.

[155] M. Khalid, Y. Wang, I. h. Ra, and R. Sankar, "Two-relay-based cooperative mac protocol for wireless ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 7, pp. 3361–3373, 2011.

[156] B. Guo, Q. Guan, F. R. Yu, S. Jiang, and V. C. M. Leung, "Energy-efficient topology control with selective diversity in cooperative wireless ad hoc networks: A game-theoretic approach," *IEEE Transactions on Wireless Communications*, vol. 13, no. 11, pp. 6484–6495, 2014.

[157] X. Liang, I. Balasingham, and V. C. M. Leung, "Cooperative communications with relay selection for qos provisioning in wireless sensor networks," in *GLOBECOM 2009 - 2009 IEEE Global Telecommunications Conference*, 2009, pp. 1–8.

[158] A. Ephremides, S. Jaggi, T. Ho, B. Shrader, and S. Y. Chung, "Network coding," *Journal of Communications and Networks*, vol. 10, no. 4, pp. 367–370, 2008.

[159] B. Sun, C. Gui, and H. Chen, "Network coding-based energy balancing for cooperative multipath routing in MANETs," in *2015 IEEE International Conference on Progress in Informatics and Computing (PIC)*, 2015, pp. 419–422.

Chapter 7

Conclusions and Future Work

MANETs are becoming more and more important, because of the growing number of wireless mobile electronic communication devices. In today's modern digital ubiquitous computing arena, the mobile devices have the ability to provide real time multimedia applications, such as digital video and audio streaming, which demand stringent QoS. However, the provision of this type of service with high quality is especially difficult in MANETs due to highly dynamic topology, limited resource availability, and energy constraints. In this thesis, efforts are made to provide QoS to improve the performance of MANETs while handling both the real time and best effort traffic.

7.1 Conclusions

In chapter 3, an adaptive and distributed congestion-aware cross-layer interaction protocol called Cross Layer Best effort QoS aware routing (CLBQ) protocol is presented. The protocol consists of rate adaptation and congestion aware optimization to improve the overall performance in terms of throughput, packet delivery, and end to end delay. The protocol implements the cross layer interaction between PHY, MAC and network layer to utilize multi-rate supports from 802.11b. The CLBQ protocol discovers multiple, congestion aware, node-disjoint routes that satisfy QoS requirements in terms of delay and chooses the links with higher data rate to achieve higher throughput while sending packets. The protocol minimizes the routing overhead by discovering multiple routes during the initial route discovery phase, which requires flooding of a RREQ message only once. Due to multiple routes available in the routing table, backup routes are always available for continuous data transmission when the primary route is broken or congested. This reduces the rediscovery of the path when the primary path fails, thereby saving network resources, bandwidth and time. The protocol also provides a load balancing procedure to balance the data rate mismatch between the source and the destination. The load balancing procedure uses an adaptive feedback mechanism that considers active queue load of a node as a congestion metric and balances the queue load during congestion by varying the data rate. This results in fewer data packet loss and reduces retransmission attempts. A detailed simulation study has been carried out to evaluate the performance of the CLBQ protocol and the results obtained are compared with the CRP protocol. Simulation results show that the CLBQ protocol significantly reduces both the packet drop ratio and the end-to-end delay without much impact on control overhead.

In chapter 4, to reduce spurious timeout due to TCP starvation as well as to improve TCP fairness, a TCP-friendly scheme called Adaptive Best Effort Traffic Scheduler for EDCA (ABET-EDCA) is presented. The main objective of the ABET-EDCA protocol is to minimize the retransmission of the TCP segments during congestion that have already been successfully received, by prioritizing TCP ACK packets. The cross-layer architecture used in ABET-EDCA utilizes the cross layer communication between PHY, MAC and network layers to enhance the TCP throughput performance in the presence of high priority multimedia UDP traffic. The protocol prioritizes the TCP ACK packets based on the queuing delay of TCP packets at the MAC layer and reassigns their priorities. To support this, AC(1) queue is implemented using heap data structure. Based on the TCP traffic congestion, the contention window is tuned for frequent and instant channel access. The TXOP parameter is also tuned to send permissible TCP data packets along with ACK packets to reach the destination and get acknowledge instantly for successfully delivered packets. This reduces the packet overflow at the node's queue, resulting in a minimal packet drop and eventually reducing packet retransmissions. The ABET-EDCA also proposes a mechanism to detect the network event occurring at the lower layers of TCP protocol stack, like link failure and buffer overflow. This information is exploited by the network and transport layer to correctly analyze the traffic behavior. The effectiveness of the ABET-EDCA protocol has been tested by extensive simulations carried out with heterogeneous traffic like FTP and VoIP. The protocol shows the improved fairness for TCP traffic compared to IEEE 802.11e and E-TCP under varying heterogeneous traffic load and node mobility. The ABET-EDCA showed a decent improvement in performance of lower priority TCP traffic under heavy network load while maintaining the QoS requirements of the higher priority traffic.

In chapter 5, Multi Objective Cross Layer optimization (MOCLO) protocol is presented which enhances the IEEE 802.11e MAC layer to prioritize and protect the important packets within each class of traffic flow. The novelty of the protocol is that the cross-layer and admission control designs are implemented such that it is able to adapt and self-configure to the dynamics of the underlying network environment. The MOCLO protocol is an on demand and congestion aware protocol, which helps to discover more stable node disjoint paths for efficient delivery of real time multimedia data. At the MAC layer, a priority scheduler for each AC is implemented to optimize the transmission of video traffic and the best effort TCP traffic by reducing the starvation of important packets. In MOCLO protocol, the traffic is split into multiple paths in order to reduce the load on the single primary path, for avoiding the congestion. MOCLO protocol exploits the multirate functionality to send the packets at higher data rates based on queue load of a node, which helps in improving the network throughput. The performance of MOCLO protocol is compared with QoS-AOMDV+EDCA protocol through extensive simulations by varying video connection time and traffic load. It is observed that MOCLO protocol exhibits superior performance in terms of video and TCP packet delay, video and TCP throughput, and frame loss percentage.

In chapter 6, Adaptive Multi QoS Cross-Layer Cooperative (AMCCR) Routing protocol is presented. AMCCR protocol is an end-to-end delay efficient, energy-aware routing protocol designed to increase the network lifetime. The protocol exploit the cross layer approach at the PHY, MAC and network layer for rate adaptation, relay selection, route discovery and resource allocations. The AMCCR uses an adaptive MAC cooperative mechanism to select the transmission mode, i.e., direct or relay mode, so that a low data rate host can be assisted by high data rate relay node. In AMCCR protocol, single relay node selection method is employed. The best relay node selection strategy is executed in distributed manner. QoS metrics used for relay selection are transmission time, channel contention metric, energy utilization factor. AMCCR is an adaptive cooperative protocol which decides whether cooperation on the link is necessary or not based on the channel state information. In case the cooperation on the link is necessary, multi QoS metric is used to determine the potential relay nodes for a cooperative transmission over each link. AMCCR protocol supports four data transmission schemes: direct, sender-relay-receiver, cooperative with AF or DF. Adaptive cooperation is supported with signal combining feature where signal received by the neighbor node is processed either by AF or DF method based on the channel conditions. AMCCR protocol also supports dual-hop half-duplex communication via selected relay node by using coding technique.

The performance of the AMCCR protocol is evaluated under two different mode i.e, cooperative routing mode and cooperative coding mode. The simulations were conducted by varying the simulation environment such as node density and traffic load. For cooperative routing mode, the performance of AMCCR is compared with CD-MAC and IEEE 802.11 DCF. The results shows that AMCCR protocol significantly improves the network throughput, average end-to-end delay and network lifetime compared to CD-MAC and IEEE 802.11 DCF in different channel conditions. For cooperative coding mode, the AMCCR protocol is compared with CODE, and IEEE 802.11 DCF protocol, The results obtained shows the substantial improvement in throughput and average end-to-end delay compared to CODE protocol.

7.2 Scope for Future Work

All of the four protocols proposed in this thesis are designed, simulated and analyzed with QoS routing design objectives in mind. It remains as a future work to analyze the design objectives whose goal is to provide QoS at other layers of TCP stack and to see how these types of design objectives would influence the optimal cross-layer design of a wireless multi-hop network. Furthermore, the implementation of all these protocols into a hardware solution is desirable for enabling a real MANETs with QoS support.

In the general, there is a trade-off between the cross-layer design complexity and performance. In the cross-layer design, the network architecture is either updated or modified and sometimes requires complete redesign and replacements. The cross-layer design creates interactions, some intended, and others unintended. Most of the cross-layer design proposed so far, the important design ideas depend on the specific network application scenario considered and the metric of interest. While there are a number of cross-layer design proposals in the literature today, there is much work to be done to finally to realize consistent QoS support irrespective of the application that the MANETs are designed for.

Another major challenge while incorporating QoS in MANETs is the unreliable wireless channel. However, It is found that the majority of QoS routing protocol evaluation studies assume a perfect physical channel, ignoring the effects of shadowing and multipath fading. Therefore, studying the impact of a more realistic physical layer model on QoS routing protocol performance is another interesting area of future work.

For the efficient use of resources such as bandwidth or time slot required for relay transmissions, full-duplex communication is expedient. AMCCR protocol is developed under the constraint of half-duplex communication can be extended to support full duplex communication.

Cooperative communication is an effective method of improving throughput, range and reliability. Cooperative communication has great potential to improve network security against eavesdropping attack. The optimal relay selection developed in this thesis can be used to assist the transmission between source and destination against the eavesdropping attack.

Abbreviations

ABET-EDCA	**A**daptive **B**est **E**ffort **T**raffic scheduler for IEEE 802.11e **EDCA**
AC	**A**ccess **C**ategory
ACK	**ACK**nowledgment
AF	**A**mplify **a**nd **F**orward
AIFS	**A**rbitration **I**nter **F**rame **S**pace
AMCCR	**A**daptive **M**ulti QoS **C**ross layer **C**ooperative **R**outing in MANETs
AODV	**A**d hoc **O**n-demand **D**istance **V**ector
AOMDV	**A**d hoc **O**n-demand **M**ultipath **D**istance **V**ector
AP	**A**ccess **P**oint
BE	**B**est **E**ffort
BER	**B**it **E**rror **R**ate
CD-MAC	**C**ooperative **D**iversity -**MAC**
CEDAR	**C**ore **E**xtraction **D**istributed **A**d hoc **R**outing
CRP	**C**ongestion adaptive **R**outing **P**rotocol
CLBQ	**C**ross **L**ayer **B**est effort **Q**oS aware routing protocol
CW	**C**ontention **W**indow
CTS	**C**lear **T**o **S**end
DCF	**D**istributed **C**oordination **F**unction
DSR	**D**ynamic **S**ource **R**outing
DF	**D**ecode **a**nd **F**orward
DIFS	**D**CF **I**nter **F**rame **S**pace
E-TCP	**E**nhanced **TCP**
EDCA	**E**nhanced **D**istributed **C**hannel **A**ccess
EDCF	**E**nhanced **D**istributed **C**ontrol **F**unction

FQMM	**F**lexible **Q**oS **M**odel **M**ANETs
FIFO	**F**irst **I**n **F**irst **O**ut
MAC	**M**edium **A**ccess **C**ontrol
MANET	**M**obile **A**d-hoc **N**etwork
MOCLO	**M**ulti **O**bjective **C**ross **L**ayer **O**ptimization
NAV	**N**etwork **A**llocation **V**ector
NS	**N**etwork **S**imulator
OSI	**O**pen **S**ystem **I**nterconnection
PHY	**P**hysical
QoS	**Q**uality **o**f **S**ervice
RE	**R**elay **E**ligible
RREP	**R**oute **REP**ly
RREQ	**R**oute **REQ**uest
RERR	**R**oute **ERR**ort
SINR	**S**ignal to **I**nterference **N**oise **R**atio
SIFS	**S**hort **I**nter **F**rame **S**pace
RTS	**R**equest **T**o **S**end
RTT	**R**ound **T**trip **T**ime
TCP	**T**ransmission **C**ontrol **P**rotocol
TXOP	**T**ransmit **O**pportunity
UDP	**U**ser **D**atagram **P**rotocol
VoIP	**V**oice **o**ver **IP**
WLAN	**W**ireless **L**ocal **A**rea **N**etwork

Brief Biography of the Candidate

Mahadev A. Gawas, is currently serving as a Lecturer in the Department of Computer Science & Information Systems, BITS Pilani K K Birla Goa Campus, Goa, India. He received his Bachelor's degree in Computer Engineering from Rayeshwar Institute Engineering and Information Technology, Goa University 2005. He did his Masters degree in Information Technology from Padre conceicao College of Engineering, under Goa University, Goa in 2007. He is currently pursuing Ph.D from BITS-Pilani, K K Birla Goa Campus, Goa. His research interests include wireless communications, multimedia communications, cross layer architecture.

Brief Biography of the Co-Supervisor

Prof. K. R. Anupama is Professor in the Department of EEE at BITS-Pilani, K K Birla Goa Campus, Goa, India. She received her Ph.D. from BITS-Pilani, Pilani Campus, Rajasthan, in 2005, on the topic of Routing Protocols for Mobile Ad Hoc Networks. She had research fellowship from Nokia Research Centre, Boston for a period of 4 years (2001-2005). She has developed a series of protocols using predictive source routing for unicast, multicast, geocast and inter-zone routing. She was Co-Investigator for 'Intelligent Water Resource Management Project at BITS, Pilani -Pilani Campus', funded by Microsoft Research (2005-2008) and Principal Investigator for 'Wireless Sensor Networks for Inaccessible Terrains', funded by UGC Major Research Project (2010-2012). She is currently Principal Investigator for 'Design and Development of Condition Based Monitoring of Pipelines using Wireless Sensor Networks' sponsored by GAIL India. She is currently guiding six PhD students in the area of Embedded Systems and Deeply Embedded Systems.

Brief Biography of the Supervisor

Dr. Lucy J. Gudino received her B.E. degree in Electronics and Communications Engineering from Kuvempu University, in 1994, and M.Tech. and Ph.D. degrees in Computer Science from Visvesvaraya Technological University, Karnataka, India, in 2003 and 2010 respectively. She is currently working as Assistant Professor in the Department of Computer Science and Information Systems at BITS-Pilani, K K Birla Goa campus, Goa, India. She has over 20 years of teaching experience in the field of Electronics and Communications, and Computer Science. Her research interests are Mobile communications, Adaptive Signal Processing, Adaptive arrays for cellular base stations, Wireless Networks (Wireless sensor networks and adhoc networks) and Digital filter design and its application in Speech Processing. Currently, she is guiding 7 PhD students in the area of Wireless Ad hoc networks and antenna array processing. She is also Co-Principal Investigator for the project “Design and Development of Condition Based Monitoring of Pipelines using Wireless Sensor Networks’ sponsored by GAIL India.

Declaration

I, GAWAS MAHADEV ANANT, declare that this thesis titled, 'Cross Layer Design Approaches for the Enhancement of QoS in MANETs' submitted by me under the supervision of Dr. Lucy J. Gudino and Prof. K. R. Anupama is a bonafide research work. I also declare that it has not been submitted previously in part or in full to this University or any other University or Institution for award of any degree.

Signature of the student :

Name of the student : **GAWAS MAHADEV ANANT**

ID number of the student : **2010PHXF0020G**

Date : 14/03/2018.

List of Publications

Internation Journals

1. Gawas Mahadev A., Lucy J. Gudino, K.R. Anupama and Joseph Rodrigues, ”A Cross-Layer Aware Node Disjoint Multipath Routing Algorithm for Mobile Ad Hoc Networks”, International Journal of Wireless and Mobile Networks (IJWMN) Vol. 6, No. 3, June 2014.
2. Mahadev A. Gawas, Lucy J. Gudino, and K. R. Anupama, “AMCCR: Adaptive Multi-QoS Cross-Layer Cooperative Routing in Ad Hoc Networks,” Journal of Computer Networks and Communications, vol. 2017, Article ID 3638920, 18 pages, 2017.

International Conferences

1. M. A. Gawas, L. J. Gudino and K. R. Anupama, ”Cross layer congestion aware multi rate multi path routing protocol for ad hoc network,” 2015 International Conference on Signal Processing and Communication (ICSC), Noida, 2015, pp. 88-93.
2. M. A. Gawas, L. J. Gudino and K. R. Anupama, ”Cross Layer Best effort QoS aware routing protocol for Ad Hoc Network,” 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Kochi, 2015, pp. 999-1005.
3. M. A. Gawas, L. J. Gudino and K. R. Anupama, ”Cross layer multi QoS metric routing for multimedia traffic in 802.11E over MANETs,” 2016 Eighth

International Conference on Ubiquitous and Future Networks (ICUFN), Vienna, 2016, pp. 582-587.

4. M. A. Gawas, L. J. Gudino and K. R. Anupama, ”Adaptive Cross Layered Cooperative Routing Model in Mobile Ad Hoc Networks”, European Conference on Networks and Communications (EUCNC), Athens, Greece, 2016.

5. M. A. Gawas, L. J. Gudino and K. R. Anupama, ”Cross layer adaptive congestion control for best-effort traffic of IEEE 802.11e in mobile ad hoc networks,” 2016 10th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP), Prague, czech republic, 2016, pp. 1-6.

6. M. A. Gawas, L. J. Gudino and K. R. Anupama, ”Cross layered adaptive cooperative routing mode in mobile ad hoc networks,” 2016 22nd Asia-Pacific Conference on Communications (APCC), Yogyakarta, Indonesia, 2016, pp. 462-469.

7. M. A. Gawas, L. J. Gudino and K. R. Anupama, ”Cross layer approach for effective multi hop broadcast in VANET,” 2017 9th International Conference on Communication Systems and Networks (COMSNETS), Bangalore, 2017, pp. 403-404.

8. M. A. Gawas, P. Hurkat, V. Goyal and L. J. Gudino, ”Cross layer approach for efficient dissemination of emergency messages in VANETs,” 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN), Milan, 2017, pp. 206-211.

9. M. A. Gawas, K. Modi, P. Hurkat and L. J. Gudino, ”QoS based multipath routing in MANET: A cross layer approach,” 2017 International Conference on Communication and Signal Processing (ICCSP), Chennai, 2017, pp. 1806-1812.

Ph.D Forum

1. Gawas Mahadev A., Lucy J. Gudino, K.R. Anupama, “Cross Layer QoS Aware Multi Rate Protocol for Ad hoc Network”, 2014 Eighth IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS): PhD Student Forum. , vol., no., pp.88-93, 16-18 March 2014.

2. Gawas Mahadev A., Lucy J. Gudino, K.R. Anupama ,“ Cross Layer Adaptive TCP Congestion Control in IEEE 802.11e QoS MAC in Mobile Ad Hoc Network”, 2015 Ninth IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS): PhD Student Forum. 15-18 December 2015.